An Ensemble meta-prediction framework to integrate multiple regression models into a current study

Tian Gu @GuTian_TianGu
Department of Biostatistics, University of Michigan

2020 JSM August 4th

Background & Settings

Background in statistical language

Goal

External

model 1 $(Y|\mathbf{X_1}) \rightarrow \hat{\beta}_1$

model 2 $(Y|\mathbf{X_2}) \rightarrow \hat{\beta}_2$

$\vdots$

model K $(Y|\mathbf{X_K}) \rightarrow \hat{\beta}_K$

+

Internal

individual-level data (Y,**X**,B)

$\rightarrow$

Fit an expanded new model

$Y|\mathbf{X}, B \rightarrow \hat{\gamma}$

For simplicity, all the following illustration will be using the case of 2 external models

Goal

External
model 1 $(\mathrm{Y}|\mathbf{X_1}) \rightarrow \hat{\beta}_1$
model 2 $(\mathrm{Y}|\mathbf{X_2}) \rightarrow \hat{\beta}_2$

+

Internal
individual-level data (Y,**X**,B)

$\rightarrow$ Fit an expanded new model
$\mathrm{Y}|\mathbf{X}, \mathrm{B} \rightarrow \hat{\gamma}$

Proposed Framework & Estimates

Key notations

$\hat{\boldsymbol{\gamma}}_{\mathbf{I}}$ maximum likelihood estimate (MLE) using the complete internal data

$\hat{\boldsymbol{\gamma}}_{\mathbf{CML}}$ the constrained ML (CML) estimate

$\hat{\boldsymbol{\gamma}}_{\mathbf{EB}}$ the empirical Bayes (EB) estimate

2-Step Framework

Proposed Framework

Step 1:

$$\text{External } \hat{\beta}_1 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^1} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^1} = \hat{W}_1\hat{\gamma}_I + (I - \hat{W}_1)\hat{\gamma}_{CML^1}$$

$$\text{External } \hat{\beta}_2 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^2} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^2} = \hat{W}_2\hat{\gamma}_I + (I - \hat{W}_2)\hat{\gamma}_{CML^2}$$

Step 2:

$$\begin{bmatrix} \hat{\gamma}_{EB^1} \\ \hat{\gamma}_{EB^2} \end{bmatrix} \sim N(\gamma_0, V),$$

$$\hat{\gamma}_0 = \text{A weighted average of } \hat{\gamma}_{EB^1} \text{ and } \hat{\gamma}_{EB^2}$$

Constrained Maximum Likelihood Estimation for Model Calibration Using Summary-Level Information From External Big Data Sources

Nilanjan CHATTERJEE, Yi-Hau CHEN, Paige MAAS, and Raymond J. CARROLL

Proposed Framework

Step 1:

$$\text{External } \hat{\beta}_1 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^1} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^1} = \hat{W}_1\hat{\gamma}_I + (I - \hat{W}_1)\hat{\gamma}_{CML^1}$$

$$\text{External } \hat{\beta}_2 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^2} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^2} = \hat{W}_2\hat{\gamma}_I + (I - \hat{W}_2)\hat{\gamma}_{CML^2}$$

Step 2:

$$\begin{bmatrix} \hat{\gamma}_{EB^1} \\ \hat{\gamma}_{EB^2} \end{bmatrix} \sim N(\gamma_0, V),$$

$$\hat{\gamma}_0 = \text{A weighted average of } \hat{\gamma}_{EB^1} \text{ and } \hat{\gamma}_{EB^2}$$

- Assumptions
  1. Known form of $g_\beta(Y|X)$
  2. $f_\gamma(Y|X,B)$ is correctly-specified
  3. $F(X,B)$ is the same in internal and external studies

