Comments Concerning Borrowing with Conditional Power Prior

jdt

Spring 2023

Contents

1 Preface

This document resulted in reading [Thompson et al., 2021] concerning the use of the KS statistic as an indicator of borrowing from historical control (placebo) data. My approach considers this problem from a frequentist perspective. The method may or may not have any utility for Bayesian approaches.

The use of the placement values and the Beta GLM has been investigated by three of my former doctoral students beginning with Stanley and Tubbs (2018).

2 Overview of this document

- Issues with Dynamic Borrowing Using Conditional Power Prior
  - What am I? Bayesian, NeoBayesian or a pretender? Tubbs' interpretation
  - How much do I borrow? How and when is this determined?
  - Under what conditions should I borrow? How and when are these conditions satisfied?
- Approach Suggested in [Thompson et al., 2021]
  - Determine the power parameter based on similarity between a percentage of the current study outcome data available at an interim look and the prior outcome data.
  - Proposed using the Kolmogorov-Smirnoff (KS) statistic as a measure of similarity.
  - An advantage of the one-sided KS similarity measure is that it can differentiate between "good" similarity (i.e., current data mean is better than the prior data mean) and less desirable similarity (i.e., current data mean is worse than the prior data mean).
  - The one-sided KS similarity measure has very limited utility. It mainly differentiates based upon the signal (means) rather than the noise (variance) in the data. Tubbs' opinion
- Propose using a commonly used ML procedure to clarify similarity in the borrowing problem.
  - ROC Literature
  - Placement Values
  - Youden Index and the KS Statistic

3 Discussion of Issues

As a non-bayesian I will politely ignore some issues, such as, the models associated with using conditional power priors and the interpretation of δ and α_{max}. Instead, I will focus on [Thompson et al., 2021] section on other methods to estimate the power parameter as found in item (3) on page 415.

> We also explored using the Kolmogorov-Smirnoff (KS) statistic as a measure of similarity. The KS statistic measures the absolute maximum distance between two empirical CDFs (here, D0

and D1). We used one minus the value of the two-sided or one-sided statistic as the similarity measure. An advantage of the one-sided KS similarity measure is that it can differentiate between "good" similarity (i.e., current data mean is better than the prior data mean) and less desirable similarity (i.e., current data mean is worse than the prior data mean) which cannot be done with the two-sided measure. As such, it borrows less in the latter case than in the former case.

As an aside, I was curious why the KS statistic was chosen over some other goodness of fit tests based upon the empirical distribution function (EDF). I suspect the choice may have arisen from the KS connection to Youden's J statistic as used in the diagnostic testing literature and expanded use in classification using machine learning methods. More on this below.

Under the section on general limitations with dynamic borrowing from prior studies, I was intrigue to learn about the importance of an immeasurable concept of patient-level exchangeability and its interpretation as "The prior and current data assume the same population parameter, albeit from distributions with potentially different variances." As a non-Bayesian, I have no idea what this statement means but have interpreted it to mean that the two populations have the same location parameter with possibly different dispersion parameters. I would argue that this approach is very short-sighted as the differences in dispersion may be more problematic than differences (or similarity) of the location parameters.

I found an interesting comment regarding borrowing from prior studies

> Rietbergen (2015) points out an important issue regarding borrowing from prior studies. That is, one should not base the amount of borrowing solely on observed outcome responses because outcome responses are influenced too much by sampling variability. For example, very different responses across prior and current data sets would cause the power parameter to approach 0, but the current and prior data sets could actually be drawn from the same population, where both sets are far in opposite tails of the population. In a related way, similar responses across sets could be as a result of sampling variability in both distributions, and neither result could be close to the "true" distribution. Rietbergen's recommendations of carefully choosing the prior datasets are relevant when using any borrowing methods.

Aside from not knowing whether "data" or "Fisher" are the true enemy here, one seemly has an impossible task. An added comment includes

> In this way, our method is limited in the same way as all dynamic borrowing methods that adjust the amount of borrowing based primarily on comparing outcome data across prior and current studies. Unfortunately, pre-specifying an amount of borrowing (i.e., static borrowing) is similarly limited because there can be no adjustment (even an adjustment to discard the prior data) if correct inference is desired.

In the next section, I present some frequentist' tools for assessing similarity between the prior and current studies. And to do so in a way that does not compromise the "correct desired inference".

4 Fundamental Concepts and Models

4.1 Diagnostic Testing methods

Suppose one has a diagnostic test Y for a disease D. Then the following terms follow:

- The sensitivity of the test is the true positive rate (TPR), given by

$$TPR = \Pr[Y \; is \; positive \mid D]$$

- 1-specificity is the false positive rate (FPR), given by

$$FPR = \Pr[Y \; is \; positive \mid \bar{D}]$$

- A test Y is said to be positive when $Y \geq c$, in which case we have

$$\text{TPR}(c) = P(Y \geq c|D)$$

$$\text{FPR}(c) = P(Y \geq c|\bar{D})$$

- The ROC is given by plotting the following x-y plot

$$ROC(c) = \{\text{TPR}(c), \text{FPR}(c)\}$$

- The AUC is the area under the ROC curve.
  - AUC can be interpreted as $P(Y_D > Y_{\bar{D}})$
  - AUC is the expected value of the non-parametric Mann-Whitney U statistic.
- In the context of borrowing, $\bar{D}$ is the prior study and D is the current study. Where I have assumed that Y on D is $\geq$ than Y on $\bar{D}$ or AUC $\geq .5$

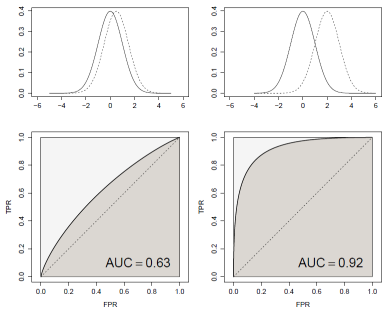
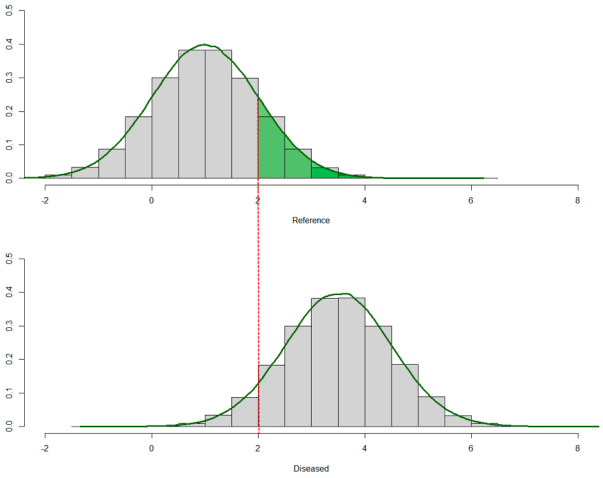


Figure 1: ROC Curve and AUC

4.1.1 Placement Values

- Placement Value of Y from the diseased group is the proportion of the non-diseased group, $\bar{D}$, with observations greater than Y: $PV = S_{\bar{D}}(Y_D)$.



