Using clinical trial and real world data to bridge efficacy to effectiveness of fingolimod in multiple sclerosis patients

Vladimir Ivanov, PhD;[1] Roman Torgovitsky, PhD;[1] Eric Tchetgen Tchetgen, PhD;[2] Bruce Church, PhD;[1] Veronica Alas, MPH, PhD;[1] Iya Khalil, PhD;[1] Valery Risson, PhD, MBA;[4] Kristijan Kahler, RPh, PhD;[3] Melvin Olson, PhD[4]

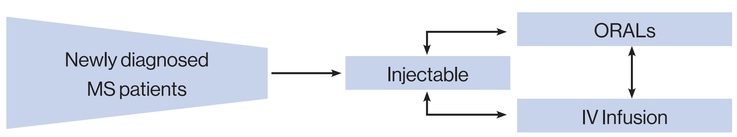
[1] GNS Healthcare, Cambridge, MA, USA; [2] Harvard University, Cambridge, MA, USA;
[3] Novartis Pharmaceuticals Corporation, East Hanover, NJ, USA; [4] Novartis Pharma AG, Basel, Switzerland

PND8

Introduction

- As more treatments come to market, demonstration of how efficacy results from clinical trials (CTs) can translate to effectiveness in the real world (RW) prior to launch is necessary to minimize negative clinical outcomes and prevent further progression of disease.
- Multiple sclerosis (MS) affects 400,000 people in the United States (US)[1] and 2.5 million people worldwide[2], and approximately 85% of patients have relapse-remitting multiple sclerosis (RRMS)[2].
- Relapse frequency and severity vary considerably among MS patients, and increased relapse frequency is associated with a higher risk of disease progression[3].
- Multiple disease modifying treatments (DMTs) for MS are currently used to manage patients' care. The typical first line of DMTs, called BRACE treatments, are injectables. The typical second line treatments are either oral treatments, intravenous infusions, or another BRACE therapy[4] (Figure 1).

Figure 1. Treatments of newly diagnosed MS patients



Objective

- We proposed and implemented an efficacy to effectiveness (E2E) causal methodology to predict the RW effectiveness of fingolimod, an oral MS treatment, prior to launch using CT data and pre-launch RW observational data.

Data Sources & Sample Selection

- Two data sources were used to implement E2E (Figure 2).

Figure 2. Data sources used for E2E analysis

3 randomized controlled clinical trials 799 patients (fingolimod, 243 MS patients)	IMS PharMetrics Plus US Data 25,000 MS patients (2007–2013)

- The first data source was three Novartis phase 3 efficacy randomized controlled CTs for fingolimod with a total of 799 patients. The data sets were combined and patients who were administered fingolimod were selected.
- The second data source was comprised of US administrative medical and pharmacy claims data from IMS PharMetrics Plus Database (2007-2013). The database is comprised of more than 87 million health plan members throughout the US. Data are de-identified in compliance with the Health Insurance Portability and Accountability Act.
  - Two cohorts of patients were created from claims data: one for the modeling (pre-launch) and another for the validation approach (post-launch). Patients prior to and after fingolimod's 2010 launch were included in the study (25,000 patients), and a previously used algorithm to identify relapses within this data source was used.[3]
  - The following inclusion and exclusion criteria were used to identify MS patients who initiated or switched to a new DMT, using claims data and based on previously published work:[5]
    - Inclusion criteria: Age 18-55 years at index date; at least one claim with a diagnosis of MS (ICD-9-CM: 340.xx) within the pre-index period; evidence of at least one diagnosis with MS (ICD-9-CM: 340.xx) in post-index period; continuous coverage for pre- and post-index period; at least one MS-related relapse in the pre-index period.
    - Exclusion criteria: Patients with index DMT during pre-index period; patients receiving more than one DMT on the index date.

