Paper PP16

Inference from weighted subjects

Andrea Parodi, Valos, Genova, Italy
Ying-Chieh (Lisa) Lee, Livanova, London, England

ABSTRACT

LivaNova is the manufacturer of Vagus Nerve Stimulation (VNS) Therapy which is an approved adjunctive treatment delivered by an implantable device to treat epilepsy by reducing the frequency of seizures.
We explored a potential clinical trial design to evaluate the effectiveness, safety, and healthcare utilization outcomes of the VNS Therapy for the treatment of drug-resistant epilepsy against the current pharmaceutical Standard of Care (SoC).
The design described is a long-term prospective, multicenter, parallel-group, observational study.
The primary endpoint is the maximum duration of continued, monthly response to treatment. The treatment choice is decided by the physician and patient, prior to receive the new treatment.
From a statistical point of view, the issues are:

- Duration of response rarely follows a common distribution;
- Accounting for baseline differences in the non-randomized treatment groups.

We analyze data with a non-parametric approach, using the overlap-weights technique for minimizing the differences between treatment groups, finding the best way of making inference from them, and comparing this approach with parametric ones.

INTRODUCTION

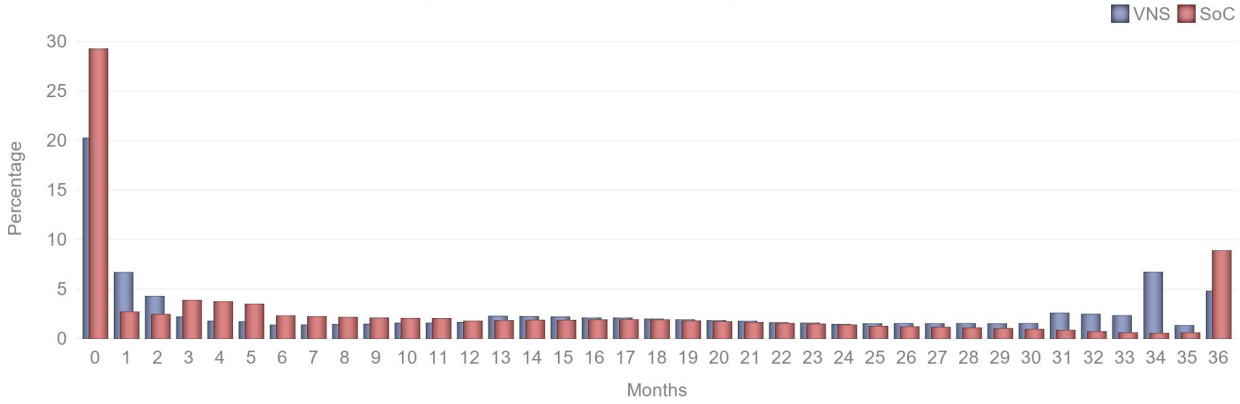
The overlap weights method described by Li et al. (2018) [1] is widely used for minimizing the differences in the baseline covariates of two populations with available Individual Patients Data (IPD). The minimization of differences in baseline covariates reduces the risk of biased comparisons.
In literature, there is no agreement on how to do inference on weighted subjects. Should the analysis be performed considering the original number of subjects? Or their sum of weights? Or other solutions?
We simulated different solutions based on a true study design, aimed at showing the difference between VNS Therapy against the SoC.

METHODS

SIMULATED POPULATIONS

Responses to therapies (VNS and SoC) have been simulated on a monthly basis for 1 000 000 subjects each, based on literature data. Each monthly response depends on baseline characteristics and the responses of the subjects in the 3 previous months.
A total of 36 months has been simulated and the maximum duration of continued response to treatment is the primary endpoint. The distributions of these durations in the 2 populations are shown below in Figure 1.

Figure 1: Maximum duration of response to treatment



BASELINE CHARACTERISTICS

The baseline characteristics of the 2 populations have been simulated based on the input parameters enlisted in the following Table 1.

Table 1: Design parameters

Parameter	VNS	SoC

Cognitive Impairment / Learning Disability	60%	40%
Number of failed therapies	Mean 6.8, SD 3.06, Min 3, Max 9	Mean 8.1, SD 4.1, Min 3, Max 9
Duration of epilepsy (years of Age minus Age at Onset)	Mean 14.4, SD 12.74, Min 0, Max 20	Mean 24.9, SD 13.4, Min 0, Max 20
Age	Mean 23.9, SD 16.55, Min 4, Max 70	Mean 36, SD 18, Min 4, Max 70
Seizure Type	Focal 70%, Generalized 30%	Focal 80%, Generalized 20%
Baseline Seizure Rate per Month	[2, 10): 28% [10, 100): 50% [100, 1000): 22%	[2, 10): 28% [10, 100): 50% [100, 1000): 22%

SAMPLE SIZE CALCULATION

NON-PARAMETRIC, CONFIDENCE INTERVALS CALCULATED USING BOOTSTRAPPING

How many subjects are needed to get a significant superiority of VNS treatment in the maximum duration of continued response to treatment, considering a type I error level of 0.025 and a power of 0.80?

