REAL-WORLD & ADVANCED ANALYTICS

Cytel

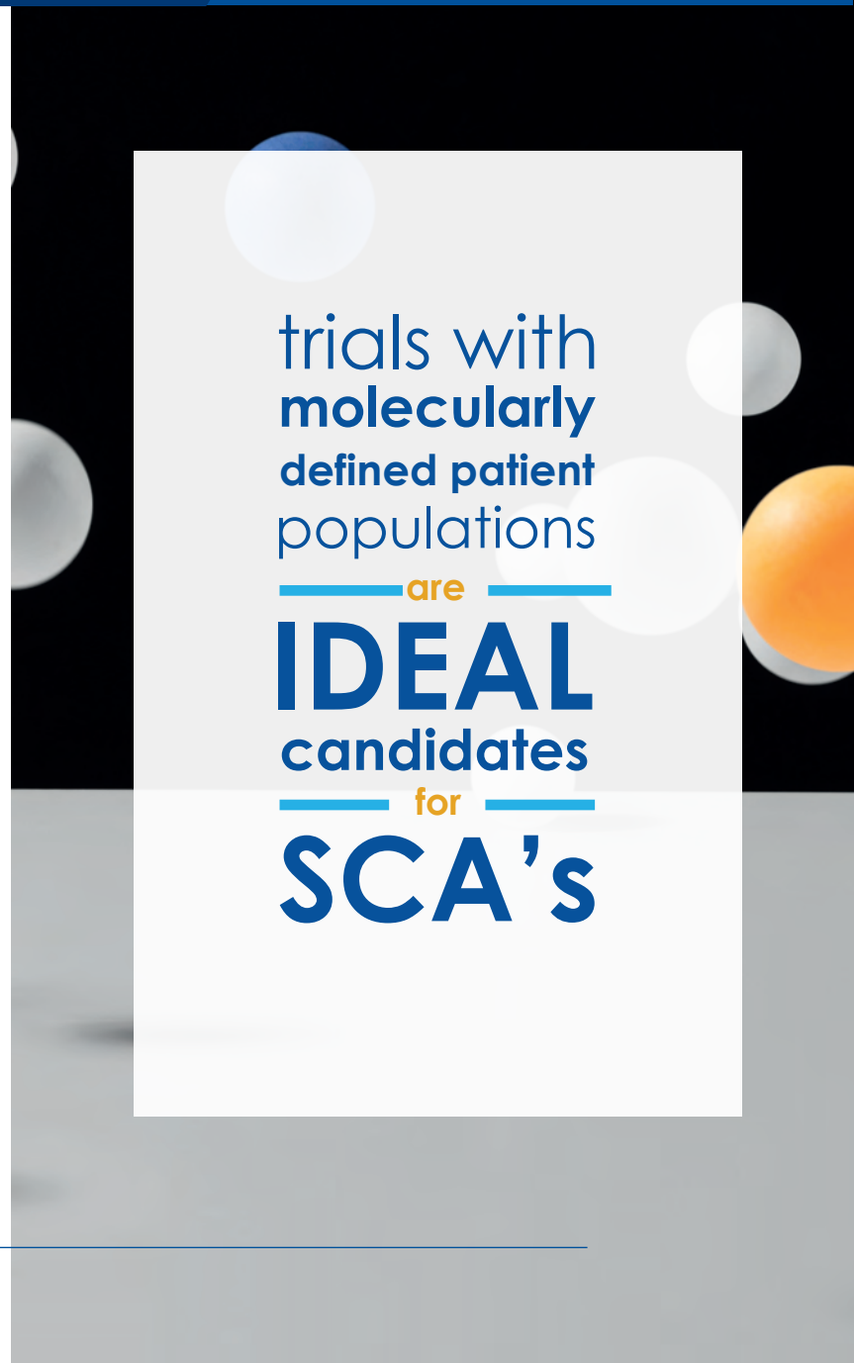
Demystifying synthetic control arms

Over the past decade, a new trend began to emerge, changing the way that clinical trials are conducted. Whereas placebo-controlled randomized control trials remain the gold standard, in some situations, single arm trials have become an accepted way of assessing a new treatment intervention. Single arm trials establish clinical benefit by demonstrating the positive effects of a new therapy or treatment, without the need to use placebo or standard of care as a control. Instead, alternative approaches of establishing the comparison are used; these have become known as external controls or synthetic control arms (SCA henceforth) and include approaches leveraging real world data from various sources or evaluations of historical clinical trial data.

