SURVIVAL-BIOEQUIVALENCE: EMPIRICAL BAYESIAN APPROACH WITH NON-CONVENTIONAL INFERENCE FRAMEWORK

Craig Jefferson, Alfred A. Bartolucci, Karan P. Singh
Department of Biostatistics, School of Public Health
The University of Alabama at Birmingham
Birmingham, Alabama, 35294-0022, U. S. A

Abstract: The approach is an empirical Bayesian methodology in a survival analysis context with an application to clinical trials. The intent of this proposed methodology is to give individuals in clinical trial equivalency research more flexibility in model selection. The methodology is an extension of Bartolucci and Singh's (1993) work to the Weibull and linear-exponential models that involves testing nonscale parameters by classical methods. For the nonscale parameter, the Thoman and Bain's (1969) method is employed for the Weibull, and a nonconventional likelihood ratio test is derived in the linear-exponential case. Pertaining to the scale parameter, the approach defines a general class of discrepancy measures for equivalency, shows specific limiting cases of the general measures, and then applies the Bayesian neighborhood null hypothesis theory to derive posterior credibility regions on the measures for both distributions.

1.   Introduction

Imagine a statistician who received a data set from a scientist. The characteristics of the data set were as follows: two treatments were applied at random to two different groups of samples, but the samples within each group were from the same homogeneous population; and the random variable of interest was survival times. Imagine also that there was ancillary information available to the statistician that the scientist had knowledge of and was pertinent to the analysis. Finally, imagine that the scientist requests that the statistician tests the hypothesis of equivalence regarding the two treatments. The above scenario describes very generally the point at which this work begins. This methodology focuses on the analysis aspect of equivalence only. It assumes that the design aspects of the study were carried out such that all parties of concern agree with the method. As a result, a formal statement is as follows:

Given a data set of two drug formulations, one as an experimental formulation and the other as the standard formulation, with the objective of demonstrating equivalence of survival times in an active control clinical trial, how does one analyze the data set assuming that the underlying distributions are either a Weibull density or a Linear-Exponential density and prior information about the parameters of these distributions are available? Furthermore, what is the correct inference?

There are several important objectives of this research. One important objective is to extend the methodology of Bartolucci and Singh (1993) to include other survival distributions such as the Weibull and Linear-Exponential. The purpose of this objective is to give researchers greater flexibility in the modeling and analysis of data.

The Bartolucci and Singh (1993) methodology is also taken a step further to derive the asymptotic joint distribution of the credibility limits so that probabilistic statements can be made about the credibility limits being a subset of the specified interval or any other interval. Finally, the approach taken in this paper differs from the Bartolucci and Singh method in terms of the calculation of the test statistic and the credibility region.

The second objective of this research is to present a technique that can be used in integrating posterior

distributions derived when using this methodology. This objective is critical because the posterior distributions derived when using this methodology are usually such that analytical methods are not applicable.

The third important objective of this research is to provide an argument for the correct inference when demonstrating equivalence is the desire of the study. The imperativeness of this objective stems from the need to correctly interpret the statistical results of an active control clinical trial.

2. Weibull and Linear-Exponential Distributions

The two-parameter Weibull probability density function may be defined (Lee 1992) as

$$f(t) = \lambda \gamma (\lambda t)^{\gamma-1} \exp(-\lambda t)^{\gamma} \qquad t, \lambda, \gamma > 0.$$
(2.1)

For the Weibull distribution with a non-scale parameter, the hypotheses are as follows:

$$H_{00}: \frac{\gamma_1}{\gamma_2} > 1 + \Delta, \quad H_{10}: 1 \le \frac{\gamma_1}{\gamma_2} \le 1 + \Delta.$$
(2.2)

The Thoman and Bains' (1969) test is used for testing the hypotheses (2.2). The two-parameter linear-exponential probability density function is given by

$$f(t) = (\lambda + \gamma t)\exp\left(-\lambda t + \frac{1}{2}\gamma t^2\right) \qquad t, \lambda, v > 0$$
(2.3)

The hypotheses for the non-linear exponential parameter of the Linear-Exponential distributions are defined as follows:

$$H'_{00}: \frac{\gamma_1}{\gamma_2} < 1 - \Delta \qquad \frac{\gamma_1}{\gamma_2} > 1 + \Delta$$

$$H'_{10}: 1 - \Delta \le \frac{\gamma_1}{\gamma_2} \le 1 + \Delta$$

The likelihood ratio test and integration techniques are developed by Jefferson (1997).