1. From the VNS and SoC populations, a number of subjects equal to *V* and *S* respectively have been randomly extracted without replacement. This step has created 10 000 datasets for VNS subjects named VNS0001-VNS10000 as well as 10 000 datasets for SoC named SOC0001-SOC10000.
2. For each pair of datasets VNS0001-SOC0001, VNS0002-SOC0002 … VNS10000-SOC10000, a logistic regression has been run, using the treatment assignment (VNS or SOC) as event, and the 6 variables of Table 1 as covariates. The predicted probability of being assigned to the VNS group is called propensity score (PS). A weight for each subject has been calculated based on the PS: for SoC subjects the Overlap Weight (OLW) is equal to PS, while for the VNS ones it is equal to 1 – PS.
3. For each pair of datasets VNS0001-SOC0001, VNS0002-SOC0002 … VNS10000-SOC10000, the weighted difference in mean of the duration of maximum response in the two treatment groups has been calculated, together with its 95% confidence intervals calculated with bootstrapping method (randomly extracting the subjects with replacement).
4. If the n^{th} confidence interval did not include 0 in its values, the n^{th} simulation is considered as a success.

The power of the analysis with *V* VNS subjects and *S* SoC subjects has been calculated by dividing the number of successes by the 10 000 simulations done.
Steps 1-4 have been repeated with increasing *V* and *S* numbers until the first value gets a power ≥ 80%. The ratio V:S has been set to 1:2.

PARAMETRIC, CONFIDENCE INTERVALS CALCULATED USING WALD OR ROBUST SANDWICH METHOD

The first 2 bullet points are in common with the first method described.

3. For each pair of datasets VNS0001-SOC0001, VNS0002-SOC0002 … VNS10000-SOC10000, the weighted difference in mean of the duration of maximum response in the two treatment groups together with its 95% confidence intervals have been calculated with SAS PROC GENMOD, assuming a parametric distribution, using Wald or Sandwich methods of calculation for confidence intervals.
4. If the n^{th} confidence interval did not include 0 in its values, the n^{th} simulation is considered as a success.

The power of the analysis with *V* VNS subjects and *S* SOC subjects has been calculated by dividing the number of successes by the 10 000 simulations done.
Steps 1-4 have been repeated with increasing *V* and *S* numbers until the first value gets a power ≥ 80%. The ratio *V:S* has been set to 1:2.
Table 2 shows the parametric distributions, the link function, and the estimation method of confidence intervals.

Table 2: Parametric distributions, link functions, and estimation method of confidence intervals