Untestable Transportability assumptions may lead to bias

- Construct constraints
  1. $p_i \equiv Pr(X = X_i, B = B_i)$, $\sum_{i=1}^n p_i = 1$
  2. Connect external score function to internal model

$$
\begin{aligned}
0 = E_{Y,X,B}\left(\frac{\partial l_\beta(Y|X)}{\partial\beta}\right) &= E_{Y,X,B}(U_\beta(Y|X)) \\
&= E_{X,B}\{E_{Y|X,B}(U_\beta(Y|X)|X,B)\} \\
&= \int_{X,B}\int_{Y|X,B} U_\beta(Y|X) f_\gamma(Y|X,B)\,dY\,dF(X,B) \\
&= \sum_{i=1}^n \int_{Y|X,B} U_\beta(Y|X) f_\gamma(Y|X,B)\,dY p_i
\end{aligned}
$$

- Maximize the likelihood $\prod_{i=1}^n f_\gamma(Y_i|X_i,B_i)p_i$ w.r.t γ and p_i subject to two constraints $\sum_{i=1}^n p_i = 1$ and $\sum_{i=1}^n \int U_\beta(Y|X) f_\gamma(Y|X,B)\,dY p_i = 0$
-

$$
\begin{aligned}
\hat\gamma_{CML} = \arg\max_{\gamma,p_i}\{&\prod_{i=1}^n f_\gamma(Y_i|X_i,B_i)p_i \\
&+ \lambda_1\left(\sum_{i=1}^n p_i - 1\right) \\
&+ \lambda_2 \sum_{i=1}^n \int U_\beta(Y|X) f_\gamma(Y|X,B)\,dY p_i\}
\end{aligned}
$$

Proposed Framework

Step 1:

$$\text{External } \hat\beta_1 \xrightarrow{\text{Chatterjee's}} \hat\gamma_{CML^1} \xrightarrow{\text{Estes'}} \hat\gamma_{EB^1} = \hat W_1 \hat\gamma_I + (I - \hat W_1)\hat\gamma_{CML^1}$$

$$\text{External } \hat\beta_2 \xrightarrow{\text{Chatterjee's}} \hat\gamma_{CML^2} \xrightarrow{\text{Estes'}} \hat\gamma_{EB^2} = \hat W_2 \hat\gamma_I + (I - \hat W_2)\hat\gamma_{CML^2}$$

Step 2:

$$\begin{bmatrix}\hat\gamma_{EB^1} \\ \hat\gamma_{EB^2}\end{bmatrix} \sim N(\gamma_0, V),$$

$\hat\gamma_0 =$ A weighted average of $\hat\gamma_{EB^1}$ and $\hat\gamma_{EB^2}$

Proposed Framework

$$Var([\hat{\gamma}_{CML^1}, \hat{\gamma}_{CML^2}, \hat{\gamma}_I]^T) = \begin{bmatrix} Var(\hat{\gamma}_{CML^1}) & Cov(\hat{\gamma}_{CML^1}, \hat{\gamma}_{CML^2}) & Cov(\hat{\gamma}_{CML^1}, \hat{\gamma}_I) \\ & Var(\hat{\gamma}_{CML^2}) & Cov(\hat{\gamma}_{CML^2}, \hat{\gamma}_I) \\ & & Var(\hat{\gamma}_I) \end{bmatrix}$$

We derived the closed-form covariance terms

Step 1:

$$\text{External } \hat{\beta}_1 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^1} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^1} = \hat{W}_1\hat{\gamma}_I + (I - \hat{W}_1)\hat{\gamma}_{CML^1}$$

$$\text{External } \hat{\beta}_2 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^2} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^2} = \hat{W}_2\hat{\gamma}_I + (I - \hat{W}_2)\hat{\gamma}_{CML^2}$$

Step 2:

$$\begin{bmatrix} \hat{\gamma}_{EB^1} \\ \hat{\gamma}_{EB^2} \end{bmatrix} \sim N(\gamma_0, V),$$

$\hat{\gamma}_0 =$ A weighted average of $\hat{\gamma}_{EB^1}$ and $\hat{\gamma}_{EB^2}$

Proposed Framework

Empirical Bayes Estimation and Prediction Using Summary-Level Information From External Big Data Sources Adjusting for Violations of Transportability

Jason P. Estes[1] · Bhramar Mukherjee[1] · Jeremy M. G. Taylor[1]