- The ROC is the CDF for the placement values.

$$\begin{aligned} F_{PV}(t) &= Pr(PV_D \leq t) \\ &= Pr(S_{\bar{D},\mathbf{X}}(Y) \leq t|\mathbf{X}, D) \\ &= Pr(Y \geq S^{-1}_{\bar{D},\mathbf{X}}(t))|\mathbf{X}, D) \\ &= S_D\left(S^{-1}_{\bar{D}}(t)\right) \\ &= ROC_{\mathbf{X}}(t) \end{aligned}$$

4.1.2 Kolmogorov–Smirnov Test of Fit

- The Kolmogorov–Smirnov (KS) Test Statistic is the “maximum vertical distance" that the observed ROC is from the diagonal line.
- The KS statistic is Youden’s J statistic given by

$$J = \text{sensitivity} + \text{specificity} - 1.$$

- Provides a measure of similarity between data sets for two populations.

4.2 ROC Regression Methods

4.2.1 Covariate-Adjusted AUC Regression

- Modeling the AUC by modifying the Mann-Whitney U statistics using GLM in the presence of covariates, Z.

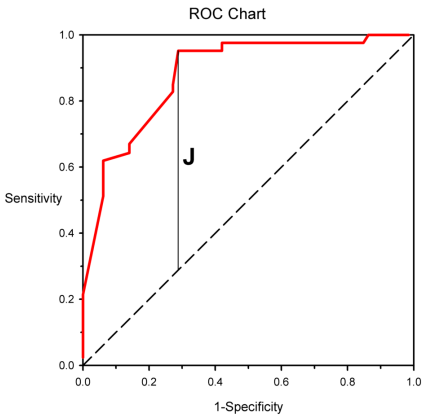


Figure 2: Youden's index.

- The Mann-Whitney (MW) U-statistic for two independent observations from two population samples, **x** and **y** is given by

$$U = \sum_{i=1}^{n} \sum_{i=1}^{m} I(x_i > y_j).$$

- AUC = $E(U \mid Z)$.

4.2.2 Beta Regression Method

- Model the placement values using beta regression [Stanley and Tubbs, 2018].
  - ([Ferrari and Cribari-Neto, 2004]) defined the Beta generalized linear model in terms of $\mu = E(Y)$ and a precision parameter $\phi = a + b$, as

$$E(Y) = \mu \text{ and } Var(Y) = \frac{\mu(1-\mu)}{1+\phi}.$$

  - The covariate adjusted beta regression model: $g(\mu_t) = \sum_{i=1}^{k} X_{ti}\beta_i = \eta_t$
  - Using the logit link, we have $\mu_t = \frac{1}{1+e^{-x_t'\beta}}$.
  - Obtain the original parameters, $\hat{a} = \frac{\hat{\phi}}{1+e^{-x_t'\hat{\beta}}}$ and $\hat{b} = \hat{\phi}\left(1 - \frac{1}{1+e^{-x_t'\hat{\beta}}}\right)$.
  - Use $\hat{a}$ and $\hat{b}$ with the CDF of this Beta distribution as the ROC.

5 Proposed Methods for Assessing Similarity for borrowing

So how does one assess the similarity between two groups when one of the groups has not been realized at the time the new study begins? What characteristics in the prior study are or are not present in the new

study? And how do these findings effect the percentage of borrowing that one can or should add to the new study data?

My approach will be to assume that the prior data characteristics are predetermined and fixed. For example, the data $Y \sim N(\mu, \sigma)$ from which I will sample (via simulation) a fixed number of observations, say n. Then for individual cases, I will sample n2 observations where $\frac{n2}{n} = r_n$ when r_n is specified. For example if $r_n = .8$ then $n2 = 80$ observations will be created for the new current study when n=100. In a similar fashion, the new data will have $N(\mu_2, \sigma_2)$ where $\frac{\mu_2}{\mu} = r_\mu$ and $\frac{\sigma_2}{\sigma} = r_\sigma$.

The parameters used in my examples were a modification of the example given in Thompson et. al. where Y is ODI. I do not try to reproduce the Bayesian approach given in their paper.

The output for my analysis are

1. Analysis with original data[1]
   (a) KDE for the two groups
   (b) Two sample tests
   (c) Goodness of fit with KS statistic
   (d) logistic regression
   (e) survival analysis
2. Analysis with Placement Values[2]
   (a) Beta Regression

5.1 Case 1

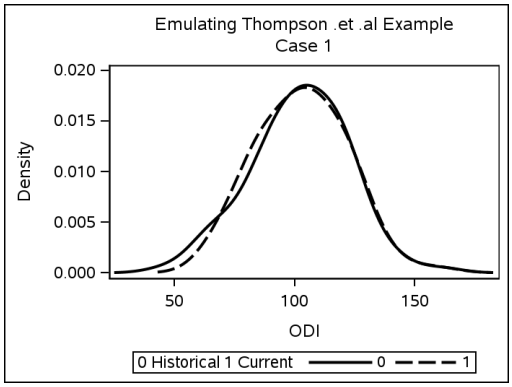
Let n=100, $\mu = 100$, $\sigma = 20$, $r_n = .8$, $r_\mu = 1$, and $r_\sigma = 1$. In which case, the two groups have the same distribution with slightly different sample sizes. One would expect that one could borrow without non-design restrictions (size and power considerations). I will annotate this case. The other cases have the same format and will likely not be annotated

5.1.1 Analysis with ODI data

Emulating Thompson .et .al Example

Case 1

[1] 2 Dimensional for Y = ODI on the two groups
[2] 1 Dimension for $PV \in [0, 1]$



The KDE for both groups. These are nearly the same, hence all the metrics should indicate that one could borrow in this case.

The TTEST Procedure

Variable: ODI

LA	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		100	102.2	20.4101	2.0410	46.5199	161.1
1		80	102.9	18.7032	2.0911	66.1654	158.5
Diff (1-2)	Pooled		−0.7546	19.6709	2.9506		
Diff (1-2)	Satterthwaite		−0.7546		2.9220		

LA	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		102.2	98.1170	106.2	20.4101	17.9202	23.7100
1		102.9	98.7592	107.1	18.7032	16.1867	22.1534
Diff (1-2)	Pooled	−0.7546	−6.5773	5.0681	19.6709	17.8224	21.9504
Diff (1-2)	Satterthwaite	−0.7546	−6.5217	5.0124			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	178	−0.26	0.7984
Satterthwaite	Unequal	174.7	−0.26	0.7965

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	99	79	1.19	0.4207

T-test assuming normal data. There is no difference in the means or variances for the two groups. I do not favor this metric due to the assumptions that are needed concerning normality may not hold in general. Although in this case they do.

Kolmogorov-Smirnov Two-Sample Statistics

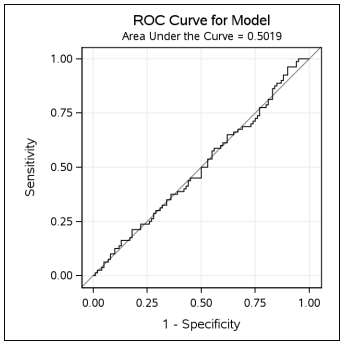
Kolmogorov-Smirnov Two-Sample Test (Asymptotic)			
KS	0.031056	D	0.062500
KSa	0.416667	Pr > KSa	0.9951

The one sided KS is .032 which favors considerable borrowing subject to the other constraints on the Bayesian model. The statistic D is simply 2*KS and KSa is the large sample adjusted statistic. Not used by me for the purposes of this problem.