Modeling Approach and REFS™ Analytical Platform

- A head-to-head comparison, or "pseudo-trial", was constructed between fingolimod and other disease-modifying therapies (DMTs), using pooled CT data for fingolimod pre-launch, and pre-launch observational data for the reference DMTs (interferons and glatiramer acetate).
- Since CT and RW data often use different coding systems, we identified common variables in CT and RW data and created common data categories to use as covariates. Figure 3 shows comorbidities aggregated to disease classes for CT and RW patients.
  - All model covariates were defined in a 1-year period before DMT initiation.
  - Endpoints (Figure 4) were defined in a 1-year period after DMT initiation.
- The E2E causal method projected the endpoints from pooled CT data to RW data. CT population standardization was completed by a weighting technique computed with GNS Healthcare's proprietary machine learning platform, Reverse Engineering and Forward Simulation (REFS™).[6]
- REFS™ uses Bayesian network inference to learn an ensemble of models directly from the data—in this case, from CT and claims data—without using *a priori* hypotheses. A large (potentially billions to trillions) space model is explored, and each proposed addition/subtraction is scored based on a maximum entropy structural prior and a complexity penalty imposed by the Bayesian Information Criterion.

Figure 3. Population comorbidity categories in pre-index period

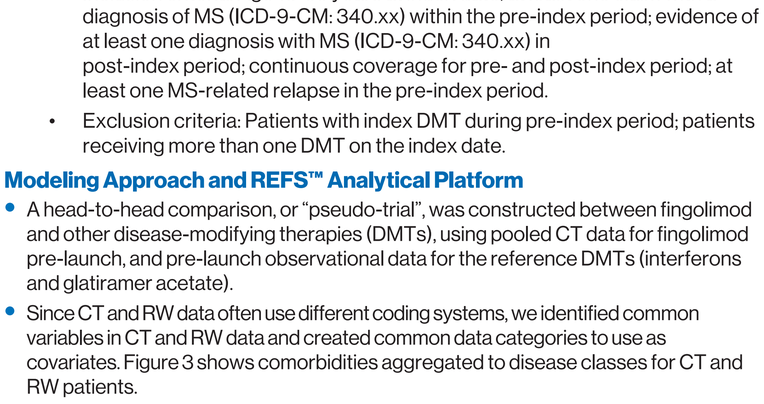
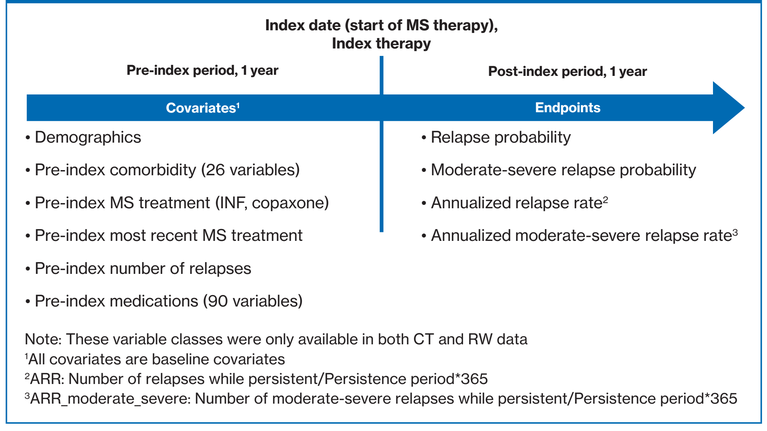


Figure 4. Model covariates and endpoints



Note: These variable classes were only available in both CT and RW data
[1]All covariates are baseline covariates
[2]ARR: Number of relapses while persistent/Persistence period*365
[3]ARR_moderate_severe: Number of moderate-severe relapses while persistent/Persistence period*365