Step 1:

$$\text{External } \hat{\beta}_1 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^1} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^1} = \hat{W}_1\hat{\gamma}_I + (I - \hat{W}_1)\hat{\gamma}_{CML^1}$$

$$\text{External } \hat{\beta}_2 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^2} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^2} = \hat{W}_2\hat{\gamma}_I + (I - \hat{W}_2)\hat{\gamma}_{CML^2}$$

Step 2:

$$\begin{bmatrix} \hat{\gamma}_{EB^1} \\ \hat{\gamma}_{EB^2} \end{bmatrix} \sim N(\gamma_0, V),$$

$\hat{\gamma}_0 =$ A weighted average of $\hat{\gamma}_{EB^1}$ and $\hat{\gamma}_{EB^2}$

To correct the potential bias due to incomparability: the larger the difference between internal & external estimates, the more $\hat{\gamma}_{EB}$ shrinks toward $\hat{\gamma}_I$

$$\begin{cases} \gamma \sim N(\gamma_0, A) \\ \hat{\gamma}_I | \gamma \sim N(\gamma, \Sigma) \end{cases} \Rightarrow \hat{\gamma}_{posterior} = A(\Sigma + A)^{-1}\hat{\gamma}_I + \Sigma(\Sigma + A)^{-1}\gamma_0$$

$$\hat{\gamma}_{EB} = \hat{A}(\hat{\Sigma} + \hat{A})^{-1}\hat{\gamma}_I + \hat{\Sigma}(\hat{\Sigma} + \hat{A})^{-1}\hat{\gamma}_{CML} = \hat{W}\hat{\gamma}_I + (I - \hat{W})\hat{\gamma}_{CML}$$

$$\begin{cases} \hat{A} = (\hat{\gamma}_I - \hat{\gamma}_{CML})(\hat{\gamma}_I - \hat{\gamma}_{CML})^T \\ \hat{\Sigma} = \hat{Var}(\hat{\gamma}_I) \end{cases}$$

Proposed Framework

Step 1:

$$\text{External } \hat{\beta}_1 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^1} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^1} = \hat{W}_1\hat{\gamma}_I + (I - \hat{W}_1)\hat{\gamma}_{CML^1}$$

$$\text{External } \hat{\beta}_2 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^2} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^2} = \hat{W}_2\hat{\gamma}_I + (I - \hat{W}_2)\hat{\gamma}_{CML^2}$$

Step 2:

$$\begin{bmatrix} \hat{\gamma}_{EB^1} \\ \hat{\gamma}_{EB^2} \end{bmatrix} \sim N(\gamma_0, V),$$

$$\hat{\gamma}_0 = \text{A weighted average of } \hat{\gamma}_{EB^1} \text{ and } \hat{\gamma}_{EB^2}$$

$$\begin{bmatrix}\hat\gamma_0^{EB1}\\ \hat\gamma_1^{EB1}\\ \vdots\\ \hat\gamma_p^{EB1}\end{bmatrix} + \begin{bmatrix}\hat\gamma_0^{EB2}\\ \hat\gamma_1^{EB2}\\ \vdots\\ \hat\gamma_p^{EB2}\end{bmatrix} \xrightarrow{?} \hat\gamma_0$$

Target evaluation metric

$$\sum_{i=1}^{n} \widehat{Var}(\underbrace{X_i}_{1\times(p+1)} \underbrace{\hat\gamma_0}_{(p+1)\times 1})$$

Goal: combine 2 $\hat\gamma_{\mathrm{EB}}$'s to minimize the expected prediction variance, which reduce the p-dimension optimization to a scalar

Proposed Framework

Step 1:

$$\text{External } \hat\beta_1 \xrightarrow{\text{Chatterjee's}} \hat\gamma_{CML^1} \xrightarrow{\text{Estes'}} \hat\gamma_{EB^1} = \hat W_1 \hat\gamma_I + (I - \hat W_1)\hat\gamma_{CML^1}$$

$$\text{External } \hat\beta_2 \xrightarrow{\text{Chatterjee's}} \hat\gamma_{CML^2} \xrightarrow{\text{Estes'}} \hat\gamma_{EB^2} = \hat W_2 \hat\gamma_I + (I - \hat W_2)\hat\gamma_{CML^2}$$