Logistic Regression

Odds Ratio Estimates			
Effect	*Point Estimate*	*95% Wald Confidence Limits*	
ODI	1.002	0.987	1.017

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	50.2	*Somers' D*	0.004
Percent Discordant	49.8	*Gamma*	0.004
Percent Tied	0.0	*Tau-a*	0.002
Pairs	8000	*c*	0.502



PROC LOGISTIC produces the non-parametric estimates of the ROC and the AUC. Related to the Mann-Whitney or Wilcoxon two sample test for location. I include it for comparison with for the Beta regression model given below.

5.1.2 Placement Values with the Beta Regression

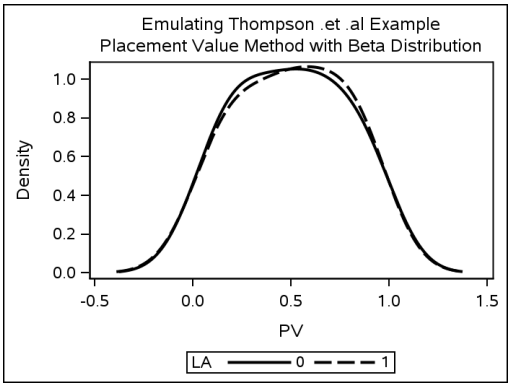
Emulating Thompson .et .al Example

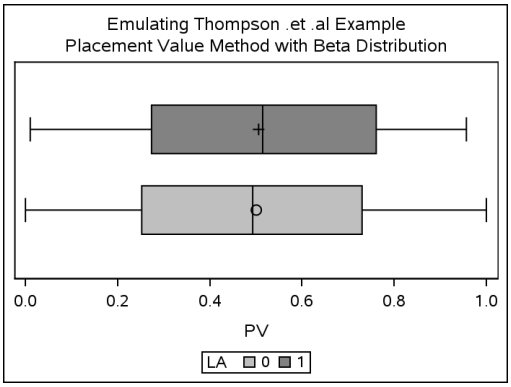
Placement Value Method with Beta Distribution

The LIFEREG Procedure

Analysis of Maximum Likelihood Parameter Estimates							
Parameter	*DF*	*Estimate*	*Standard Error*	*95% Confidence Limits*		*Chi-Square*	*Pr > ChiSq*
Intercept	1	4.7036	0.0188	4.6667	4.7405	62414.3	<.0001
Scale	1	0.1785	0.0134	0.1541	0.2068		

Analysis of Maximum Likelihood Parameter Estimates							
Parameter	*DF*	*Estimate*	*Standard Error*	*95% Confidence Limits*		*Chi-Square*	*Pr > ChiSq*
Weibull Scale	1	110.3475	2.0776	106.3498	114.4955		
Weibull Shape	1	5.6028	0.4206	4.8362	6.4909		



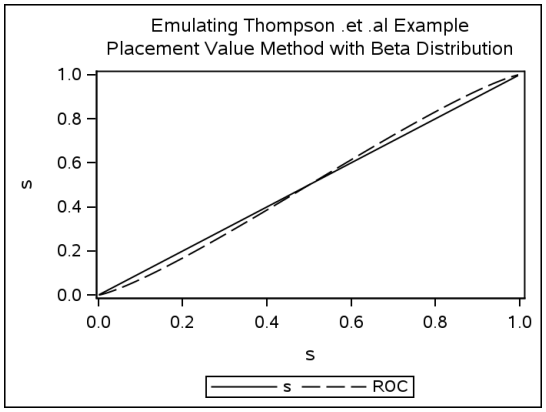


The survival function for the prior study is used as a transformation of the current study data. I present the results for both groups. Note the transformed prior data should be $U(0,1)$ whereas the transformed current study will be modeled using a Beta distribution as a generalized linear model. In this case, the transformed data is very similar to the $U(0,1)$.

The GLIMMIX Procedure

Parameter Estimates					
Effect	*Estimate*	*Standard Error*	*DF*	*t Value*	*Pr > |t|*
Intercept	0.000053	0.1156	79	0.00	0.9996
Scale	2.4976	0.3394	.	.	.

Variable	*Mean*	*Variable*	*Maximum*
auc	0.4999868	youden	0.0320753



The results are nearly identical to what was found using data for both groups. In general, the ROC curve using the placement values is smoother and the Youden index is slightly smaller than the KS in the logistic model. I suspect the difference is the use of EDF (step functions) whereas the beta model uses the incomplete Beta function and the diagonal line when using the output of the beta regression model. In these case, I only specify the intercept on the right hand side of the model. One could use covariates if desired.

5.1.3 Discussion for Case 1

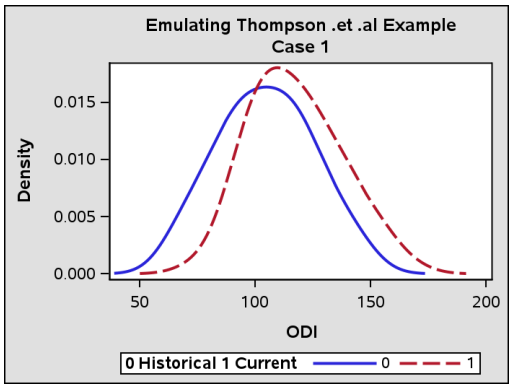
In this case, all the indicators suggest that one can borrow as needed from the prior data.

5.2 Case 2

Let n=100, $\mu = 100$, $\sigma = 20$, $r_n = .8$, $r_\mu = 1.15$, and $r_\sigma = 1$. In which case, the mean for the current study is 15% larger than the mean for the prior study. The variances for the two studies are the same.

Emulating Thompson .et .al Example

Case 2



The separation in the KDE functions suggest that the separation is too great for borrowing, yet we do not know how the KS and Youden index reflect this separation

The TTEST Procedure

Variable: ODI

LA	*Method*	*N*	*Mean*	*Std Dev*	*Std Err*	*Minimum*	*Maximum*
0		100	104.5	20.5635	2.0564	62.8093	149.8
1		80	117.4	19.7908	2.2127	73.1478	168.3
Diff (1-2)	*Pooled*		−12.8921	20.2242	3.0336		
Diff (1-2)	*Satterthwaite*		−12.8921		3.0207		

LA	*Method*	*Mean*	*95% CL Mean*		*Std Dev*	*95% CL Std Dev*	
0		104.5	100.4	108.6	20.5635	18.0549	23.8881
1		117.4	113.0	121.8	19.7908	17.1280	23.4417
Diff (1-2)	*Pooled*	−12.8921	−18.8786	−6.9056	20.2242	18.3238	22.5680
Diff (1-2)	*Satterthwaite*	−12.8921	−18.8545	−6.9297			

Method	*Variances*	*DF*	*t Value*	*Pr > |t|*
Pooled	Equal	178	−4.25	<.0001
Satterthwaite	Unequal	172.01	−4.27	<.0001

Equality of Variances				
Method	*Num DF*	*Den DF*	*F Value*	*Pr > F*
Folded F	99	79	1.08	0.7267