3. Applications

An analysis of data from a published clinical study was performed. The objectives of the trial were:

(1) To compare the efficacy and toxicity of the combination of idarubicin(IDR) plus cytarabine (CA) to daunorubicin (DNR) plus CA for remission induction in previously untreated acute myelogenous leukemia.

(2) To compare the combination of IDR, CA and thioguanine (TG) to DNR, CA, and TG in consolidation.

(3) To compare IDR plus CA to DNR plus CA for intensification treatment during maintenance;

Vogler et. al, (1992). With respect to the context of this paper, the analysis performed is focused on the efficacy aspect of objective 1, where IDR (treatment 2) is the experimental drug and DNR (treatment 1) is the reference therapy.

Data Set Information

The total number of observations involved in the analysis was 225. The total number of survival times recorded for IDR was 109. Pertaining to IDR, there were 104 non-censored and 5 censored survival times. When compared to DNR, the total number of survival times was 116. There were 104 non-censored and 12 censored survival times for DNR. The unit of measurement regarding survival times was in months.

Regarding the testing of the non-scale parameters, Thoman and Bain's (1969) methodology was used for the shape parameters of the Weibull distribution, and a likelihood ratio test was developed for the non-exponential linear parameter of the Linear-exponential model (Jefferson 1997). The results contained in Tables 3, 4, and 5 pertain to the testing of the non-exponential linear parameter for the Linear-Exponential model.

Concerning the testing of the scale parameters, after deriving the posterior distribution for $\eta = \lambda_1/\lambda_2$ for each model, $E(\eta \mid D)$ and $E\{[\eta - E(\eta|D)]^2|D\}$ were computed for each model. $Z_{0.05}$ and $Z_{0.025}$ are the multipliers needed for computing the 90% and 95% limits, respectively, based on asymptotic normal distribution theory. The $E(\eta|D)$ has an asymptotic normal distribution (Jefferson 1997, for the Weibull and linear-exponential models, respectively). Consequently, the 90% and 95% credibility limits were calculated and compared to a predetermined interval (0.8, 1.2). The selection of the interval (0.8, 1.2) is from the bioavailability protocol guideline produced by the FDA (1977).

The ML estimates for the shape parameters of the Weibull distribution for treatment 1 and treatment 2, respectively are $\hat{\gamma}_1 = 0.815$ and $\hat{\gamma}_2 = 0.91$. Employing the test given by Thomas and Bain (1969),

$$\hat{k} = 1.11656, \text{ where } \hat{k} = \underline{\gamma}_1/\underline{\gamma}_2.$$

A researcher should reject H_{00}: non equivalence of the shape parameters and conclude that a necessary condition for equivalence has been established with respect to the shape parameter of the Weibull distribution.

Regarding testing the hypothesis $H_{01:}$ (non equivalence), concerning the scale parameter ratio $\eta = \gamma_1/\gamma_2$ the general case posterior distribution of η was derived. Credibility limits were calculated for sixteen different prior parameters combinations of the I_{γ_1} for the Weibull model at the 90% and 95% levels. The four selected values for the I_{γ_1} prior distribution parameters n_0 and t_0 were 1.7, 4, 6, 9.5 and 10, 17, 24, 33, respectively. The I_{γ_1} prior density parameters were selected for comparison purposes with the results obtained by Bartolucci and Singh (1993) and Singh (1996).

