Natural History Study – A Gateway to Treat Rare Disease

Tabassum Ambia, Senior Statistical Programmer, Alnylam Pharmaceuticals, Inc.

ABSTRACT

A Natural History Study is a study that follows a group of people over time who have, or are at risk of developing, a specific medical condition or disease. It bears significant importance in the discovery, marketing and post-marketing phases of a drug. There are different types of natural history studies which may help to determine the requirement of the adequate treatment in a target population or to assess the outcome of a treatment in real life. Real world evidence (RWE) data are the primary source of health information to be used in a natural history study. Statistical analysis is focused on both incidence and prevalence of the disease involving different procedures to determine the distribution of characteristics or events over a certain time and association with covariates.

A rare disease is one which affects a very small percentage of the population. Recently FDA has emphasized the importance of natural history studies for the development of orphan drugs for rare diseases. Natural History Studies can also help clinical research to treat non-rare disease.

This paper will discuss the basics of natural history studies in line with rare disease and orphan drugs, real-world evidence data for natural history studies in pharmaceutical setting, types of natural history studies, analysis techniques and limitation, difficulties during the development of orphan drugs which can be mitigated through natural history studies and usage of natural history studies in Randomized Controlled Trials (RCTs) to develop the treatments of rare disease per FDA guidance.

NATURAL HISTORY STUDY

A Natural History Study is an observational study that follows a group of people over time who have, or are at risk of developing, a specific medical condition or disease. Traditionally a natural history study is defined as the course of progression of the disease in a target population without proper intervention. This includes the onset of the disease until its resolution or the patients' death.

In natural history studies, researchers examine how a disease or medical condition develops over time. Researchers follow participants who are at risk for or have a specific disease or condition to determine how specific factors, e.g. family history, age, occupational exposure and treatment of interest, affect the development or progression of the disease or condition. Data collected during a natural history study also provide information on how best to treat different stages of a disease.

REAL WORLD EVIDENCE FOR NATURAL HISTORY STUDY

Real world evidence (RWE) in medicine means evidence obtained from real world data (RWD), which are observational data obtained outside the context of randomized controlled trials (RCTs) and generated during routine clinical practice. These data are obtained from medical records comprising of clinician's observation, electronic health records, claims and billing activity, product and disease registry, patient-generated data including in home use settings, phone calls, survey etc. These are the clinical evidence associated with the proper usage, benefits and risk of a medical product. RWD help in early discovery of conditions and requirement of a drug, trial design, trial feasibility, method of trial execution, market access and post-market assessment of a treatment.

RARE DISEASE

A rare disease is one which affects a very small percentage of population – defined as the one affecting fewer than 200,000 in USA at a given time, or fewer than 1 in 2000 in Europe.[2] Many rare diseases appear early in life, and about 30% of children with rare diseases will die before reaching their fifth birthday.[2]

ORPHAN DRUG

An orphan drug is a pharmaceutical agent developed to treat medical conditions which, because they are so rare, usually would not be profitable to produce without government assistance.[3]

USES OF NATURAL HISTORY STUDY

- **Drug Development:**
1. **Identification of the Target Population:** Rare diseases often have certain genotypic and/or phenotypic heterogeneity which are hard to identify. Natural history studies may detect important features or events in patients which predict disease progression more precisely. The rates of attacks and progression of disease in the specific population help to determine the inclusion criteria, the stage of disease required to be treated, particular endpoints, frequency of data collection, duration of a trial and more.
2. **Determination and Observation of Clinical Outcome Assessments:** Natural history studies help to assess the safety and effect of a treatment, observing and analyzing how a patient feels, functions and survives. The following types of outcome assessments are considered:
     - Clinician-reported outcomes
     - Observer reported outcomes
     - Patient-reported outcomes
     - Performance outcomes

