Mitigating study power loss in disrupted clinical trials: Leveraging external data via the propensity score-integrated approach

Heng Li, Wei-Chen Chen, Nelson Lu, Changhong Song, Chenguang Wang*, Ram Tiwari, Yunling Xu, Lilly Yue

Division of Biostatistics, FDA/CDRH

* JHU

March 23, 2021

Outline

- A novel statistical procedure for leveraging external data in medical product evaluation
  - Today's focus: Propensity score-integrated approach
- Application: mitigating study power loss for a clinical study cut short due to the COVID-19 pandemic (this study will be called the "current study" henceforth)
- Concluding remarks

Basic Statistical Ingredients

- Power prior (Bayesian)
  - Borrow external subjects with discounting
- Composite likelihood (frequentist)
  - Incorporate external subjects with down-weighting
- Propensity score methodology
  - Balance subject characteristics (covariates) between two groups: current study subjects and external subjects

- Propensity score-integrated approach – Today's focus
  - Propensity score-integrated power prior – Bayesian
  - Propensity score-integrated composite likelihood – Frequentist
  - Goal: the augmentation of the current study with external data while maintaining study integrity

Bayesian Power Prior

- A power prior is constructed as

$$\pi(\boldsymbol{\theta}|D_0, \alpha) \propto [L(\boldsymbol{\theta}|D_0)]^{\alpha} \pi_0(\boldsymbol{\theta})$$

  - $\boldsymbol{\theta}$ **:** parameter of interest
  - $L(\boldsymbol{\theta}|D_0)$: likelihood of the external data
  - $\pi_0(\boldsymbol{\theta})$: initial prior distribution for $\boldsymbol{\theta}$
  - α**: *power prior parameter,*** $\mathbf{0 \le \alpha \le 1}$

- α: control how much external data to leverage
  - $\alpha = 0$: no leverage
  - $\alpha = 1$: full leverage

- Question: how and when to determine α for a prospective investigational study?

Ref. Chen, M-H and Ibrahim, J.G., (2000) Power Prior Distribution for Regression Models. Statistical Science, 15(1): 46-60

Composite Likelihood

- General form (weighted product of probability density functions):

$$L(\theta|Y) = \prod_i f(y_i|\theta)^{\lambda_i}$$

where λ_i is nonnegative weight to be chosen, and can be used to discount subject info from external data source.

- We set:
  - $\lambda_i = 1$, if the subject *i* is from the investigational study
  - $0 < \lambda_i \leq 1$, if the subject *i* is from the external data source
    - E.g. If $\lambda_i = 0.6$, 60% of this subject's info is leveraged and 40% discounted.
- Question: how and when to determine λ for a prospective investigational study?

Ref. Lindsay, BG (1988). Composite likelihood method. *Contemporary mathematics*, 80(1): 221-239.
Varin et al (2011). An overview of composite likelihood methods. *Statistica Sinica*, P5-42.

Propensity Score Methodology

- A ground-breaking statistical innovation for the *design* and *analysis* of observational studies, developed by Rosenbaum and Rubin in 1983 (Rosenbaum and Rubin, 1983).
- Propensity score (PS): Conditional probability of receiving treatment A rather than treatment B, given a collection of observed baseline covariates.
- Replace a (large) number of observed confounding covariates with one scalar function of these covariates: the propensity score.
- Goal in observational studies: Simultaneously balance many observed covariates between the two treatment groups, and then reduce bias in treatment comparison on outcomes.
- Our Goal: Simultaneously balance many observed covariates between the external subjects as one group and current study subjects as the other group, to make leveraging external subjects more justified

Propensity Score Methods

- Propensity score is estimated through statistical modeling of relationship between **the covariates** and **membership of the groups between which covariates are to be balanced.**
- Commonly used PS methods in the regulatory settings:
  - Matching on propensity scores
  - Stratification on propensity scores
  - Weighting using propensity scores
- All these methods can separate study *design* from outcome *analysis*.
- In our application, propensity score stratification is used, propensity scores are estimated by independent statistician blinded to outcome data

Propensity Score-Integrated Approach

- A new methodology for leveraging external data to augment a prospective investigational study.
  - PS-integrated power prior (PS + PP) - Bayesian
  - PS-integrated composite likelihood (PS + CL) – Frequentist
- Used to
  - augment a single-arm investigational study with external data,
  - augment one or both arms of an RCT,
  - with the option of discounting/down-weighting information from external data.
- PS -> Study design
- PP or CL -> Outcome analysis

PS-Integrated Approach – Study Design

- Define PS as the conditional probability of being in the current study vs external source, given subject baseline covariates.
- Use PS to design the study:
  - Select comparable subjects from external data source
  - Stratify the pooled current study and external subjects
  - Specify the amount by which information of external subjects is discounted/down-weighted (i.e., determine α in PP or λ in CL)
- Only baseline covariates are used in the above: Outcome free!