E2E Method

- Drug performance results of the "pseudo-trial" were estimated using a two-stage process:
  - Potential confounders were identified in the pseudo-trial, and their joint distribution was modeled. We assumed that the confounders in the pseudo-trial were the same confounders, which can be identified from the pre-launch observational data.
  - Expected (mean) outcome was modeled in the CT data given the confounders.
- We described CT and RW environments by the vector , where
  - *P* defines observable baseline patient characteristics variables in CT and RW.
  - *E* defines observable baseline healthcare system characteristics variables and post-baseline observable patient-healthcare interaction variables in CT and RW.
  - We assumed that variables contain all information about differences between CT and RW settings.
- *Contrast of interest:* We compare fingolimod (from CT data) to the existing DMTs among all patients in the RW:
  $E(Y_{A=A_{CT}} - Y_{A=A_{RW}} \mid Env=RW)$
  where Y_A is a counterfactual value of outcome Y if the treatment A was prescribed, Env stands for environment, e.g. CT or RW
- Derivation of $E(Y_{A=A_{CT}} - Y_{A=A_{RW}} \mid Env=RW)$
  - Conditioning on P and E and integrating out
    $\int E(Y_{A=A_{CT}}, P, E, Env=CT)\, f(P,E \mid Env=RW)\, dP\, dE$
  - Estimation of regression and multi-dimensional density might be a difficult problem due to curse of dimensionality
- A semi-parametric approach focused on estimating the logistic regression:
  $\text{Prob}(A = A_{ct}, Env = CT \mid P,E)$ and $\text{Prob}(A = A_{rw1}, Env = RW \mid P,E)$
- Closed form solution:
  $\psi E(w \mid [Env=CT, A=CT] Y)$, with weights which are used to standardize CT population to RW population
- A logistic regression model was learned using REFS™ and weights were estimated.

Results

- Our model estimated the RP in the RW to be between 0.12 (for moderate-severe relapses) to 0.24 (all relapses). This compares to the 0.11 to 0.22 seen in the CTs. Hence, the correction for the difference in the joint distributions of covariates between the CTs and the RW data pre-launch has not had a large effect on the point estimates of RP (Table 3).

Modeling Approach Validation

- The validation using post-launch administrative data rests on the assumption that characteristics of the MS population do not change from the pre-2010 to post-2010 period.
- We used post-launch administrative data to re-compute the weights and estimated RP and ARR, if fingolimod were prescribed to the entire post-launch fingolimod population. The model estimated RW RP between 0.13 (moderate-severe relapses) to 0.29 (all relapses). The observed RP for fingolimod in the RW was 0.24 (Table 4).

Statistical Assumptions & Limitations

- E2E methodology rests on the following statistical assumptions:
  - **Sufficient variables**. There are a finite number of relevant observable variables. Variables $P=(P_1, ..., P_p)$, $E=(E_1, ..., E_p)$ which define patient population characteristics, baseline healthcare characteristics and post-baseline healthcare—patient interactions in CT and RW. We assume that the vector (P,E) contains all information about the differences between CT and RW settings.
  - **All sufficient variables are observed**. All variables (confounders) required for the conditional exchangeability to hold are observed.
  - **Common support for sufficient variables**. Probability densities for all components of the covariate vector (P, E) in the clinical trial and the real world are non-zero on the same interval. For example, for the covariate "age," we assume that density of the variable age for CT data and the density of the covariate age for RW data are non-zero on the same interval 18-65.
  - **Randomized trial**. In the randomized trial counterfactual $Y_A = A_{ct}$ is independent of the drug A that was prescribed in reality.
  - **Consistency**. The consistency assumption states that the counterfactual value of the outcome Y, e.g., RP, under treatment A is the same as the observed outcome Y under treatment A.
- There is no diagnosis code for relapse in administrative RW data. A previously used algorithm was used to identify relapses, which works best for identifying moderate-to-severe relapses.
- Administrative claims data may be prone to undercounting relapses. Hence, estimated RP and ARR were most likely underestimated.

Table 1. CT and pre-launch RW populations used in modeling

	CT	RW
Number of patients	243	7,471
Mean age (years)	37 ± 7.9	42 ± 8.6

Table 2. Means for REFS™ selected model covariates in the CT data before and after weights application

Variable	Weighted CT	CT
Age at index date	41.78	37.80
Antianxiety agents	0.15	0.08
Antidepressants	0.30	0.19
Hypnotics	0.12	0.15
ADHD/anti-narcolepsy/anti-obesity/anorexiants	0.12	0.06
Analgesics- Opioid	0.12	0.13
Analgesics – Anti-inflammatory	0.16	0.18
Musculoskeletal therapy agents	0.17	0.07
Musculoskeletal and connective tissue disorders	0.45	0.23
Nervous system disorder	0.89	0.63
Psychiatric disorders	0.37	0.32
Cardiac disorders	0.10	0.07
Reproductive system and breast disorders	0.20	0.11
Eye disorders	0.26	0.20
Gastrointestinal disorders	0.30	0.20
General disorders administration site conditions	0.19	0.13
Glatiramer acetate therapy	0.10	0.08
INF therapy	0.21	0.30

