Leveraging Real World Data in the Evaluation of Diagnostic Devices via the Propensity Score-Integrated Composite Likelihood Approach

Changhong Song[a], Wei-Chen Chen[a], Heng Li[a], Nelson Lu[a], Ram Tiwari[b], Chenguang Wang[c], Yunling Xu[a], Lilly Yue[a]

a: Division of Biostatistics, Center for Devices and Radiological Health, U.S. Food and Drug Administration
b: Bristol Myers Squibb
c: Division of Biostatistics and Bioinformatics, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University

Abstract

For evaluating the performance of an investigational diagnostic device, real-world data can be leveraged to augment a traditional study consisting of prospectively enrolled subjects, using recently developed propensity score-integrated composite likelihood (PSCL) approach, where a two-stage outcome-free principle is applied to separate study design and data analysis, and the propensity score methodology is used to select comparable subjects from real-world data in terms of baseline covariates. For data analysis, the PSCL approach is applied to estimate the diagnostic performance (sensitivity, specificity, PPV, and NPV). The approach will ultimately increase the validity of study design and interpretability of the study results in the diagnostic area, thereby advancing the current regulatory evaluation practice.

Method

- Two types of study designs are discussed with leveraging real world data to augment a traditional *investigational diagnostic device (IDT)* study of prospectively enrolled subjects [1].
- Apply propensity score-based methods along with *two-stage study design* (outcome free) [2] to select comparable subjects from real-world data source and separate study design and data analysis.
- Estimate sensitivity/specificity/PPV/NPV via PSCL.

Type 1 Study Design

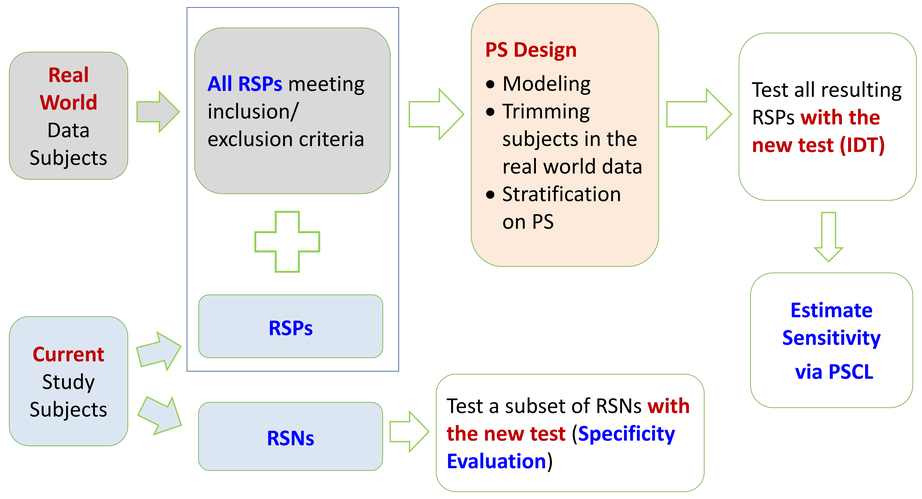
- All current study subjects are tested *with the reference standard test first,* then all of reference standard positives (RSPs) and a subset of reference standard negatives (RSNs) are tested with the IDT.
- Based on prespecified design parameters (e.g., nominal number of real world RSPs to be leveraged), use propensity score (PS) modeling, trimming, and stratification to select and discount real world RSPs.
- Estimate sensitivity via PSCL.