Leveraging External Data -- Caveats

- It is critical to assess whether leveraging external data fits for purpose for the specific objectives of the current study.
- Sufficient external data quality and integrity are essential for regulatory decision-making – *relevance and reliability.*
- Outcome-free planning is critical – *trial integrity and transparency.*
- Early consultation with relevant FDA review division is important (FDA guidance documents).

- The propensity score-integrated approach can be used
  - To augment single-arm study with external data
  - To augment one or both arms of RCT with external data

A Illustrative Example

- A single-arm clinical study
- Primary endpoint: one-year adverse event
- Parameter of interest: θ, probability of a subject experiencing adverse event(s) within the first year.
- Associated hypothesis testing:

$$H_0: \ \theta \geq 36\% \quad vs: \quad H_a: \ \theta < 36\%$$

- Sample size determination
  - Assume $\theta = 0.30$
  - Set: power = 80%; significance level = 0.05
  - Then, N = 380

A Illustrative Example (cont.)

- The enrolment was stopped at **290** due to the pandemic, and that it is not practical to reopen the enrolment at a later time.
- It was proposed to
  - borrow **90 = 380 - 290** subjects from a registry for this device in Europe (the device was approved in EU), using the PS-integrated approaches.
  - identify an independent statistician who was blinded to the outcomes data.
- Based on the subject inclusion/exclusion criteria specified in the current study, **941** subjects were selected from the registry.
- With the covariate data of **1,231** (290 + 941) subjects from the current study and registry, a propensity score model was created by the independent statistician, using logistic regression.
- Propensity score is defined as the probability of a subjects being in the current study instead of the registry conditional on all the covariates.

Table 1. Sample Sizes in PS Stratum

	1	2	3	4	5	Total
Current Study (n)	58	58	58	58	58	290
Registry (n)	281	210	154	187	109	941

- Five PS strata were formed, and balance for each covariate was checked using numerical and graphical methods.

- Note:
  - 90 external subjects were to be leveraged, but 941 identified.
  - Only partial info from each of 941 external subjects could be leveraged.
  - Partial? How much? Depending on what?

Step 1 – Split 90 nominal subjects into 5 PS strata

Step 2 – Determine power parameter α or exponent λ within each stratum

Step 1. Split 90 nominal subjects

- Split 90 nominal subjects into 5 PS strata, *proportional* to the similarity of external and the current subjects in terms of baseline covariates.

- The similarity is measured by an *overlapping coefficient*, the overlapping area of propensity score distributions of the two groups of subjects.

Overlapping Coefficients

	PS Stratum					
	1	2	3	4	5	Total
Current Study (n)	58	58	58	58	58	290
Registry (n)	281	210	154	187	109	941
Overlap Coeff	0.87	0.78	0.86	0.84	0.77	

Standardized Overlapping Coefficients

	PS Stratum					
	1	2	3	4	5	Total
Current Study (n)	58	58	58	58	58	290
Registry (n)	281	210	154	187	109	941
Overlap Coeff	0.87	0.78	0.86	0.84	0.77	
Std. Overlap Coef.	21%	19%	21%	20%	19%	100%

Splitting 90 Nominal External subjects

	PS Stratum					
	1	2	3	4	5	Total
Current Study (n)	58	58	58	58	58	290
Registry (n)	281	210	154	187	109	941
Overlap Coeff	0.87	0.78	0.86	0.84	0.77	
Std. Overlap Coef.	21%	19%	21%	20%	19%	100%
Subjects Leveraged	19	17	19	18	17	90

(90 x 21% =19)

(90 x 19% = 17)

- The number of external subjects allocated to each PS stratum is proportional to their *standardized overlapping coefficient.*

Step 2. Determining How Much Info to Leverage

- The info leveraged from each individual external subject depends on how many external subjects in that PS stratum.
- The power prior parameter for each individual external subject, α, is inversely proportional to the sample size of external subjects in the PS stratum.