Step 2:

$$\begin{bmatrix}\hat\gamma_{EB^1}\\ \hat\gamma_{EB^2}\end{bmatrix} \sim N(\gamma_0, V),$$

$$\hat\gamma_0 = \text{A weighted average of } \hat\gamma_{EB^1} \text{ and } \hat\gamma_{EB^2}$$

$$\begin{bmatrix} \hat{\gamma}_0^{EB1} \\ \hat{\gamma}_1^{EB1} \\ \vdots \\ \hat{\gamma}_p^{EB1} \end{bmatrix} + \begin{bmatrix} \hat{\gamma}_0^{EB2} \\ \hat{\gamma}_1^{EB2} \\ \vdots \\ \hat{\gamma}_p^{EB2} \end{bmatrix} \overset{?}{\Longrightarrow} \hat{\gamma}_0$$

Naïve approach:
Inverse variance weighed estimate,
ignoring the correlation between $\hat{\gamma}_{EB}$'s

$$V = Var\begin{bmatrix} \hat{\gamma}_{EB^1} \\ \hat{\gamma}_{EB^2} \end{bmatrix} = \cancel{\begin{bmatrix} Var(\hat{\gamma}_{EB^1}) & Cov(\hat{\gamma}_{EB^1}, \hat{\gamma}_{EB^2}) \\ & Var(\hat{\gamma}_{EB^2}) \end{bmatrix}}$$

Proposed Framework

Step 1:

$$\text{External } \hat{\beta}_1 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^1} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^1} = \hat{W}_1\hat{\gamma}_I + (I - \hat{W}_1)\hat{\gamma}_{CML^1}$$

$$\text{External } \hat{\beta}_2 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^2} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^2} = \hat{W}_2\hat{\gamma}_I + (I - \hat{W}_2)\hat{\gamma}_{CML^2}$$

Step 2:

$$\begin{bmatrix} \hat{\gamma}_{EB^1} \\ \hat{\gamma}_{EB^2} \end{bmatrix} \sim N(\gamma_0, V),$$

$$\hat{\gamma}_0 = \text{A weighted average of } \hat{\gamma}_{EB^1} \text{ and } \hat{\gamma}_{EB^2}$$

Optimal Covariance Weighted Estimate

Selective Coefficient Learner

Proposed Estimates

$$\text{Weight 1}\begin{bmatrix}\hat{\gamma}_0^{EB1}\\ \hat{\gamma}_1^{EB1}\\ \vdots \\ \hat{\gamma}_p^{EB1}\end{bmatrix} + (1-\text{weight 1})\begin{bmatrix}\hat{\gamma}_0^{EB2}\\ \hat{\gamma}_1^{EB2}\\ \vdots \\ \hat{\gamma}_p^{EB2}\end{bmatrix}$$

$$\begin{bmatrix}\text{weight 1}\ \hat{\gamma}_0^{EB1} + (1-\text{weight 1})\ \hat{\gamma}_0^{EB2}\\ \text{weight 2}\ \hat{\gamma}_1^{EB1} + (1-\text{weight 2})\ \hat{\gamma}_1^{EB2}\\ \vdots \\ \text{weight p+1}\ \hat{\gamma}_p^{EB1} + (1-\text{weight p+1})\ \hat{\gamma}_p^{EB2}\end{bmatrix}$$

Proposed Framework

Step 1:

$$\text{External } \hat{\beta}_1 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^1} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^1} = \hat{W}_1\hat{\gamma}_I + (I - \hat{W}_1)\hat{\gamma}_{CML^1}$$

$$\text{External } \hat{\beta}_2 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^2} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^2} = \hat{W}_2\hat{\gamma}_I + (I - \hat{W}_2)\hat{\gamma}_{CML^2}$$

Step 2:

$$\begin{bmatrix}\hat{\gamma}_{EB^1}\\ \hat{\gamma}_{EB^2}\end{bmatrix} \sim N(\gamma_0, V),$$

$\hat{\gamma}_0 =$ A weighted average of $\hat{\gamma}_{EB^1}$ and $\hat{\gamma}_{EB^2}$