Kolmogorov-Smirnov Two-Sample Statistics

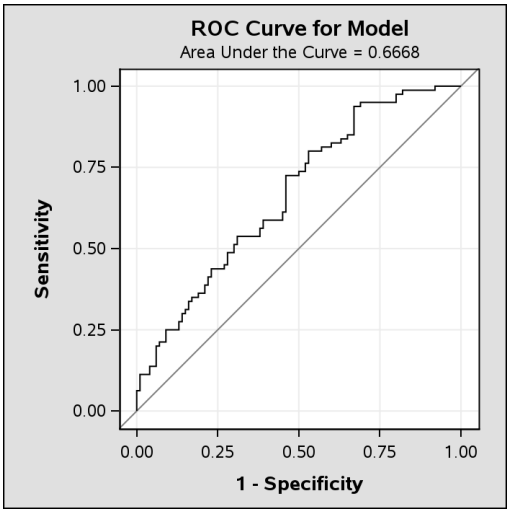
Kolmogorov-Smirnov Two-Sample Test (Asymptotic)			
KS	0.134164	*D*	0.270000
KSa	1.800000	*Pr > KSa*	0.0031

Frequentist AUC

The LOGISTIC Procedure

Odds Ratio Estimates			
Effect	*Point Estimate*	*95% Wald Confidence Limits*	
ODI	1.032	1.016	1.048

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	66.7	*Somers' D*	0.333
Percent Discordant	33.3	*Gamma*	0.334
Percent Tied	0.0	*Tau-a*	0.166
Pairs	8000	*c*	0.667

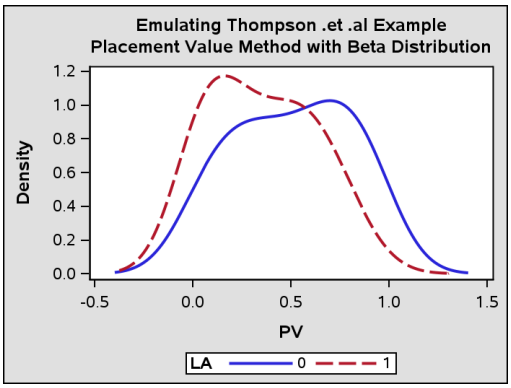


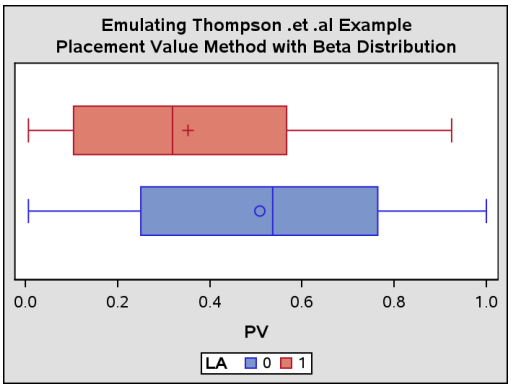
This is the classic ROC for cases where the two density functions are not similar in the sense we need for borrowing

Placement Value Method with Beta Distribution

The LIFEREG Procedure

Analysis of Maximum Likelihood Parameter Estimates							
Parameter	*DF*	*Estimate*	*Standard Error*	*95% Confidence Limits*		*Chi-Square*	*Pr > ChiSq*
Intercept	1	4.7269	0.0186	4.6905	4.7633	64797.9	<.0001
Scale	1	0.1757	0.0135	0.1511	0.2043		
Weibull Scale	1	112.9478	2.0974	108.9109	117.1343		
Weibull Shape	1	5.6921	0.4379	4.8954	6.6184		



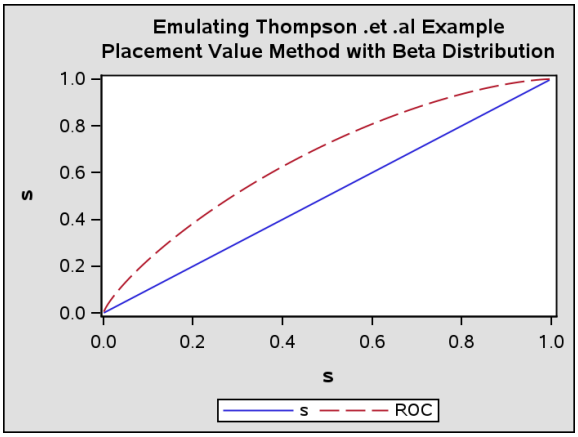


When the two density functions are separated and the location for the new study exceeds the location for the prior study, then the placement values tend to pile up close to the left endpoint, zero. In theory one could have a ZIP model but this hardly suggest that one should borrow from the prior study data.

Parameter Estimates					
Effect	*Estimate*	*Standard Error*	*DF*	*t Value*	*Pr > |t|*
Intercept	−0.6642	0.1247	79	−5.32	<.0001
Scale	2.3286	0.3255	.	.	.

Placement Value Method with Beta Distribution

Variable	*Mean*	*Variable*	*Maximum*
auc	0.6602015	youden	0.2268607



5.2.1 Discussion for Case 2

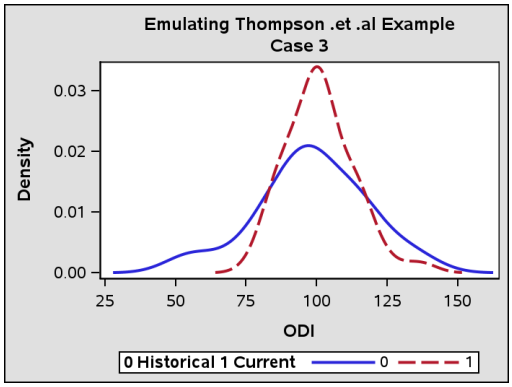
In this case, the indicators suggest that one should not borrow from the prior study as the mean shift towards the prior data could adversely effect the results of the new study. If one assumes that the active treatment group should exceed the values for the control or placebo in the current study, then the borrowing might falsely exaggerate the treatment effect in the new study.

5.3 Case 3

Let n=100, $\mu = 100$, $\sigma = 20$, $r_n = .8$, $r_\mu = 1.$, and $r_\sigma = .6$. In which case, the means for the two studies are the same and the variance for the new study is 40% smaller than the variance in the prior study.

Emulating Thompson .et .al Example

Case 3



This is an example of a case where one should not borrow because the variability in the prior study could adversely effect the final results in the current study. The large variability potentially “stretches" the location of thee current study as to fail to find a drug effect in the current study. This is seen in the following ROC curves.

The TTEST Procedure

Variable: ODI

LA	*Method*	*N*	*Mean*	*Std Dev*	*Std Err*	*Minimum*	*Maximum*
0		100	98.3249	19.5070	1.9507	47.8872	142.6
1		80	101.1	11.9509	1.3361	76.7519	138.8
Diff (1-2)	*Pooled*		−2.7385	16.5839	2.4876		
Diff (1-2)	*Satterthwaite*		−2.7385		2.3644		

LA	*Method*	*Mean*	*95% CL Mean*		*Std Dev*	*95% CL Std Dev*	
0		98.3249	94.4542	102.2	19.5070	17.1273	22.6608
1		101.1	98.4038	103.7	11.9509	10.3429	14.1555
Diff (1-2)	*Pooled*	−2.7385	−7.6475	2.1705	16.5839	15.0256	18.5058
Diff (1-2)	*Satterthwaite*	−2.7385	−7.4064	1.9294			

Method	*Variances*	*DF*	*t Value*	*Pr > |t|*
Pooled	Equal	178	−1.10	0.2724
Satterthwaite	Unequal	167.49	−1.16	0.2484

Equality of Variances				
Method	*Num DF*	*Den DF*	*F Value*	*Pr > F*
Folded F	99	79	2.66	<.0001

The locations are similar yet the variances are not. If one was just using this metric one might be inclined to borrow from the prior study.