	PS Stratum					
	1	2	3	4	5	Total
Current Study (n)	58	58	58	58	58	290
Registry (n)	281	210	154	187	109	941
Subjects Levd. (n)	19	17	19	18	17	90
α (or λ)	0.07	0.08	0.12	0.10	0.15	

(19/281 = 0.07)

Leveraging External Data Planning – Finished

- We know
  - The PS stratum each subject would belong to.
  - How much info each external subject could contribute.
  - Study operating characteristics: 80% power; 5% Type I error rate.

	PS Stratum					
	1	2	3	4	5	Total
Current Study (n)	58	58	58	58	58	290
Registry (n)	281	210	154	187	109	941
α (or λ)	0.07	0.08	0.12	0.10	0.15	

Outcome Analysis (Power Prior)

- After the clinical outcome was observed from all the subjects, the final analysis was conducted, based on the PS study design:

	PS Stratum					
	1	2	3	4	5	Total
Current Study (n)	58	58	58	58	58	290
Registry (n)	281	210	154	187	109	941
α	0.07	0.08	0.12	0.10	0.15	

- Apply the power prior approach within each stratum to get stratum-specific posterior distribution, which are then combined to complete the inference for the parameter of interest.

- The posterior probability of $\theta < 36\%$ is 96.9%, which meets the study success criterion.

Outcome Analysis (Composite Likelihood)

- After the clinical outcome was observed from all the subjects, the final analysis was conducted.

	PS Stratum					
	1	2	3	4	5	Total
Current Study (n)	58	58	58	58	58	290
Registry (n)	281	210	154	187	109	941
λ	0.07	0.08	0.12	0.10	0.15	

- Apply the composite likelihood approach to get stratum-specific parameter estimates, which are then combined to complete the inference for the parameter of interest.
- Maximum likelihood estimate of $\theta = 31\%$, *p*-value = 0.01.

2
2

Concluding Remarks

- Novel statistical methods play a critical role in leveraging external data to support regulatory decisions.
- Propensity score-integrated approaches can be applied to incorporate external data for a prospective investigational clinical study.
- Propensity score-integrated approach can be utilized to mitigate study power loss due to the COVID-19 pandemic.

US FDA Guidance

- US Food and Drug Administration (2020), "FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Public Health Emergency," available at https://www.fda.gov/regulatory-information/search-fda-guidance-documents/fda-guidance-conduct-clinical-trials-medical-products-during-covid-19-public-health-emergency.

- US Food and Drug Administration (2020), "Statistical Considerations for Clinical Trials During the COVID-19 Public Health Emergency Guidance for Industry," available at https://www.fda.gov/regulatory-information/search-fda-guidance-documents/statistical-considerations-clinical-trials-during-covid-19-public-health-emergency-guidance-industry.

PS-Integrated Approaches - Reference

- Wang, C., Li, H., Chen, W., Lu, N., Tiwari, R., Xu, Y., Yue, L. (2019). Propensity Score-Integrated Power Prior Approach for Incorporating Real-World Evidence in Single-Arm Clinical Studies. *Journal of Biopharmaceutical Statistics*, 29 (5),731-748.

- Wang, C., Lu, N. Chen, W., Li, H. Tiwari, R., Xu, Y., Yue, L. Propensity Score-Integrated Composite Likelihood Approach for Incorporating Real-World Evidence in Single-Arm Clinical Studies. *Journal of Biopharmaceutical Statistics*, 30(3), 2020.

- Chen, W., Wang, C., Li, H., Lu, N. Tiwari, R., Xu, Y., Yue, L. Propensity Score-Integrated Composite Likelihood Approach for Augmenting the Control Arm of a Randomized Controlled Trial by Incorporating Real-World Data. *Journal of Biopharmaceutical Statistics*, 30(3), 2020.

- Li, H., Chen, W.C., Lu, N.T., Song, C., Wang, C., Tiwari, R., Xu, Y., and Yue, L.Q. Mitigating Study Power Loss Caused by Clinical Trial Disruptions Due to the COVID-19 Pandemic: Leveraging External Data via Propensity Score-Integrated Approaches. *Statistics in Biopharmaceutical Research*, 13, (2021) DOI:10.1080/19466315.2020.1860813

Thank You!