Type 2 Study Design

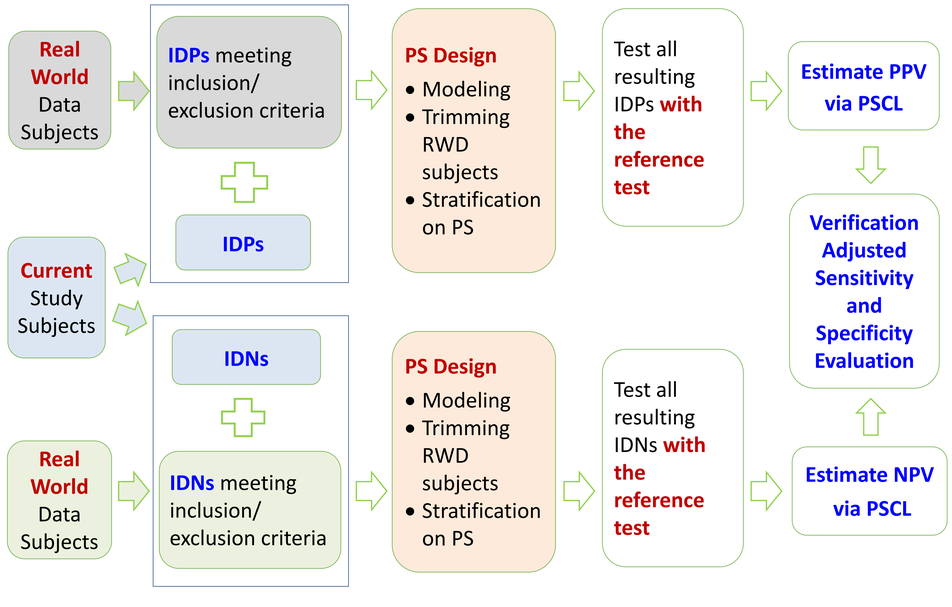
- All current study subjects are tested *with the IDT first,* then all of IDT positives and a subset of IDT negatives are tested with the reference standard test.
- For IDT Positives (IDPs) and IDT Negatives (IDNs):
  - Based on prespecified design parameters (e.g., nominal number (A1 and A2) of real world IDPs and IDNs to be leveraged), use PS modeling, trimming, and stratification to select and discount real world IDPs and IDNs.
  - Estimate PPV and NPV via PSCL.
- Estimate verification adjusted sensitivity and specificity based on PS-integrated PPV and NPV estimates [1].

Step by Step

Type 1 Study Design



Type 2 Study Design



Example

Type 2 study design that augmented IDPs and IDNs with real world data.

- For IDPs, 415 subjects were needed.
  - Planned to enroll 290 IDPs in the current study (undertake reference test) and borrow 125 from real world data (A1 = 415 - 290)
  - From 1984 identified IDPs from real world data, 125 nominal IDPs were borrowed.
- For IDNs, 660 subjects were needed.
  - Planned to enroll 450 IDNs in the current study (undertake reference test) and borrow 210 from real world data (A2 = 660 - 450).
  - From 6713 identified IDNs from real world data, 210 nominal IDNs were borrowed.

A total of 11 baseline covariates related to test performance were identified to establish PS modeling. Table 1 gave the PS stratification.

Table 1: *Stratum-specific design specifications for an example of type 2 study design.*

PS Stratum	1	2	3	4	5	Total
Number of current study IDPs	58	58	58	58	58	290
Number of real world IDPs	75	231	278	393	1007	1984
Overlapping coefficient	0.75	0.85	0.85	0.79	0.81	
Standardized overlapping coefficient	18%	21%	21%	19%	20%	100%
***Nominal number of leveraged IDPs* (A1)**	23.0	26.4	26.4	24.3	24.9	125
Discounting weights via PSCL	0.31	0.11	0.09	0.06	0.02	
Number of current study IDNs	90	90	90	90	90	450
Number of real world IDNs	42	223	419	745	5284	6713
Overlapping coefficient	0.71	0.77	0.87	0.81	0.58	
Standardized overlapping coefficient	19%	20%	23%	21%	16%	100%
***Nominal number of leveraged IDNs* (A2)**	40.0	43.0	48.9	45.5	32.6	210
Discounting weights via PSCL	0.95	0.19	0.12	0.06	0.01	

Analysis Results:

- $\widehat{PPV} = 0.502$ $(se = 0.021)$, $\widehat{NPV} = 0.983$ $(se = 0.004)$ via PSCL.
- $\hat{P}(Test = +) = 0.176$ $(se = 0.011)$ via Binomial distribution.
- The verification adjusted sensitivity and specificity were calculated.
  - $\widehat{Sensitivity} = 86.5\%$ with 95% CI [80.2%, 92.8%].
  - $\widehat{Specificity} = 90.3\%$ with 95% CI [88.9%, 91.6%].
- The success criteria (Sensitivity>0.80 & Specificity>0.80) were met.

Conclusion

The methodological innovation described in this presentation can significantly advance regulatory science, and greatly improve study design and statistical data analysis where real world data are leveraged in the area of diagnostic devices.

References:

1. Song C, Li H, Chen WC, Lu N, Tiwari R, Wang C, Xu Y, Yue LQ. Principled leveraging of external data in the evaluation of diagnostic devices via the propensity score-integrated composite likelihood approach. Pharm Stat. 2023 Mar 5. doi: 10.1002/pst.2295.
2. Wang C, Lu N, Chen WC, Li H, Tiwari R, Xu Y, Yue LQ. Propensity score-integrated composite likelihood approach for incorporating real-world evidence in single-arm clinical studies. J Biopharm Stat. 2020 May 3;30(3):495-507.