   A natural history study helps to determine the precise outcomes to assess the progression or symptoms of a disease. This helps to determine the ideal time and a more effective clinical pathway to treat patients.
3. **Identification or Development of Biomarkers:** Natural history studies can help to identify or develop biomarkers which can help with the diagnosis of a rare disease, observing the course of a disease and predictive of treatment response, or useful in patient and dose selection.
4. **Design of Externally Controlled Studies with Natural History Study Data:** In order to qualify for market approval, sponsors need to demonstrate the evidences of safety and efficacy of a drug from well-controlled investigations. This requires a valid comparison to a control. Natural history studies can help the sponsor select the ideal control group to differentiate between outcome of the investigational drug from effects of other factors. The observational data also helps to limit bias.
5. **Post-marketing Assessment:** Not only the development of the drug, but how the drug actually performs on the patients in the real world, is important. Real world evidence data through natural history studies help to observe the actual effect on the patients of interest following the release of a drug.
6. **Usage Assessment:** Observational studies help to determine the proper usage of a drug by selecting the right indications. It is also helpful in determining or answering questions regarding the safety or adverse events leading to the appropriate use of the drug.

- **Other Uses:**
  Natural History Studies may also help patients by improving communication pathways, identifying disease specific centers for excellence, evaluating the standard of care, improving patient care, estimating disease characteristics with prevalence and help with disease tracking.

TYPES OF NATURAL HISTORY STUDY

- **Retrospective and Prospective Natural History Study:**
1. **Retrospective studies:** Focus on the on available data where evaluation has already occurred. Retrospective studies may be considered beneficial as a quicker means to collect information versus implementing a prospective study. However, these studies may have additional constraints, e.g. limited data due to depending on the information already available, not able to determine data collection methods/standards or variables, data elements may lack comparability due to being collected at different points of time, data may be prone to referral bias if site only included serious patients, may have length biased sampling if some patients have been in database for short time frame and others much longer.

2. **Prospective Studies:** Focus on the future evaluation over time based on prespecified data reflecting the current standard. Prospective studies can address the common limitations of retrospective studies as the standard, method, procedure and duration of data collection can be determined. However, due to the nature of the study, additional time/cost is needed to implement a prospective study.

- **Cross-Sectional Studies and Longitudinal Natural History Studies:**
1. **Cross-sectional studies**: Data are collected at a specified time period. Age at incidence distribution is a common analysis. Other covariates may be measured too. Prevalent cohort is more useful for rare disease. Time of sampling may correspond with covariates, so sometimes time is considered an outcome variable, but this may lead to bias. It is often used in the drug industry because the general course of a disease can be followed by sampling at different stages of a disease, can provide the range and severity of the manifestations of the disease and their evaluation, information about treatment for immediate response adequate for a certain stage of a disease. Cross-sectional studies may be prone to length biased sampling due to data being collected at a specific time rather than over longer time frame to accommodate rapidly progressing subtypes.
2. **Longitudinal studies**: Data are collected at different points over time. Often longitudinal studies focus on the timing of recurrent events during disease progression and identify prognostic factors. Data can be analyzed over different stages of disease progression and reduce bias. Longitudinal studies may solve the limitations of cross-sectional studies but duration of study is often longer.

ANALYSIS TECHNIQUES

Analysis of a natural history study focuses on both incidence and prevalence of the disease but more importantly the incidence. Descriptive statistics including frequency, mean, correlation and regression, time to event analysis including survival analysis and Kaplan-Meier Plots are often used.

Often during an event of death, subject is censored or dropped from further analysis. Event of death should not be used as a censoring point because it is unlikely for the death to be independent of the time to event. Or if death was independent, most standard methods for estimating time to the event are in fact targeted to natural history parameters in a world where death does not occur. The event of death should be analyzed further to determine any probable association with the treatment of interest or relation with certain characteristics of a target population. Generalized estimating equations (GEE) may be used for longitudinal data truncated by death.

Example of an exploratory natural history study:

Pulling this altogether, let's consider an example, anonymized multi-center, multi-national, observational, natural history study which was to characterize the natural history and clinical management of patients with a rare genetic disease of interest. We used descriptive statistics to assess the following in 2 parts:

- Part A - During recurrent events of pain for 1 year:
  - Biomarker levels between and during pain events
  - Signs and symptoms between and during pain events
  - Medical history and family history
  - Quality of Life assessment and healthcare utilization
  - Laboratory parameters including chemistry, hematology and urinalysis
  - Exploratory biomarker sample collection
- Part B – Long term analysis for up to 3 years:
  - Pain intensity and impact as measured by Brief Pain Inventory form
  - Changes in disease activity as measured by survey instruments