Table 1 contains the Weibull model results concerning scale parameter equivalence. All of the credibility regions (both 90% and 95%) for the Weibull model are within the predetermined interval (0.8, 1.2). As a result, a researcher would reject H_{01} and conclude that a necessary condition for equivalence has been established with respect to the scale parameters for the Weibull density. The overall inference, regarding the findings of this clinical trial with the underlying assumption of a Weibull distribution, is that a necessary condition for equivalence has been established since both parameters of the Weibull distribution are equivalent among the two treatment therapies.

Linear-Exponential Model

Table 2 contains the Linear-Exponential model results with respect to scale parameter equivalence. Each of the credibility regions are a subset of the specified interval. As a result, in each of those cases, a researcher would reject H_{01} and conclude that a necessary condition for equivalence has been established with respect to the scale parameters while assuming the linear-exponential model as the underlying distribution. The capstone inference, concerning the analysis and results of this clinical trial with the underlying distributional assumption of the linear-exponential, is that a necessary condition for

equivalence has been established since both parameters of the linear-exponential density are equivalent among the two treatment therapies.

These findings support the hypothesis of equivalence. However, the inference regarding the findings of this trial is that of a necessary condition which is: for two therapies to be equivalent in their effectiveness to the treatment of a disease, then a credibility region must be such that it is within the predetermined interval. The ML estimates for the non-exponential linear parameters were zero, indicating that a single parameter exponential distribution as a more appropriate model for the given data set. However, to illustrate the linear-exponential model, 0.01, was selected as an estimate for $\hat{\gamma}_1$ and $\hat{\gamma}_2$ and 0.0101 was selected for of the combined data sets. Continuing, the Table 3 and 4 contain the values of $\lambda_{j^*} = E(\lambda_j)$, $j = 1,2$. Table 5 contains the values of $-2 \log(k_1)$ and $-2 \log(K_2)$ where k_j are the likelihood ratio statistics,

$$K_j = L(\hat{\gamma}_1, \hat{\gamma}_2 | \lambda_{*j})/L(\hat{\gamma}, \hat{\gamma} | \lambda_{*j}), j=1,2,$$

respectively. As expected, because of the selection of $\hat{\gamma}_1$, $\hat{\gamma}_2$, $\hat{\gamma}$, H_{00} is rejected all combinations of n_0 and t_0 since $-2 \log(k_j) < 3.841$ $(X_1^2{}_{,0.05})$. Pertaining to the testing of η in the hypothesis as H_{01}:, while assuming an underlying linear-exponential density, the posterior density in the general case was derived. The values of n_0 and t_0 were the same as in the Weibull model case, and they were selected for the same reasons. The 90% and 95% credibility regions were computed for the sixteen different combinations of the prior parameter of the I_{γ_1} density.

4. **Discussion**

These findings support the hypotheses of equivalence for the scale and non-scale parameters. However, the inference regarding the findings of this trial is that of a necessary condition which is: for two therapies to be equivalent in their effectiveness to the treatment of a disease, then a credibility region must be such that it is within the predetermined interval.

The results of the data analysis by both models of this clinical trial indicates that a necessary condition has been established for equivalence using this methodology. The findings using this methodology are consistent with the findings of the classical methods. In this particular clinical trial, the experimental therapy was selected for its potential efficacy equivalence to the standard therapy and less severe side effects.

This paper demonstrates the increased flexibility of the Bartolucci and Singh (1993) methodology by extending to the Weibull and linear-exponential models. The methodology of Bartolucci and Singh (1993) also allows researchers to include their knowledge of the active agents in the compound by the selection of the prior parameters of the I_{γ_1} distribution, namely, n_0 and t_0. For detailed discussions on equivalence and bioequivalence the reader is referred to Anderson and Hauck (1983), Bartolucci and Dickey (1977), Blackwelder (1982), HSU (1983), Makuch and Johnson (1989), Metzler (1974), Munk (1993), Schuirmann (1987), Westlake (1979, 1981).