Table 3. Probability of relapse (PR) and annualized relapse rate (ARR) of fingolimod (pre-launch) projected to RW in comparison with RW DMT (interferons, glatiramer acetate)

	Probability of relapse (RP)	Probability of moderate-severe relapse (RPm-s)	Annualized Relapse Rate (ARR)	ARR moderate-severe
CT (Projected)	0.24	0.12	0.32	0.14
CT (Observed)	0.22	0.11	0.28	0.13
RW (RefP- pre-2010)	0.27		0.47	

Note: RefP- pre-2010 is the patient population prior to fingolimod's 2010 launch.

Table 4. RP and ARR validation model estimates and observed values (post-2010)

	Probability of relapse	Probability of moderate-severe relapse	ARR	ARR moderate-severe
CT (Projected)	0.29	0.13	0.36	0.15
CT (Observed)	0.22	0.10	0.28	0.13
RW-fingolimod > 2010 (RefP-F- post-2010)	0.24		0.34	

Note: Results of modeling in bold, i.e., CT (Projected). RefP-F- post-2010 is the patient population after fingolimod's 2010 launch.

Conclusions

- We implemented an E2E approach to estimate endpoints of interest (RP and ARR) from CT by standardizing to the reference population which allows making a population level prediction of effectiveness in the RW while the drug is still in the development stage.
- It should be noted that a previously conducted RW study has shown that the treatment effect seen in MS on relapses in the RW is similar to that seen in CTs.[5] This may also suggest that the correction for the difference in joint distributions of covariates between CTs and RW data is not large in this disease area and corroborates the findings seen here.
- The GNS Healthcare E2E methodology is generalizable to other CT outcomes and to other RW populations. The key to generalization is the generation of appropriate mappings between CT and RW outcomes.
- This project demonstrates that these mappings can be developed in the MS space. Further research of this type in other disease areas is needed to know the extent of general applicability that can be achieved.

References

1. Hersh CM, Fox RJ. Multiple Sclerosis. Cleveland Clinic Center for Continuing Education. Available at: http://www.clevelandclinicmeded.com/medicalpubs/diseasemanagement/neurology/multiple_sclerosis/ [Last accessed 13 October 2016].
2. Bainbridge JL, Rieckmann P. Multiple sclerosis. In: DiPiro JT, Talbert R, Yee G, Matzke G, eds. Pharmacotherapy: a pathophysiologic approach. New York: McGraw-Hill Medical, 2008:913–26.
3. Richards RG, Sampson FC, Beard SM, et al. A review of the natural history and epidemiology of multiple sclerosis: implications for resource allocation and health economic models. Health Technol Assess 2002;6:1–73.
4. Marziniak M, Meuth S. Current perspectives on interferon beta-1b for the treatment of multiple sclerosis. Adv Ther 2014;31:915-931. doi:10.1007/s12325–014–0149–1.
5. Bergvall N, Makin C, Lahoz R, Agashivala N, Pradhan A, Capkun G, et al. Comparative effectiveness of fingolimod versus interferons or glatiramer acetate for relapse rates in multiple sclerosis: a retrospective US claims database analysis. Curr Med Res Opin 2013;29:1647–56.
6. Anderson JP, Parikh JR, Shenfeld DK, Ivanov V, Marks C, Church BW, Laramie JM, Mardekian J, Piper BA, Willke RJ, Rublee DA. Reverse engineering and evaluation of prediction models for progression to type 2 diabetes: An application of machine learning using electronic health records. J Diab Sci Tech 2016;10:6–18.

Acknowledgement

The authors thank Daniel Shenfeld of GNS Healthcare (Cambridge, MA, USA) for intellectual and technical support and Maninder Reddy Karka for designing the poster layout.

Poster presented at ISPOR 19th Annual European Congress, 29 October – 2 November 2016, Vienna, Austria.

Scan to download this poster