The intended analysis outputs comprised of descriptive statistics including frequency tables for the categorical responses and tables containing mean, standard deviation, median, 1st and 3rd quantiles, minimum, maximum for continuous responses. Besides the disposition, demographics and medical history tables, we created tables for questionnaire analysis which pertained to the baseline disease characteristics and Quality of Life (symptoms such as pain, number of attacks, ability to work, hospital visit), Brief Pain Inventory – Short Form score which required corresponding calculations assessing the

response to multiple questions for each parameter. We also developed listings for demographics, eligibility and questionnaire responses; and figures displaying the average number of attacks by visit, percent of subjects with disease status change and forest plot of risk factors for annual attack rate.

Datasets were mapped to CDISC Standard Data Tabulation Model (SDTM). At the beginning, we decided to generate the intended analysis result outputs directly from the SDTMs instead of creating Analysis Data Model Datasets (ADaM) as we did not have any treatment and analysis method was different from a typical RCT. However, we faced significant difficulties during analysis which are discussed in the next section.

CHALLENGES

A complexity factor for analyzing rare disease trials is that the sample size is extremely small. A study usually needs to be conducted in multiple centers, presumably with a wide range of demographics with different standards and situations over the duration of the trial. Some studies experience a higher risk of loss to follow-up, which compounds the impact to the intended sample size. During programming, data across the population may need to be pulled from several different studies conducted at different timepoints. This may lead to programming difficulties while appending and merging subjects from different sites with different reference numbers, parameter names and standards. This requires careful observation and familiarization with the data, precise programming and innovative end point selection to reach the objectives.

Unique Subject Identifier (USUBJID) in our study was different between Part A and Part B. Subjects who enrolled in Part B from Part A, had the Subject Identifier for the Study (SUBJID) from Part A recorded under PARTANUM in SUPPDM dataset. Also, medical history which recoded the disease of our interest was not captured for Part A subjects in our Part B data. So we merged the datasets in order to pull the precise medical history of Part A subjects. This requirement was detected following careful observation of the data.

We did not create ADaM datasets initially but observed the SUBJID mismatch after generating the first round of tables. This required us to move back to the data manipulation stage and fix the issue. In order to avoid such situations for future reruns of tables, we decided to follow our usual standard of creating ADaM datasets as inputs to the creation of tables, listings and figures (TLFs). In that way, such issues could be detected early and corrected at the beginning before reaching the TLF stage.

We also decided to check the ADaM datasets through Pinnacle 21. Although this would lead to errors because of some common variables expected would be missing, it helped us follow the data standard as closely to RCTs as possible despite having no treatment.

The analysis helped us determine the incidence of the disease and association with the treatment to have a better idea about the effect of our treatment on the intended population.

CONCLUSION

With increasing research to discover the treatments for rare disease, natural history studies help to meet new challenges and the requirements of the clinical development regulatory framework. Selection of the right patient sample at a certain stage of the disease, observation of the disease progression and selection of the ideal parameters to determine the outcome through natural history studies allow overall more effective strategy of drug development, patient care and product marketing.

Using similar data and programming standards from RCT trials for analysis of Natural History trials add to overall efficiency in Programming and Analysis of these unique studies.

REFERENCES

1. NIH. 2017. "FAQ About Rare Diseases". Publication from GARD – Genetic and rare Disease information Center. Available at: https://rarediseases.info.nih.gov/diseases/pages/31/faqs-about-rare-diseases
2. Wikipedia. "Rare Disease". Available at: https://en.wikipedia.org/wiki/Rare_disease

3. Wikipedia. "Orphan Drug". Available at: https://en.wikipedia.org/wiki/Orphan_drug
4. FDA. 2019. "Rare Diseases – Natural History Studies for Drug Development". Publication from Food and Drug Administration. Available at: https://www.fda.gov/media/122425/download

CONTACT INFORMATION

Your comments and queries are valued and encouraged. Please contact the author at:

Tabassum Ambia
Alnylam Pharmaceuticals, Inc.
101 Main Street
Cambridge, MA 02142
Email: tambia@alnylam.com , ambia@bu.edu