Kolmogorov-Smirnov Two-Sample Statistics

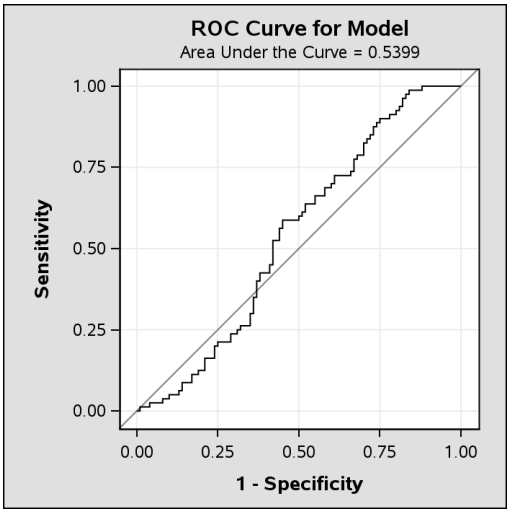
Kolmogorov-Smirnov Two-Sample Test (Asymptotic)			
KS	0.074536	*D*	0.150000
KSa	1.000000	*Pr > KSa*	0.2700

One might have some borrowing with KS = .15 but this could lead to unwanted issues. I would not borrow.

The LOGISTIC Procedure

Odds Ratio Estimates			
Effect	*Point Estimate*	*95% Wald Confidence Limits*	
ODI	1.010	0.992	1.029

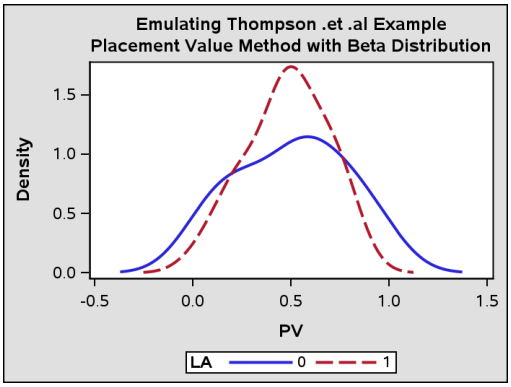
Association of Predicted Probabilities and Observed Responses			
Percent Concordant	54.0	*Somers' D*	0.080
Percent Discordant	46.0	*Gamma*	0.080
Percent Tied	0.0	*Tau-a*	0.040
Pairs	8000	*c*	0.540

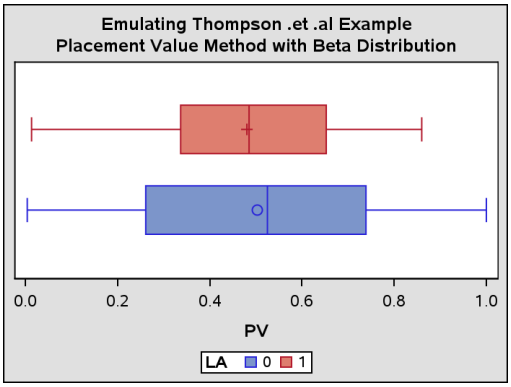


This pattern indicates that there are regions in the data where there are lack of similarity in the two studies.

The LIFEREG Procedure

Analysis of Maximum Likelihood Parameter Estimates							
Parameter	*DF*	*Estimate*	*Standard Error*	*95% Confidence Limits*		*Chi-Square*	*Pr > ChiSq*
Intercept	1	4.6647	0.0183	4.6289	4.7006	65066.0	<.0001
Scale	1	0.1735	0.0133	0.1492	0.2016		
Weibull Scale	1	106.1368	1.9410	102.3999	110.0100		
Weibull Shape	1	5.7646	0.4426	4.9593	6.7006		





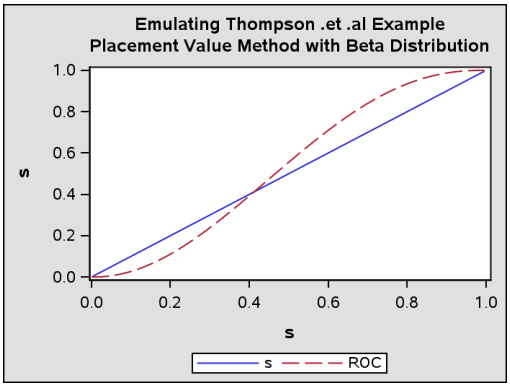
It helps to remember that the density for the prior study data is a constant on (0,1). In this case, the density for the current study placement values is very much an inverted U on (0,1) with mean close to .5.

Placement Value Method with Beta Distribution

The GLIMMIX Procedure

Parameter Estimates					
Effect	*Estimate*	*Standard Error*	*DF*	*t Value*	*Pr > |t|*
Intercept	−0.1186	0.09319	79	−1.27	0.2068
Scale	4.6386	0.6680	.	.	.

Variable	*Mean*	*Variable*	*Maximum*
auc	0.5296188	youden	0.1426434



5.3.1 Discussion for Case 3

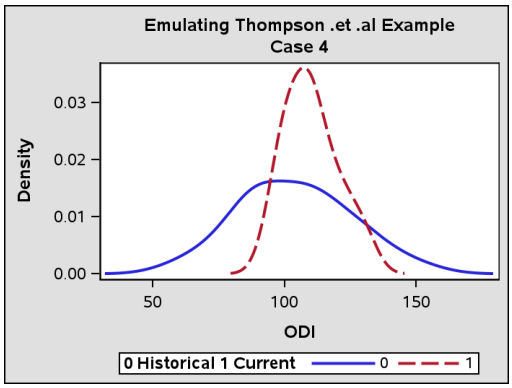
In this case, the usual indicators for the location parameter suggest that one could borrow from the prior study. However, the KS statistic (.15 with the ODI and .14 with the PV) with the ROC curves suggest that borrowing will likely increase the study variance for the control group which would likely hide any potential benefit that the active drug may have in the new study. The AUC is "close" to .5 and is misleading as there are regions in the ROC space where the two groups are somewhat dissimilar.

5.4 Case 4

Let n=100, $\mu = 100$, $\sigma = 20$, $r_n = .8$, $r_\mu = 1.1$, and $r_\sigma = .6$. In which case, the mean for the current study is 10% larger than the mean for the prior study and the variance for the new study is 40% smaller than the variance in the prior study.

Emulating Thompson .et .al Example

Case 4



The TTEST Procedure

Variable: ODI

LA	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		100	103.8	21.2464	2.1246	56.0257	155.1
1		80	110.1	10.0413	1.1227	91.9022	133.2
Diff (1-2)	*Pooled*		−6.3431	17.1993	2.5799		
Diff (1-2)	*Satterthwaite*		−6.3431		2.4030		

LA	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		103.8	99.5772	108.0	21.2464	18.6545	24.6815
1		110.1	107.9	112.4	10.0413	8.6903	11.8936
Diff (1-2)	*Pooled*	−6.3431	−11.4342	−1.2520	17.1993	15.5831	19.1924
Diff (1-2)	*Satterthwaite*	−6.3431	−11.0919	−1.5944			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	178	−2.46	0.0149
Satterthwaite	Unequal	147.58	−2.64	0.0092

Equality of Variances				
Method	*Num DF*	*Den DF*	*F Value*	*Pr > F*
Folded F	99	79	4.48	<.0001