5.   References

Anderson, S. and Hauck, W. (1983). A new Procedure for Testing Equivalence in Comparison Bioavailability and Other Clinical Trials. *Communications in Statistics* - Theory and Methods 12, 2663-2692.

Bartolucci, A. A. and Dickey, J. M. (1977). Comparative Bayesian and Traditional Inference for Gamma Modeled Survival Data. *Biometrics* 3:2, 343 - 354.

Bartolucci, A.A. and Singh, K.P. (1993). Therapeutic Equivalence - A Bayesian Approach. Proceedings of the Biopharmaceutical Section of the American Statistical Association. 349 - 357.

Blackwelder, W.C. (1982). "Proving the Null Hypothesis" in Clinical Trials. *Controlled Clinical Trials* 3: 345 - 353.

Food and Drug Administration (1977). The Bioavailability Protocol Guideline for ANDA and NDA Submission. *Division of Biopharmaeutics, Drug Monographs/Bureau of Drugs, Food and Drug Admission.*

Hsu, J.P. (1983). The Assessment of Statistical Evidence from Active Control Clinical Trials. Proceedings of the Biopharam Session of the American Statistical Association Meeting in Toronto. pp 12 - 16.

Jefferson, C. (1997). An Analysis for Estimating a Necessary Condition for Equivalence in Active Control Clinical Trials, Ph.D. Dissertation, *The University of Alabama at Birmingham, USA.*

Makuch, R. and Johnson, M. (1989). Issues in Planning and Interpreting Active Control Equivalence Studies. *Journal of Clinical Epidemiology,* 503 - 571.

Metzler, C. M. (1974). Bioavailability - a Problem in Equivalence. Biometrics 30: 309-317.

Munk, A. (1993). An Improvement on Commonly Used Tests in Bioequivalence Assessment. *Biometrics 49:* 1225 - 1230.

Schuirmann, D. L. (1987). A comparison of the Two One-Sided Tests Procedure and the Power Approach for Assessing the Equivalence of Average Bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics 15,* 657 - 680.

Singh, K. (1996). Therapeutic Equivalence Using A Rich Family of Prior Distributions in Lifetime Data: Models in Reliability and Survival Analysis, N. P. Jewell, A. C. Kimber, M-L. T. Lee and G. A. Whitemire (eds), Netherlands: *Kluwer Academic Publishers,* pp 315 - 322.

Thoman, D. R. and Bain, L. J. (1969). Two Sample Tests in the Weibull Distribution. *Technometrics:* 11, 4, 805 - 815.

Vogler, W. R. et.al. (1992). A Phase III Trial Comparing Idarubicin and Daunorbicin in Combination with Cytarabine in Acute Myelogenous Leukemia: A Southeastern Cancer Group Study. *Journal of Clinical Oncology:* 1103 - 1111.

Westlake, W. J. (1979). Statistical Aspects of Comparative Bioavailability Trials. *Biometrics 35:* 273 - 280.

Westlake, W. J. (1981). Response to T. B. L. Kirkwood: Bioequivalence Testing - a Need to Rethink. *Biometrics 37:* 589-594.

TABLE 1
WEIBULL MODEL RESULTS FOR $\hat{\mu}_n$
(a = 90% Credibility Region, b = 95% Credibility Region)

