Novel Statistical Methods for Incorporating Real-World Data:

Mitigating Study Power Loss Caused by Clinical Trial Disruptions Due to the COVID-19 Pandemic

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

Division of Biostatistics, FDA/CDRH

*: JHU

ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop

Sep. 22-25, 2020

Outline

- Novel statistical approaches in leveraging RWD in medical product evaluation
  - Today's focus: Propensity score-integrated approaches
- Application: mitigating study power loss caused by clinical trial disruptions due to the COVID-19 pandemic, by leveraging RWD.
- Concluding remarks

Novel Statistical Methods in Leveraging RWD

- Propensity score methodology
  - Identify and construct a control group
  - Form both treatment and control arms
  - Augment treatment or control arm
- Bayesian inference
  - Borrow external info. for treatment or control arm
- Composite likelihood approach
  - Down-weight patient info obtained from RWD
- Propensity score-integrated approaches – Today's focus
  - Propensity score-integrated power prior – Bayesian
  - Propensity score-integrated composite likelihood – Frequentist
  - Be used to augment an investigational study leveraging RWD, with the option of down-weighting info. from RWD.

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 the collection of confounding covariates with one scalar function of these covariates: the propensity score.
- Goal: Simultaneously balance many observed covariates between the two treatment groups, and then reduce bias in treatment comparison on outcomes.

Propensity Score Methods

- Propensity score is estimated through statistical modeling of relationship between **treatment group membership** and **covariates.**
- Commonly used PS methods in the regulatory settings:
  - Matching on propensity scores
  - Stratification on propensity scores
  - Inverse probability of treatment weighting using propensity scores
- All these methods can be used for both study design and outcome analysis, and can separate study *design* from outcome *analysis*.

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 borrow
  - $\alpha = 0$: no borrow
  - $\alpha = 1$: full borrow

- 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 Approach

- 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 patient info from external data source.

- We set:
  - $\lambda_i = 1$, if the patient *i* is from the investigational study
  - $0 < \lambda_i \leq 1$, if the patient *i* is from the external data source
    - E.g. If $\lambda_i = 0.6$, 60% of this patient's info is borrowed 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. *Statistics Sinica*, P5-42.

Propensity Score-Integrated Approaches

- A new methodology for leveraging RWD to augment a prospective investigational study, and save sample size required in 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 the control arm in an RCT,
  - with the option of down-weighting information from external data.
- PS -> Study design
- PP or CL -> Outcome analysis

PS-Integrated Approaches – Study Design

- Define PS as the conditional probability of being in the investigational study, given patient baseline covariates.
- Use PS to leverage external data source and design an investigational study:
  - Select comparable patients from external data source
  - Determine the weights used to down-weight information of external patients
- The selection of external patients and determination of weights are based on patient baseline covariates: Outcome free!

Application - Mitigating Study Power Loss

- The COVID-19 has major impacts on ongoing clinical trials, and imposed significant challenges caused by clinical trial disruptions.
- Some studies suspended enrolment.
- If restarting the enrolment is impractical, it may still be possible to rescue the disrupted study by mitigating the study power loss - How?
- Some suggestions (Meyer et al., 2020):

  1). using patients in the current study only and including
  - additional follow-up time (e.g., for time-to-event endpoints), or
  - additional repeated measurements (e.g., for longitudinal endpoints).

  2). leveraging external patients such as those extracted from
  - a historical clinical study, or
  - a RWD source.

Ref. Statistical Issues and Recommendations for Clinical Trials Conducted During the COVID-19 Pandemic (Meyer et al, 2020, *SBR*)

Mitigating Power Loss by Leveraging External Data

- 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 – *trail integrity and transparency.*
- Early consultation with relevant FDA review division is important (FDA guidance documents).
- Statistically, it is borrowing patients from an external data source to augment the current study data.
- The propensity score-integrated approaches can be used
  - Augment single-arm study
  - Augment control-arm of RCT

A Hypothetical Example

- A single-arm clinical study
- Primary endpoint: one-year adverse event
- Parameter of interest: θ, proportion of patients who experienced adverse event(s) within the one-year period.
- 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 Hypothetical 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** patients 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 patient inclusion/exclusion criteria specified in the current study, **941** patients were selected from the registry.
- With the covariate data of **1,231** (290 + 941) patients from the current study and registry, a propensity score model was created by the independent statistician, using logistic regression.
- Five PS strata were formed, and balance for each covariate was checked using numerical and graphical methods.

Table 1. Sample Size 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

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

Step 1 – Split 90 nominal patients into 5 PS strata

Step 2 – Determine how much to borrow within each PS stratum

Step 1. Split 90 nominal patients

- Split 90 nominal patients into 5 PS strata, *proportional* to the similarity of external and the current patients 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 patients.

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 Patients

	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%
Patients Borrowed	19	17	19	18	17	90

(90 x 21% =19)

(90 x 19% = 17)

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

Step 2. Determining How Much Info to Borrow

- The info borrowed from each individual external patient depends on how many external patients in that PS stratum.
- The power prior parameter for each individual external patient, α, is inversely proportional to the sample size of external patients 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
Patients Borr. (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 RWD Planning – Finished

- We know
  - The PS stratum each patient would belong to.
  - How much info each external patient 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 patients, 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 patients, 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 estimate, which are then combined to complete the inference for the parameter of interest.
- Maximum likelihood estimate of $\theta = 31\%$, *p*-value = 0.01.

2
3

Concluding Remarks

- Novel statistical methods play a critical role in leveraging RWD to support regulatory decisions.
- Propensity score-integrated approaches can be applied to incorporate RWD for a prospective investigational clinical study.
- Propensity score-integrated approaches 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.

Thank You!