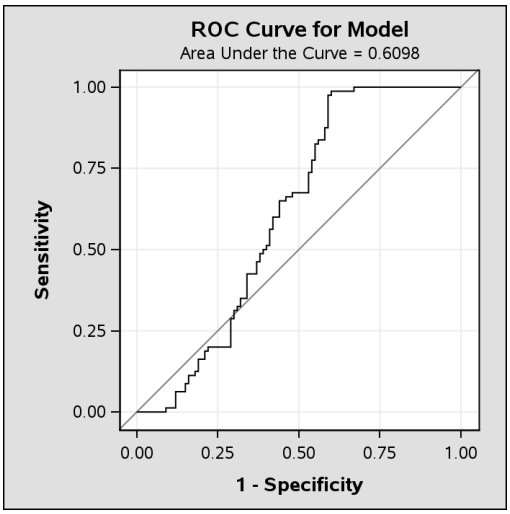
Kolmogorov-Smirnov Two-Sample Statistics

Kolmogorov-Smirnov Two-Sample Test (Asymptotic)			
KS	0.192550	*D*	0.387500
KSa	2.583333	*Pr > KSa*	<.0001

The LOGISTIC Procedure

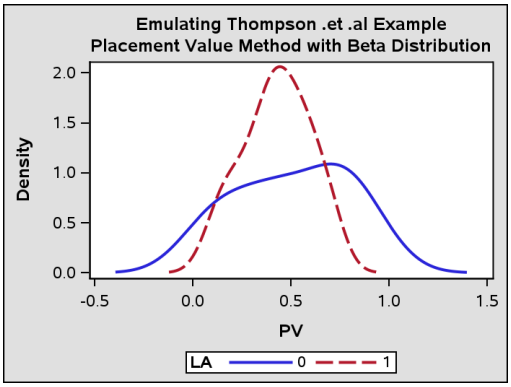
Odds Ratio Estimates			
Effect	*Point Estimate*	*95% Wald Confidence Limits*	
ODI	1.022	1.004	1.041

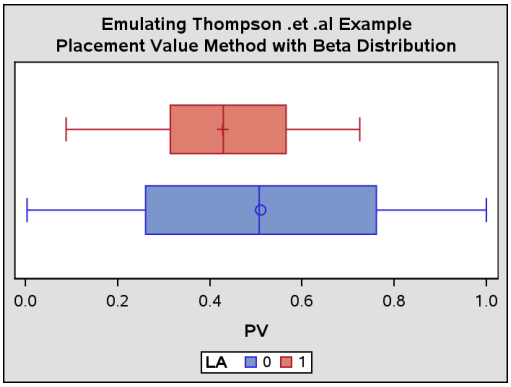
Association of Predicted Probabilities and Observed Responses			
Percent Concordant	61.0	*Somers' D*	0.220
Percent Discordant	39.0	*Gamma*	0.220
Percent Tied	0.0	*Tau-a*	0.109
Pairs	8000	*c*	0.610



The LIFEREG Procedure

Analysis of Maximum Likelihood Parameter Estimates							
Parameter	*DF*	*Estimate*	*Standard Error*	*95% Confidence Limits*		*Chi-Square*	*Pr > ChiSq*
Intercept	1	4.7224	0.0198	4.6837	4.7611	57126.5	<.0001
Scale	1	0.1868	0.0141	0.1610	0.2166		
Weibull Scale	1	112.4381	2.2216	108.1672	116.8777		
Weibull Shape	1	5.3542	0.4049	4.6167	6.2096		

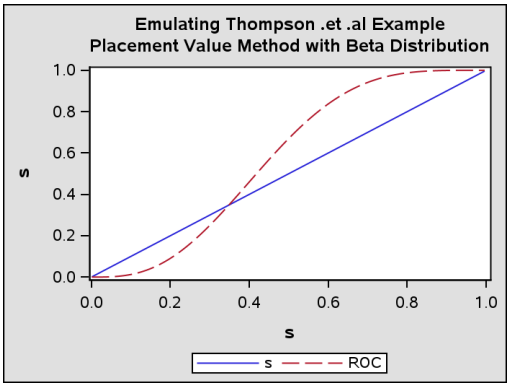




The GLIMMIX Procedure

Parameter Estimates					
Effect	*Estimate*	*Standard Error*	*DF*	*t Value*	*Pr > |t|*
Intercept	−0.3022	0.07629	79	−3.96	0.0002
Scale	7.7026	1.1486	.	.	.

Variable	*Mean*	*Variable*	*Maximum*
auc	0.5749922	youden	0.2476196



5.4.1 Discussion for Case 4

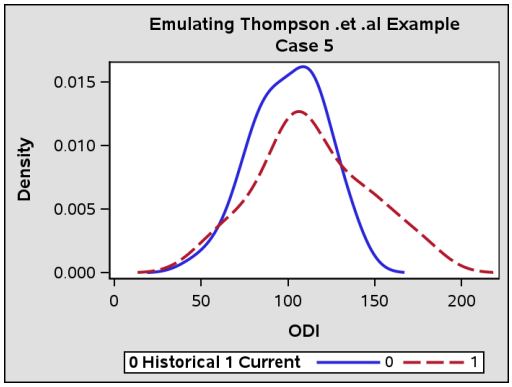
This case is similar to case 2 where one should not borrow. The large variance in the prior study does not hide the difference between the two studies when the variance for the current study is comparatively smaller. The next case considers the case where the variances are reversed, i.e. $r_\sigma = 1.6$.

5.5 Case 5

Let n=100, $\mu = 100$, $\sigma = 20$, $r_n = .8$, $r_\mu = 1.1$, and $r_\sigma = 1.6$. In which case, the mean for the current study is 10% larger than the mean for the prior study and the variance for the new study is 60% larger than the variance in the prior study.

Emulating Thompson .et .al Example

Case 5



The TTEST Procedure

Variable: ODI

LA	*Method*	*N*	*Mean*	*Std Dev*	*Std Err*	*Minimum*	*Maximum*
0		100	100.6	21.1692	2.1169	43.0657	142.9
1		80	114.2	31.8901	3.5654	48.2663	183.3
Diff (1-2)	*Pooled*		−13.6312	26.4689	3.9703		
Diff (1-2)	*Satterthwaite*		−13.6312		4.1465		

LA	*Method*	*Mean*	*95% CL Mean*		*Std Dev*	*95% CL Std Dev*	
0		100.6	96.3718	104.8	21.1692	18.5867	24.5918
1		114.2	107.1	121.3	31.8901	27.5994	37.7730
Diff (1-2)	*Pooled*	−13.6312	−21.4662	−5.7962	26.4689	23.9816	29.5362
Diff (1-2)	*Satterthwaite*	−13.6312	−21.8337	−5.4287			

Method	*Variances*	*DF*	*t Value*	*Pr > |t|*
Pooled	Equal	178	−3.43	0.0007
Satterthwaite	Unequal	131.48	−3.29	0.0013

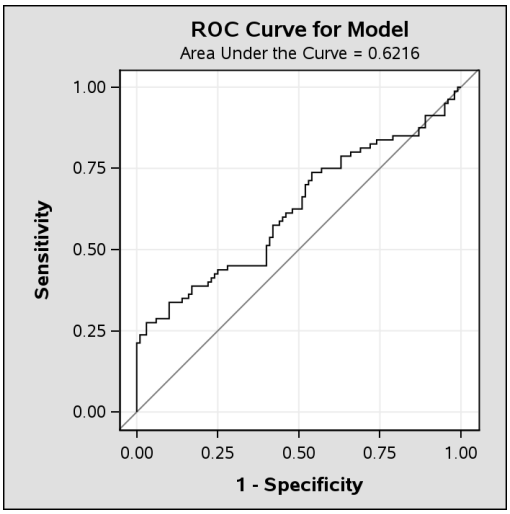
Equality of Variances				
Method	*Num DF*	*Den DF*	*F Value*	*Pr > F*
Folded F	79	99	2.27	0.0001