The potential simplification of clinical development by using single arm trials and SCAs does not mean, however, that the application of the approach is straightforward. It requires a careful assessment of trade-offs between data content and their availability, as well as the determination of the optimal methodology. One retrospective study found that single arm trials in oncology often do not meet criteria for clinical benefit as established by the European Medicines Agency. Others have questioned whether non-randomized trials are equipped to handle the needs of unbiased comparisons, or whether they are simply a way to deal with unmet recruitment targets or the over- bureaucratization of the regulatory process.

In contrast, there's a case to be made for making the most use of available real world and historical trial data, if the alternative is to conduct experiments on patients that only marginally improve robustness of findings. There are also legitimate cases, such as rare disease and/or genetic conditions, where recruitment is a challenge. Finally, enrolling into a control arm might be viewed as un-ethical (e.g., in pediatric conditions or some oncology indications). SCAs provide a solution to these issues, but require appropriate use.

SCAs are best suited for situations when a single arm trial is run in a patient population which is molecularly defined, allowing for a clearly defined historical or real-world control group to be created. They are also useful in situations where RCTs opt to enroll some, but fewer patients into the control arm (e.g. a 4:1 allocation ratio of historical data to newly enrolled patients; see Case Study on Page 15). Advanced statistical methods are applied to historical trial or real world data to build the SCA in a way that allows for the appropriate comparison with data gathered during the execution of the single arm trial. As such, this approach to arriving at the comparison between the new and the currently available treatments requires substantial amount of scientific and operational rigour, and there is a need for continuing education to understand the various ways SCAs are constructed.



This ebook is an effort to explain common strategies for the construction of SCAs, and to give a high-level overview of how statisticians think through challenges encountered during trial design. Awareness of these varying practices might enable the non-specialist to determine under what circumstances a placebo-controlled RCT is most appropriate and under what conditions to use single arm trial combined with an SCA. The goal is to create a framework to help readers explore various uses of data and how these approaches can help in achieving the goals of clinical development programs.

TABLE OF CONTENTS

INTRODUCTION

Regulators in both the United States and Europe have responded positively to the use of SCAs in clinical development. While implementation of this method for regulatory purposes might be a new development, the majority of statistical and mathematical theories used for the design of SCAs are decades old and familiar to the scientific community. **The desire to speed up and lower the cost of drug development, coupled with increased availability of rich real-world data, contributed to the increased openness towards using SCAs as supplementary evidence to accompany regulatory submissions using single arm trial data only.**

For instance, in oncology, where enrichment and stratification strategies had resulted in nearly 24% of oncology trials resembling rare disease studies, SCA has led to the expedited approval of several new therapies. AstraZeneca's EGFR modulator Tagrisso received FDA approval within three years from the first patient receiving a dose, while Loxo's Vitrakvi received approval in five years.

The use of SCA has also enabled investigators to leverage combined data from uncontrolled single-arm trials and post-market studies to create much smaller trials. Pfizer's Ibrance (for men with breast cancer) was the result of a Synthetic Control Arm composed of a post-market extension of a similar therapy for women, combined with results from three databases. In non-oncology indications, single-arm trials are often used in combination with natural history studies. Using established knowledge of natural disease progression, Novartis was able to obtain approval for Zolgensma (gene therapy for spinal muscular atrophy) after enrolling only 22 patients. BioMarin's Brineura also obtained approval using a natural history study in the comparator arm, after enrolling 24 patients. It is now the only drug on the market to combat Batten's disease.

INTRODUCTION

Just as various adaptations suit different needs in adaptive trial design, the types of data available to sponsors, the desired sample size, and the anticipated length of a trial might determine the right approach to designing an SCA. Once the appropriate data sources are identified, a careful assessment needs to be made about the potential biases which are inherent in real world data or present in historical trial data as these data were not collected for the purposes of the comparison with a new single arm trial.

