Adaptive Sample Size Re-estimation Incorporating Real-World Evidence in Clinical Trials for Small Population or High Unmet Medical Need

Ran Liao, Junjing Lin, Margaret Gamalo

Eli Lilly and Company, Takeda Pharmaceuticals, Pfizer Inc.

Abstract

For the clinical studies with small patient population, to have a fully powered randomized controlled trial is further complicated due to variability. As such, sample size re-estimation has been a useful tool to evaluate if during an interim look the trial sample size needs to be increased to achieve enough power to detect an expected treatment effect. Furthermore, borrowing or extrapolating information from real-world evidence or real-world data has been applied in trial design and analysis. These two innovative approaches can be combined to form an even more powerful technique, i.e., high quality real-world data, if leveraged properly, has the potential to generate real-world evidence to assist interim decision-making, lower enrollment burden, and reduce study timeline and costs.

With proper borrowing from historical control, some of the challenges in these high unmet medical need studies could be resolved considerably, e.g., decrease in trial sample size, shorter trial duration, patient protection from exposure to potential futile treatment. We examine the strategy in pediatric Type II diabetes trials where recruitment has been challenging and the completion is hardly on time. Simulations under various scenarios are conducted to assess the **borrowing strategy, that includes the matching method**, **in combination of sample size re-estimation**. The type I error for each strategy is reported and compared to demonstrate how it is controlled. Comparison of power and reductions in sample size are reported to demonstrate the advantages of proposed method.

Motivation: Pediatric T2 Diabetes

Total pediatric T2D population (~20,000-25,000 U.S.*; 100%)

Socioeconomically challenged population with many inherent barriers to participation

Patients out of reach of study PIs

- Receiving care at clinical care-only facilities
- Receiving care at research sites without dedicated pediatric T2D teams (e.g., T1D focused)

# of patients at sites active in T2D clinical trials (~5,000-6,000 U.S.*; 20%)

"Pre-Screen" Failure

- Unlikely to be eligible
  - Controlled on metformin or insulin
  - A1C or liver enzymes too high
  - Age > 17 years
- Unable to reach

Challenge: Small population, high unmet medical need

# of patients approached for screening (~1,000-5,000 U.S.*; <5%)

Screen Failure

- No show
- Unwilling to participate
- Cannot comply with study visits
- Inclusion/exclusion criteria not met

Challenge: Hard to recruit and long enrollment period

Eligible subjects for pediatric T2D trials (~500-600; 2%)

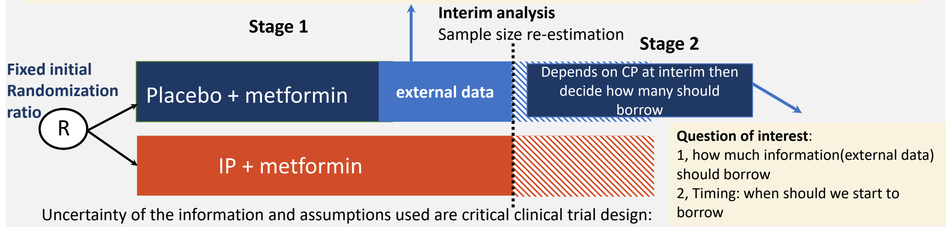
Study	Enroll #.	Start	Primary Completion	Recruitment Duration
Ellipse Trial(Liraglutide)	135	Nov 2012	Nov 2017	~5 years
NCT01760447(Sitagliptin)	110	Feb 2013	Sep 2019	~6.5 years
NCT01472367(Sitagliptin)	140	Dec 2011	Sep 2019	~8 years
NCT01485614(Sitagliptin)	190	Feb 2012	Oct 2019	~7.5 years
TODAY trial*	699	May 2004	Feb 2011	~5.5 years

Plot was cited from: Nadeau et al. Youth-Onset Type 2 Diabetes Consensus Report: Current Status, Challenges, and Priorities. 2016. Diabetes Care.

Sample Size Re-estimation(SSR) with External Borrowing

External data was borrowed through propensity score matching:

- Additional placebo/control patients was "borrowed" from the external controls, for maintaining a one-to-one randomization between the treatment arm and active control, by matching the new treatment and control units based on a set of measured covariates, ie, model-based pairing of treatment and control units that are similar in terms of their observable pretreatment characteristics.