Kolmogorov-Smirnov Two-Sample Statistics

Kolmogorov-Smirnov Two-Sample Test (Asymptotic)			
KS	0.121741	*D*	0.245000
KSa	1.633333	*Pr > KSa*	0.0096

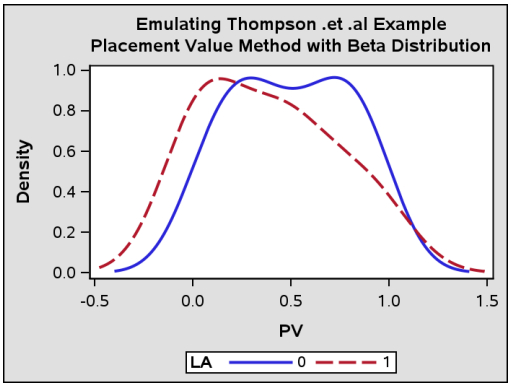
Odds Ratio Estimates			
Effect	*Point Estimate*	*95% Wald Confidence Limits*	
ODI	1.020	1.008	1.032

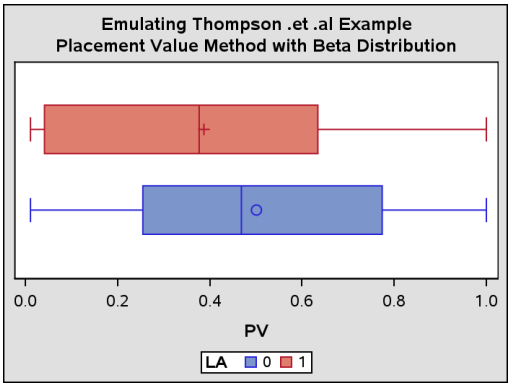
Association of Predicted Probabilities and Observed Responses			
Percent Concordant	62.2	*Somers' D*	0.243
Percent Discordant	37.8	*Gamma*	0.243
Percent Tied	0.0	*Tau-a*	0.121
Pairs	8000	*c*	0.622



The LIFEREG Procedure

Analysis of Maximum Likelihood Parameter Estimates							
Parameter	*DF*	*Estimate*	*Standard Error*	*95% Confidence Limits*		*Chi-Square*	*Pr > ChiSq*
Intercept	1	4.6911	0.0190	4.6539	4.7283	60974.6	<.0001
Scale	1	0.1804	0.0142	0.1545	0.2105		
Weibull Scale	1	108.9735	2.0702	104.9905	113.1076		
Weibull Shape	1	5.5445	0.4373	4.7504	6.4714		



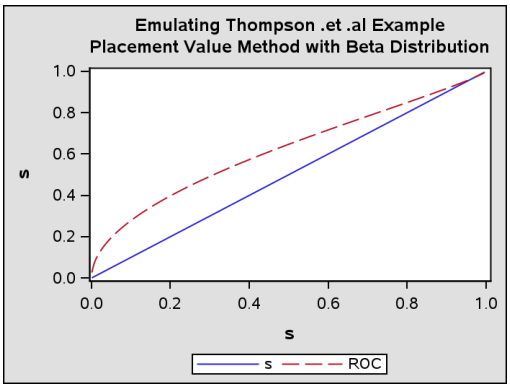


Placement Value Method with Beta Distribution

The GLIMMIX Procedure

Parameter Estimates					
Effect	*Estimate*	*Standard Error*	*DF*	*t Value*	*Pr > |t|*
Intercept	−0.4898	0.1406	78	−3.49	0.0008
Scale	1.3329	0.1759	.	.	.

Variable	Mean	Variable	Maximum
auc	0.6200679	youden	0.1977978



5.5.1 Discussion for Case 5

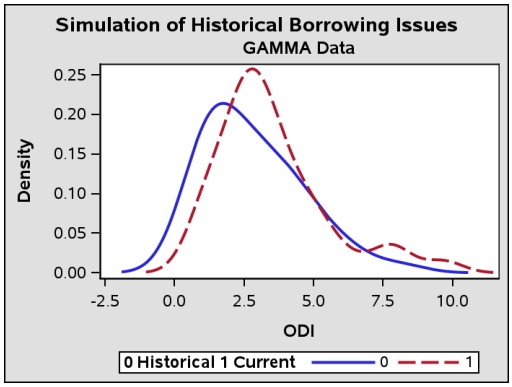
Having a large variance in the current study is seldom helpful or beneficial. In this case, the right hand tail of the current study doesn't overlap with the prior study. Borrowing would likely add observations to the left hand tail of the new study and further increase its variance. The KS statistics suggest that one would have limited borrowing.

5.6 Gamma

In this case I have replaced the normal data with data from a Gamma distribution with one parameter . Let n=100, $\lambda = 2$, $r_n = .8$, $r_\lambda = 1.1$. In which case, the mean for the current study is 10% larger than the mean for the prior study. This distribution has a longer right-hand tail.

Simulation of Historical Borrowing Issues

GAMMA Data



The TTEST Procedure

Variable: ODI

LA	*Method*	*N*	*Mean*	*Std Dev*	*Std Err*	*Minimum*	*Maximum*
0		100	2.8347	1.7868	0.1787	0.3044	8.3601
1		80	3.4906	1.9764	0.2210	0.6807	9.7579
Diff (1-2)	*Pooled*		−0.6559	1.8733	0.2810		
Diff (1-2)	*Satterthwaite*		−0.6559		0.2842		

LA	*Method*	*Mean*	*95% CL Mean*		*Std Dev*	*95% CL Std Dev*	
0		2.8347	2.4802	3.1893	1.7868	1.5688	2.0757
1		3.4906	3.0508	3.9304	1.9764	1.7105	2.3410
Diff (1-2)	*Pooled*	−0.6559	−1.2104	−0.1013	1.8733	1.6973	2.0904
Diff (1-2)	*Satterthwaite*	−0.6559	−1.2170	−0.0947			

Method	*Variances*	*DF*	*t Value*	*Pr > |t|*
Pooled	Equal	178	−2.33	0.0207
Satterthwaite	Unequal	161.12	−2.31	0.0223

Equality of Variances				
Method	*Num DF*	*Den DF*	*F Value*	*Pr > F*
Folded F	79	99	1.22	0.3402

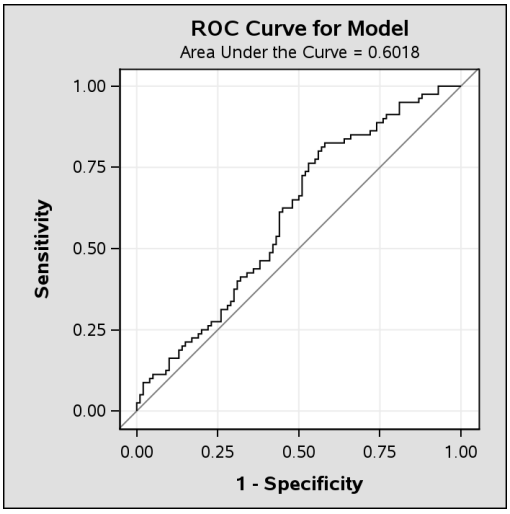
Kolmogorov-Smirnov Two-Sample Statistics

