C-177

POWER VS PRECISION WHEN ESTIMATING RISK DIFFERENCES USING OBSERVATIONAL DATA? A CASE STUDY FROM TARGET-IBD

YASMMYN D. SALINAS[1], KAYLA HENDRICKSON[1], TERESA BUFFORD[1], DEREK GAZIS[1], JULIE M. CRAWFORD[1], JENNIFER CHRISTIAN[1], AND KATHLEEN HURWITZ[1]

[1] Target RWE, Durham, North Carolina, United States

TARGET RWE
BETTER EVIDENCE • BETTER HEALTH

Background and Objective

- Traditional power-based sample size calculations are modeled after randomized controlled trials.
- In studies using real-world data, sample sizes are often fixed and adjustment for confounding is necessary.
- Precision—the amount of random variation in the study estimates—may be more informative in observational settings as compared to power-based sample size calculations.
- **Objective**: To compare expected statistical power and precision for estimating the risk difference (RD) for surgery or hospitalization among patients with inflammatory bowel disease (IBD) who achieved early vs. late remission following initiation of an advanced therapy.

Methods

- Data were obtained from TARGET-IBD, a large US-based registry of patients with clinician-diagnosed ulcerative colitis (UC) or Crohn's disease (CD).
- The analysis included 316 UC and 478 CD patients who initiated an advanced therapy (i.e., a biologic, biosimilar, or small molecule) after April 1, 2017.

Power Analysis

- **Sufficient** power defined as ≥80% power
- Computed using G*Power 3.1.9.6 (Faul et al. 2009) using the **effective sample size (ESS)** in treatment group a
- $\boldsymbol{ESS_a} = \frac{N_a}{VIF_a}$ | where $\boldsymbol{VIF_a}$ = **variance inflation factor** in treatment group a

Precision Analysis

- **Sufficient** precision defined as 95% confidence interval (CI) excluding the null value
- **95% CI** for weighted RD = $\widehat{RD} \pm 1.96\sqrt{Var(\widehat{RD})}$
- $\boldsymbol{Var(\widehat{RD})}$ $= VIF_1 * \sigma_1^2 + VIF_2\sigma_2^2$ | where $\boldsymbol{VIF_a}$ = **variance inflation factor** in treatment group a

Calculating VIF due to application of Inverse Probability of Treatment (IPT) weights

$$\boldsymbol{VIF_a} = N_a \sum_{i=1}^{N} \widehat{W}_i{}^2 I(A_i = a) \Big/ \left\{\sum_{i=1}^{N} \widehat{W}_i I(A_i = a)\right\}^2$$

- $\boldsymbol{N_a}$ is the sample size in treatment group a
- $\boldsymbol{W_i}$ is IPT weight for subject i

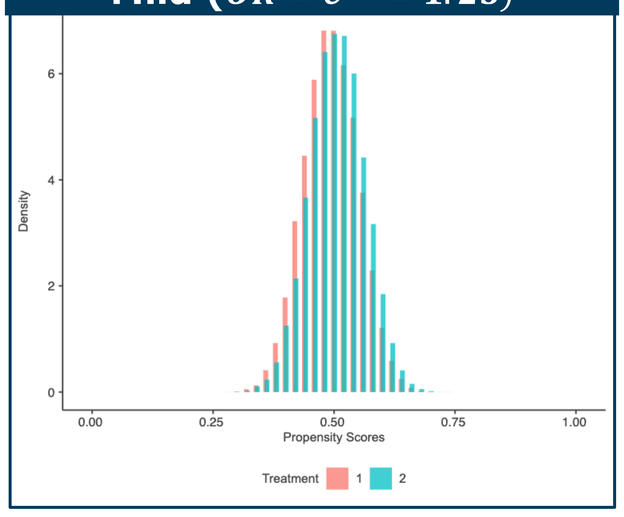
Simulating IPT weights

$$\boldsymbol{W_i} = \frac{A}{P(A=1|X)} + \frac{1-A}{1-P(A=1|X)}$$

- Estimated using the logit model: $P(A = 1|X) = 1/1 + \exp(-\beta_0 - \beta_1 * X)$
- Inputs: β_1 and prevalence of treatment $[P(A = 1)]$