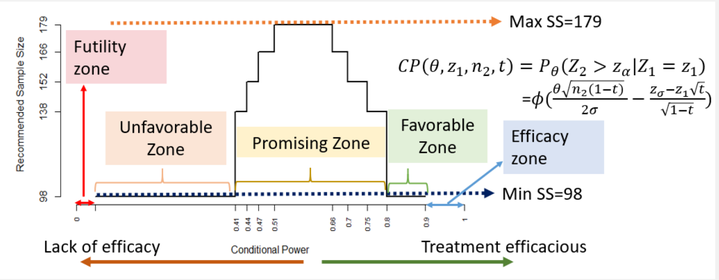


Question of interest:
1, how much information(external data) should borrow
2, Timing: when should we start to borrow

Uncertainty of the information and assumptions used are critical clinical trial design:

- Example: changing medical practice might significantly affect assumed event rate for endpoint
- Example: estimated rates based on historical data might no longer reflect the current medical environment

SSR Decision Rule

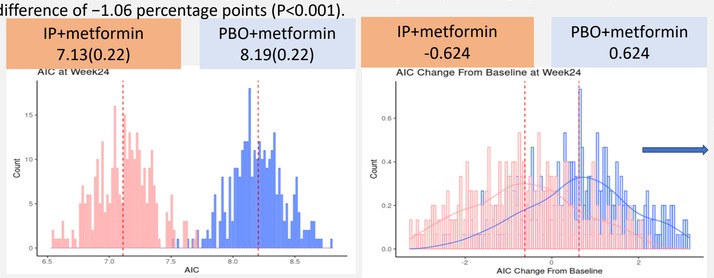


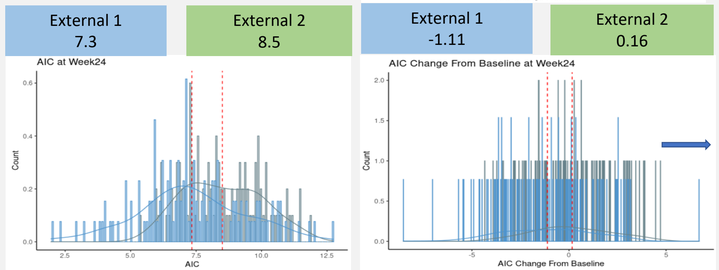
- Allow trial to stop early for efficacy or futility
- Increase sample size only when interim results is promising (assessed by condition power)
- Use external control data: historical trials often have comparable I/E and similar key endpoints

Simulation Setup

Simulated RCT:

- Endpoint of interest: change from baseline glycated hemoglobin at week 24
- Two assumption: 1, Delta=0.7 (1.2), power =80% sample size per arm 47; 2, Delta= 1.2(2.5), power=80%, sample size per arm 69
- Based on recent clinical trial (Ellipse Trial): At the 26-week analysis, the mean glycated hemoglobin level had decreased by 0.64 percentage points with liraglutide and increased by 0.42 percentage points with placebo, for an estimated treatment difference of −1.06 percentage points (P<0.001).





Results and Conclusion

Scenario 1: outcome weakly related to baseline covariates

IA zone	Prob of IA zone	Cond Power (fixed)	Cond Power (adaptive)	Expected Sample Size (fixed)	Expected Sample Size (adaptive)	
					Not borrow	borrow
Futility*	8.9%	4%	3.5%	138	98	74
Unfavorable	13.1%	25%	24.9%	138	138	114
Promising	20.2%	63%	62.7%	138	169	145
Favorable	9.2%	85%	85.4%	138	138	114
Efficacy**	48.6%	98%	98.0%	138	98	74

Expected sample size w/ borrowing = 97.
Assuming the enrollment rate from a recent trial
W/out IA & W/out borrowing: ~61 mos
W/ IA & W/out borrowing: ~44 mos to finish enrollment for IA (70% IF)
W/ IA & W/ borrowing: ~43 mos to finish enrollment for IA

Scenario 2: outcome strongly related to baseline covariates

IA zone	Prob of IA zone	Cond Power (fixed)	Cond Power (adaptive)	Expected Sample Size (fixed)	Expected Sample Size (adaptive)	
					Not borrow	borrow
Futility*	2.6%	4%	0%	138	98	74
Unfavorable	5.9%	26%	24.7%	138	138	114
Promising	13.3%	65%	73.2%	138	168	144
Favorable	7.4%	86%	85.2%	138	138	114
Efficacy**	70.8%	98%	100%	138	98	74

Expected sample size w/ borrowing = 88.
Assuming the enrollment rate from a recent trial
W/out IA & W/out borrowing: ~61 mos
W/ IA & W/out borrowing: ~44 mos to finish enrollment for IA (70% IF)
W/ IA & W/ borrowing: ~39 mos to finish enrollment for IA