n_0

t_0		1.7	4	6	9.5
10		1.00365	0.99271	0.98105	1.03805
	a	(0.8936, 1.1137)	(0.8834, 1.1020)	(0.8733, 1.0888)	(0.9265, 1.1496)
	b	(0.8726, 1.1347)	(0.8625, 1.1230)	(0.8526, 1.1095)	(0.9052, 1.1710)
17		0.99920	1.00607	0.9952	1.02424
	a	(0.8900, 1.1084)	(0.8959,1.1162)	(0.8896, 1.1095)	(0.9141,1.1344)
	b	(0.8691, 1.1293)	(0.8748, 1.1373)	(0.8685, 1.1305)	(0.8930, 1.1555)
24		0.98335	0.98258	0.98747	1.00998
	a	(0.8740, 1.0927)	(0.8740, 1.0912)	(0.8779, 1.0971)	(0.9002, 1.1198)
	b	(0.8530, 1.1137)	(0.8532, 1.1120)	(0.8569, 1.1180)	(0.8792, 1.1408)
33		1.0037	0.98778	1.01231	0.99385
	a	(0.8905, 1.1102)	(0.8774, 1.0982)	(0.9033, 1.1214)	(0.8849, 1.1028)
	b	(0.8695, 1.1313)	(0.8562, 1.1193)	(0.8824, 1.1422)	(0.8640, 1.1237)

TABLE 2
LINEAR-EXPONENTIAL MODEL RESULTS FOR $\hat{\mu}_n$
(a = 90% Credibility Region, b = 95% Credibility Region)

n_0

t_0		1.7	4	6	9.5
10		1.01586	0.99635	0.93210	1.00906
	a	(0.9035, 1.1283)	(0.8876, 1.1051)	(0.8242, 1.0400)	(0.8971, 1.1210)
	b	(0.8819, 1.1498)	(0.8668 1.1259)	(0.8035, 1.0607)	(0.8757, 1.1424)
17		1.03510	1.01845	0.99202	1.03438
	a	(0.9257, 1.1445)	(0.9070, 1.1299)	(0.8815, 1.1026)	(0.9247, 1.1440)
	b	(0.9048, 1.1654)	(0.8857, 1.1512)	(0.8603, 1.1237)	(0.9037, 1.1650)
24		0.97751	0.96017	1.01327	1.02109
	a	(0.8687, 1.0863)	(0.8501, 1.0702)	(0.9003, 1.1263)	(0.9119, 1.1303)
	b	(0.8479, 1.1072)	(0.8291, 1.0913)	(0.8786, 1.1479)	(0.8910, 1.1512)
33		1.05063	0.97425	0.99604	1.01763
	a	(0.9397, 1.1615)	(0.8658, 1.0827)	(0.8865, 1.1056)	(0.9097, 1.1256)
	b	(0.9185, 1.1828)	(0.8451, 1.1035)	(0.8656, 1.1265)	(0.8890, 1.1463)

TABLE 3
LINEAR-EXPONENTIAL RESULTS FOR TREATMENT 1,
$\lambda^*_1 = E(\lambda_1)$

	n_0			
t_0	1.7	4	6	9.5
10	0.1010	0.0995	0.0990	0.0977
17	0.1261	0.1227	0.1267	0.1234
24	0.1448	0.1445	0.1421	0.1419
33	0.1687	0.1667	0.1672	0.1659

TABLE 4
LINEAR-EXPONENTIAL RESULTS FOR TREATMENT 2,
$\lambda^*_2 = E(\lambda_2)$

	n_0			
t_0	1.7	4	6	9.5
10	0.0939	0.0931	0.0923	0.0911
17	0.1181	0.1165	0.1165	0.1145
24	0.1371	0.1365	0.1362	0.1350
33	0.1591	0.1583	0.1574	0.1561

TABLE 5
LINEAR-EXPONENTIAL RESULTS FOR $a = -2 \log(k_1)$ and $b = -2 \log(k_2)$

		n_0			
t_0		1.7	4	6	9.5
10	a	1.8058	1.8163	1.8198	1.8291
	b	1.8570	1.8630	1.8690	1.8781
17	a	1.6491	1.6684	1.6457	1.6644
	b	1.6954	1.7050	1.7050	1.7173
24	a	1.5515	1.5530	1.5648	1.5658
	b	1.5952	1.5933	1.5944	1.6009
33	a	1.4447	1.4529	1.4509	1.4563
	b	1.4855	1.4890	1.4930	1.4988