Proposed Estimates

Optimal Covariance Weighted Estimate

$$\text{Weight } 1 \begin{bmatrix} \hat{\gamma}_0^{EB1} \\ \hat{\gamma}_1^{EB1} \\ \vdots \\ \hat{\gamma}_p^{EB1} \end{bmatrix} + (1-\text{weight } 1) \begin{bmatrix} \hat{\gamma}_0^{EB2} \\ \hat{\gamma}_1^{EB2} \\ \vdots \\ \hat{\gamma}_p^{EB2} \end{bmatrix}$$

Weighting across different $\hat{\gamma}_{EB}$'s: same weight for all coefficient within one $\hat{\gamma}_{EB}$

Selective Coefficient Learner

$$\begin{bmatrix} \text{weight } 1 \ \hat{\gamma}_0^{EB1} + (1-\text{weight } 1)\ \hat{\gamma}_0^{EB2} \\ \text{weight } 2 \ \hat{\gamma}_1^{EB1} + (1-\text{weight } 2)\ \hat{\gamma}_1^{EB2} \\ \vdots \\ \text{weight } p{+}1 \ \hat{\gamma}_p^{EB1} + (1-\text{weight } p{+}1)\ \hat{\gamma}_p^{EB2} \end{bmatrix}$$

Weighting across each coefficient: different weight for coefficients within one $\hat{\gamma}_{EB}$

Proposed Framework

Step 1:

$$\text{External } \hat{\beta}_1 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^1} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^1} = \hat{W}_1\hat{\gamma}_I + ($$

$$\text{External } \hat{\beta}_2 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^2} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^2} = \hat{W}_2\hat{\gamma}_I + (I - \hat{W}_2)\hat{\gamma}_{CML^2}$$

Step 2:

$$\begin{bmatrix} \hat{\gamma}_{EB^1} \\ \hat{\gamma}_{EB^2} \end{bmatrix} \sim N(\gamma_0, V),$$

$\hat{\gamma}_0 =$ A weighted average of $\hat{\gamma}_{EB^1}$ and $\hat{\gamma}_{EB^2}$

Proposed Estimates

Optimal Covariance Weighted Estimate

$$\hat{\gamma}_0 = Weight_1 \hat{\gamma}_{EB^1} + Weight_2 \hat{\gamma}_{EB^2}$$

$$Weight = \arg\min_{Weight} Var(X^T \hat{\gamma}_0)$$

$$s.t. \begin{cases} Weight1^T = 1 \\ Weight \in (0,1) \end{cases}$$

Selective Coefficient Learner

$$\hat{\gamma}^*_{X_j} = \sum_{k \in E_j} \frac{\hat{V}ar(\hat{\gamma}_{X_j^k})}{\sum_{k \in E_j} \hat{V}ar(\hat{\gamma}_{X_j^k})} \hat{\gamma}_{X_j^k}$$

$$\hat{\gamma}_{SC-Learner} = [\hat{\gamma}^*_{X_0}, \hat{\gamma}^*_{X_1}, \dots, \hat{\gamma}^*_{X_p}, \hat{\gamma}_B]^T$$

Proposed Framework

Step 1:

$$\text{External } \hat{\beta}_1 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^1} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^1} = \hat{W}_1 \hat{\gamma}_I + (I - \hat{W}_1) \hat{\gamma}_{CML^1}$$

$$\text{External } \hat{\beta}_2 \xrightarrow{\text{Chatterjee's}} \hat{\gamma}_{CML^2} \xrightarrow{\text{Estes'}} \hat{\gamma}_{EB^2} = \hat{W}_2 \hat{\gamma}_I + (I - \hat{W}_2) \hat{\gamma}_{CML^2}$$

Step 2:

$$\begin{bmatrix} \hat{\gamma}_{EB^1} \\ \hat{\gamma}_{EB^2} \end{bmatrix} \sim N(\gamma_0, V),$$

$$\hat{\gamma}_0 = \text{A weighted average of } \hat{\gamma}_{EB^1} \text{ and } \hat{\gamma}_{EB^2}$$