Kolmogorov-Smirnov Two-Sample Test (Asymptotic)			
KS	0.121741	*D*	0.245000
KSa	1.633333	*Pr > KSa*	0.0096

The LOGISTIC Procedure

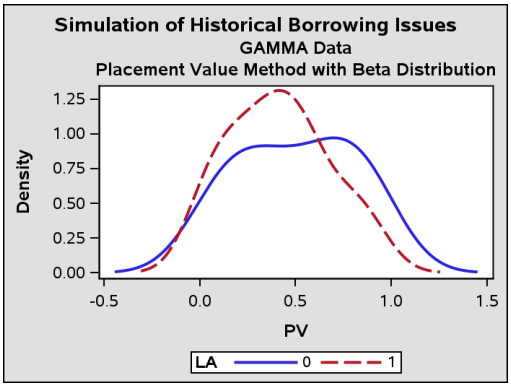
Odds Ratio Estimates			
Effect	*Point Estimate*	*95% Wald Confidence Limits*	
ODI	1.206	1.026	1.418

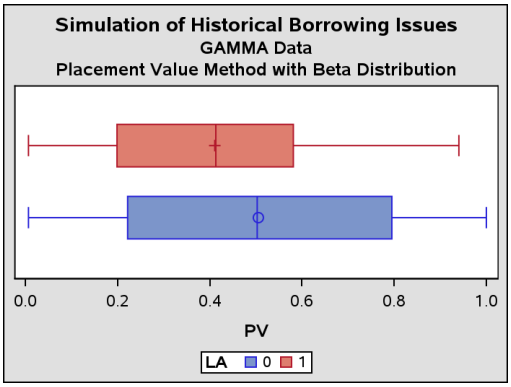
Association of Predicted Probabilities and Observed Responses			
Percent Concordant	60.2	*Somers' D*	0.204
Percent Discordant	39.8	*Gamma*	0.204
Percent Tied	0.0	*Tau-a*	0.101
Pairs	8000	*c*	0.602



The LIFEREG Procedure

Analysis of Maximum Likelihood Parameter Estimates							
Parameter	*DF*	*Estimate*	*Standard Error*	*95% Confidence Limits*		*Chi-Square*	*Pr > ChiSq*
Intercept	1	1.1575	0.0635	1.0331	1.2819	332.64	<.0001
Scale	1	0.6012	0.0468	0.5161	0.7002		
Weibull Scale	1	3.1819	0.2019	2.8097	3.6034		
Weibull Shape	1	1.6634	0.1294	1.4281	1.9374		



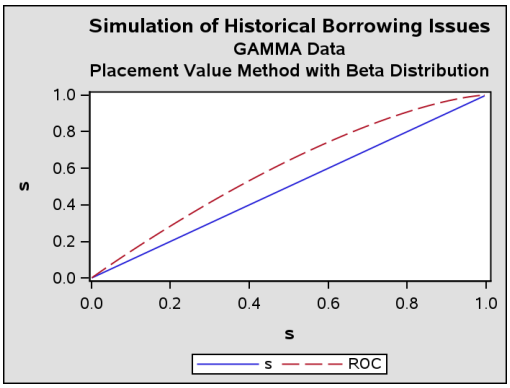


Placement Value Method with Beta Distribution

The GLIMMIX Procedure

Parameter Estimates					
Effect	*Estimate*	*Standard Error*	*DF*	*t Value*	*Pr > |t|*
Intercept	−0.3967	0.1185	79	−3.35	0.0012
Scale	2.4608	0.3378	.	.	.

Variable	Mean	Variable	Maximum
auc	0.5978970	youden	0.1441946



5.6.1 Discussion for Gamma

In this case, the indicators suggest that one should not borrow from the prior study. I included this example as it illustrates the distribution free method used in the Mann-Whitney approach to computing the ROC/AUC. Likewise, the placement value approach using the beta regression method. It should be noted that both approaches allow for covariates when computing the ROC/AUC. The placement value method allows covariates when computing the Youden index.

6 Conclusions and What is next?

These methods assume that the endpoint is a continuous random variable and that the prior study and the current study are comparable in terms of their location and dispersion parameters. Assuming that the underlying distribution is normal is a nicety but hardly realistic or necessary. Similarity as defined by the KS method given in the paper is a function of the "closeness" of the location parameters. Indeed, this closeness is a major consideration but it is far from being the only consideration. I would surmise that ignoring information about the precision or dispersion for the two groups might be a major reason that borrowing fail when completing the current study.

So if I was in control of the world, what would I do? I would do everything possible to "understand" the prior study data[3]. To the extent that if you gave me data that possibly came from a current study I would "know" the extent to which I would borrow from the prior study. This extent might mean, do not borrow. So how could I do this without compromising or offending the Bayesian paradigm and its "correct inference"?

I would suggest we take a reverse engineering approach as I did when I generated the cases in this document. Specify the rates, r_μ, r_σ, and r_n relative to the fixed prior study under consideration. In which case, one would produce a similarity metric for these cases using simulated current study data. For example suppose $\mu_0 = 100$, $\sigma_0 = 20$ and $n_0 = 250$. Let $r_\mu = 1.1$, $r_\sigma = 1$, and $r_n = .5$ then simulate data for the

[3] I mean the observed data and not the prior density function for the parameters. You may be able to work magic with the borrowing induced posterior function but I can't!

current study with $\mu_1 = 110$, $\sigma_1 = 20$ and $n_1 = 125$. Suppose we determine that one should borrow with these parameters. One could complete the Bayesian model and determine the posterior densities.[4]

James and I talked briefly and I told him that the placebo is a “TERRIBLE DRUG" and one never knows what you will get. My recommendation is “never borrow" from a prior placebo study when you have none or very limited data for the control in the new study. The placebo may induce similar location parameters but who knows what the variance or dispersion will be. The placebo has little reason to force homogeneity in the response. The very fact that people acknowledge the existence of a “placebo effect" ensures that the dispersion in the placebo group is not homogeneous.

I apologize for the randomness of these concluding comments. I think the similarity metric needs to incorporate covariates that are shared by the placebo groups of interest. This would enable one to introduce a “blocking type effect" to handle the confounding that may account for the lack of homogeneity in the response outcomes. If this effect is present, then one may need to have block or strata defined similarity metrics for borrowing within a block (strata). The Youden type metric for the logistic regression approach would work, although it is not my first choice. I would prefer to work with the placement values and the beta regression models where each ROC, hence Youden index, is a function of the covariates.

7 Reference

References

[Ferrari and Cribari-Neto, 2004] Ferrari, S. and Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 31(7):799–815.

[Stanley and Tubbs, 2018] Stanley, S. and Tubbs, J. (2018). Beta regression for modeling a covariate adjusted roc. *J. of Applied Mathematics and Statistics*, 6(4):110.

[Thompson et al., 2021] Thompson, L., Chu, J., Xu, J., Li, X., Nair, R., and Tiwari, R. (2021). Dynamic borrowing from a single prior data source using the conditional power prior. *J. OF BIOPHARMACEUTICAL STATISTICS*.

[4] Since the above analysis is based upon simulated current study data none of the Bayesian paradigm assumptions are violated for the actual study under consideration.