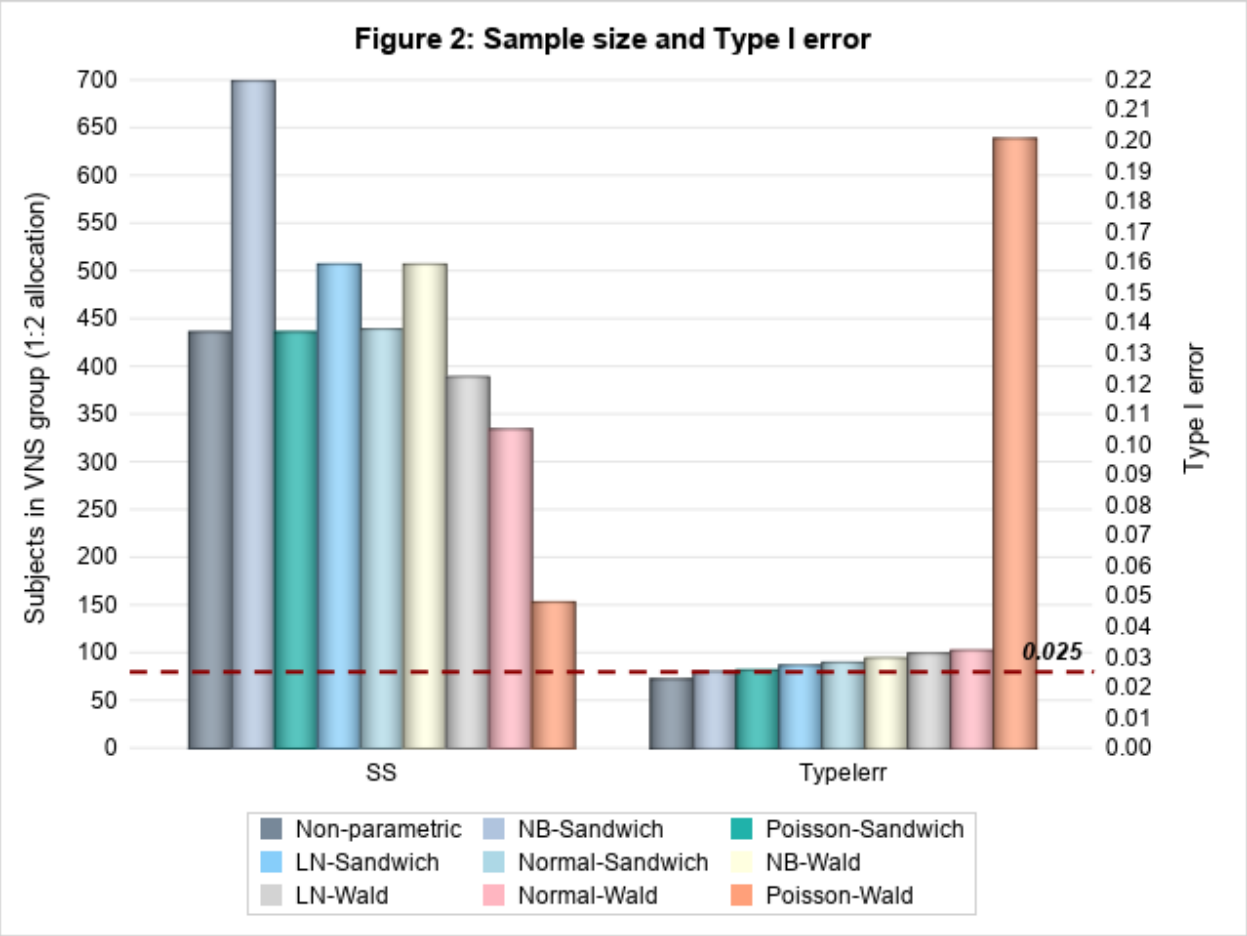
SAS color	Parameter	Link functions	CI calculation
Pink	Normal	Identity	Wald
Blue	Normal	Identity	Sandwich
Grey	Normal (on the logarithm of the endpoint)	Identity	Wald
SkyBlue	Normal (on the logarithm of the endpoint)	Identity	Sandwich
Salmon	Poisson	Log	Wald
SeaGreen	Poisson	Log	Sandwich
Yellow	Negative-binomial	Log	Wald
SteelBlue	Negative-binomial	Log	Sandwich

ASSESSING THE RELIABILITY OF THE METHODS

To assess if the used methods are reliable, the Type I error inflation has been evaluated.

Based on the sample sizes found in the Sample Size Calculation section, the power is calculated using as population the dataset including both the VNS and SoC ones.
The expected power should be close to the 0.025 chosen as alpha for sample size calculation (see Fig. 2).

RESULTS



The non-parametric method with bootstrap confidence intervals is the only one ensuring a Type I error without inflation. All other parametric methods resulted in Type I error > 0.025.
The negative-binomial model with robust estimation of confidence intervals got close (0.0255) but with a number of VNS subjects needed equal to 700 (against the 437 of non-parametric). Also, the Poisson model with robust estimation performed quite well (0.026) with the same number of subjects needed for the non-parametric one.
The worst one is the Poisson model with Wald confidence interval, with a Type I greater than 0.20.

CONCLUSION

The results obtained in this analysis are not generalizable to other scenarios, but are useful to understand what is the correct inference method to be used in this analysis. The distribution of the endpoint is clearly not following any known distribution and the results can be influenced by this fact: other simulations with the same goal will be planned for assessing the robustness of this result when dealing with a different distribution of the endpoint.

REFERENCES

1. Li, F., Morgan, K. L., & Zaslavsky, A. M. (2018). Balancing covariates via propensity score weighting. *Journal of the American Statistical Association, 113*(521), 390-400.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:
Andrea Parodi
Valos s.r.l.
Via Ceccardi 4/31
16149 Genova (Italy)

Work Phone: +39 010 2534160
Email: andrea.parodi@valos.it
Web: www.valos.it

Ying-Chieh (Lisa) Lee
LivaNova PLC
4th Floor, 20 Eastbourne Terrace,
London, W2 6LG, United Kingdom
Work Phone: +44 7823 345403
Email: lisa.lee@livanova.com
Web: www.livanova.com