Simulation/Real Data Analysis Results and Conclusion

Simulation Setting Snapshot

Overall performance: $\overline{V}(Xi\hat{\gamma})$, SSE, Brier Score

Covariate-wise performance: variance decrease for single coefficient

		X1	X2	X3	X4	B
I: homogeneous populations with correct external models	Internal					
	External 1					
	External 2					
	External 3					

		X1	X2	X3	X4	B
II: homogeneous populations with uncertainties in external models	Internal					
	External 1 (small sample size)					
	External 2					
	External 3					

			X1	X2	X3	X4	B
III: heterogenous X distribution	**1**	Internal					
		External 1					
		External 2 (different B|X)					
		External 3					
	2	Internal					
		External 1					
		External 2					
		External 3 (different Y|X, B)					

			X1	X2	X3	X4	X5	X6	X7	X8	X9	B
IV: homogeneous populations with correct external models	**1**	Internal										
		External 1										
		External 2										
		External 3										

			X1	X2	X3	X4	B1	B2	B3	B4	B5
IV: homogeneous populations with correct external models	**2**	Internal									
		External 1									
		External 2									
		External 3									

Abbreviation: $\overline{V}(Xi\hat{\gamma})=\frac{1}{N}\sum_{i=1}^{N}\hat{V}(Xi\hat{\gamma})$, estimated mean variance of predicted log odds over observations in the validation dataset;

SSE=$\sum_{i=1}^{N}(p_{i0}-\hat{p}_i)^2$, sum of squared error over observations in the validation dataset;

Scaled Brier score= $\sum_{i=1}^{N}(Y_i-\hat{p}_i)^2/\sum_{i=1}^{N}(Y_i-\bar{Y})^2$, scaled Brier score over observations in the validation dataset

X represents the predictors that have been used at lease in one external model; B represents the predictors that are only available in the internal study; grey indicates that the external model did not use this predictor.

Summary of simulation & prostate cancer data analysis

- Incorporating 3 correct external model information
  - Overall: ↓ ~25%
  - Covariate-wise: ↓ ~30% each
- Incorporating 2 correct & 1 incorrect external model information
  - Low-weight for the wrong external model information
  - Overall: ↓ 15-25% depending on dimension
  - Covariate-wise: ↓ 15-30% each
- Regarding both overall and covariate-wise performance, ***Optimal Covariate Weighted Estimate (OCWE)*** and ***Selective Covariate Learner (SC-Learner)*** are better than
  - Direct regression using internal data only
  - EB estimate solely (one-time shrinkage)
  - Inverse variance weighted estimate
- OCWE vs. SC-Learner
  - SC-Learner has better covariate-wise performance when the dimension difference of external calculators was large (i.e. certain external calculators used more X's than other calculators)

e.g.

	X1	X2	X3	X4	X5	X6	X7	X8	X9	B
Internal										
External 1										
External 2										
External 3										

In Conclusion

- The proposed framework is flexible and robust
  - It can incorporate external models that use a slightly different set of covariates
  - It can identify the most relevant external information and diminish the influence of information that is less compatible with the internal data
  - It nicely balances the bias-variance tradeoff while preserving the most efficiency gain

- The 2 proposed estimates--Optimal Covariate Weighted Estimate (OCWE) and Selective Covariate Learner (SC-Learner) are efficient
  - More efficient than the naïve analysis of the internal data and other naïve combinations of external estimators
  - More efficient than incorporating single external information

Reference

Chatterjee, N., Chen, Y.-H., Maas, P. & Carroll, R. J. (2016). Constrained maximum likelihood estimation for model calibration using summary-level information from external big data sources. *Journal of the American Statistical Association*, 111(513), 107–117.

Estes, J.P. & Mukherjee, B. & Taylor, J.M.G. (2018). Empirical Bayes estimation and prediction using summary-Level information from external big data sources adjusting for violations of transportability. *Statistics in Biosciences*, 10(3), 568–586.

Gu, T. & Taylor, J.M.G. & Cheng, W. & Mukherjee, B. (2019). Synthetic data method to incorporate external information into a current study. *Canadian Journal of Statistics*, 47(4), 580-603.

Another approach to solve the same genre of problem by converting the external information into the synthetic data (no external model coefficient required)

Thank you!