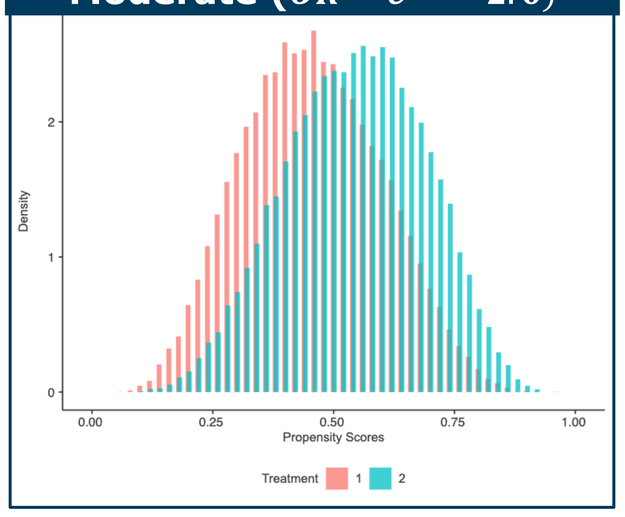
Relationship between β_1 (degree of confounding) and the Variance Inflation Factor

Mild ($OR = e^{\beta_1} = 1.25$)



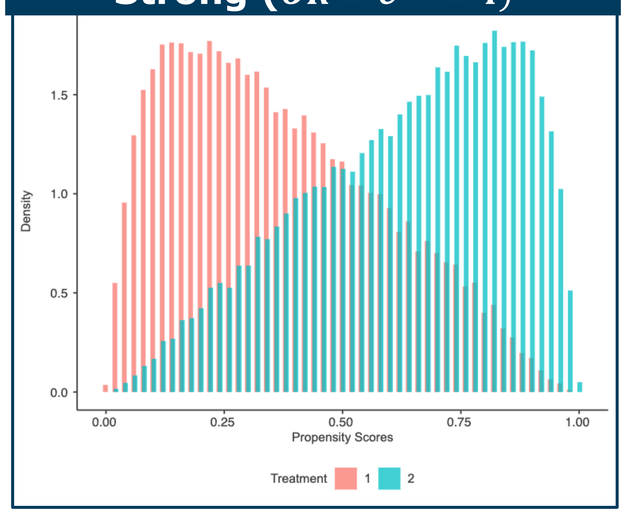
$\boldsymbol{VIF_1 = 1.01; VIF_2 = 1.01}$

Moderate ($OR = e^{\beta_1} = 2.0$)



$\boldsymbol{VIF_1 = 1.14; VIF_2 = 1.14}$

Strong ($OR = e^{\beta_1} = 4$)



$\boldsymbol{VIF_1 = 1.66; VIF_2 = 1.82}$

Input parameters for power and precision analyses

- **Total sample size:** 316 UC and 478 CD patients
- **Prevalence of early remission ($P(A=1)$):** 25–45% (UC) and 32–52% (CD)
- **Baseline risk (r_2):** 14% (UC) and 48% (CD)
- **Effect size (RD):** 10%
- **Degree of confounding:** Moderate
- **Censoring rate**: 10%

Results

			Early remission (A=1)				Late Remission (A=2)							
Disease	N total	P(A=1)	N1	VIF1	ESS1	R1	N2	VIF2	ESS2	R2	Power	95% CI	Sufficient Power?	Sufficient Precision?
CD	478	0.32	153	1.06	130	0.38	325	1.21	242	0.48	46%	0.10 (-0.002, 0.202)	✗	✗
CD	478	0.42	201	1.09	166	0.38	277	1.16	215	0.48	50%	0.10 (0.003, 0.197)	✗	✓
CD	478	0.52	249	1.13	198	0.38	229	1.11	186	0.48	51%	0.10 (0.004, 0.196)	✗	✓
UC	316	0.25	79	1.03	69	0.04	237	1.26	169	0.14	63%	0.10 (0.032, 0.168)	✗	✓
UC	316	0.35	111	1.06	94	0.04	205	1.20	154	0.14	74%	0.10 (0.035, 0.165)	✗	✓
UC	316	0.45	142	1.10	116	0.04	174	1.14	137	0.14	79%	0.10 (0.034, 0.166)	✗	✓

Conclusions

- Precision and power analyses yielded contradictory assessments of study feasibility.
- If we had only calculated power, we may not have proceeded with the study.
- Our findings demonstrate that real-world data can provide valuable insights even in studies with limited power.
- **Precision—not power—should guide feasibility assessments in observational studies.**