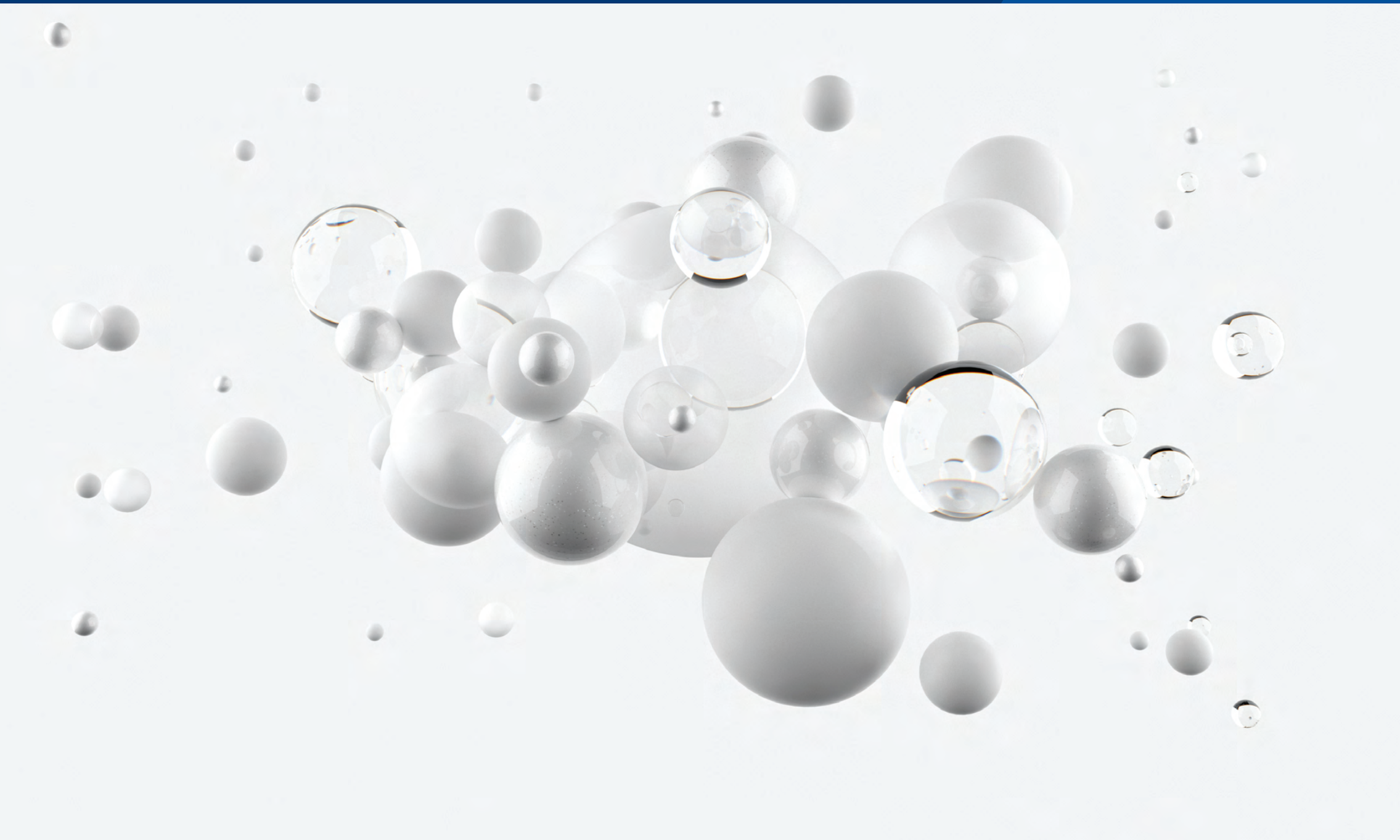
These methods can be used together; e.g., propensity scoring can also be used within Bayesian dynamic borrowing, in order to assign weights to external control data, that better reflect how they match to the population in the treatment group. This is particularly useful when multiple sources of external data are being pooled and works well with aggregate level external control data.
(See pages 11 and 14 for more information.)

The methods described in this ebook offer analytical approaches to correct for possible biases:

- Propensity scoring corrects for bias arising from the fact that the populations receiving an experimental therapy may not be the same as the population receiving control.
- Bayesian dynamic borrowing is particularly useful when external data are not entirely homogeneous with the characteristics of patients enrolled in your trial.

METHODS

Cytel

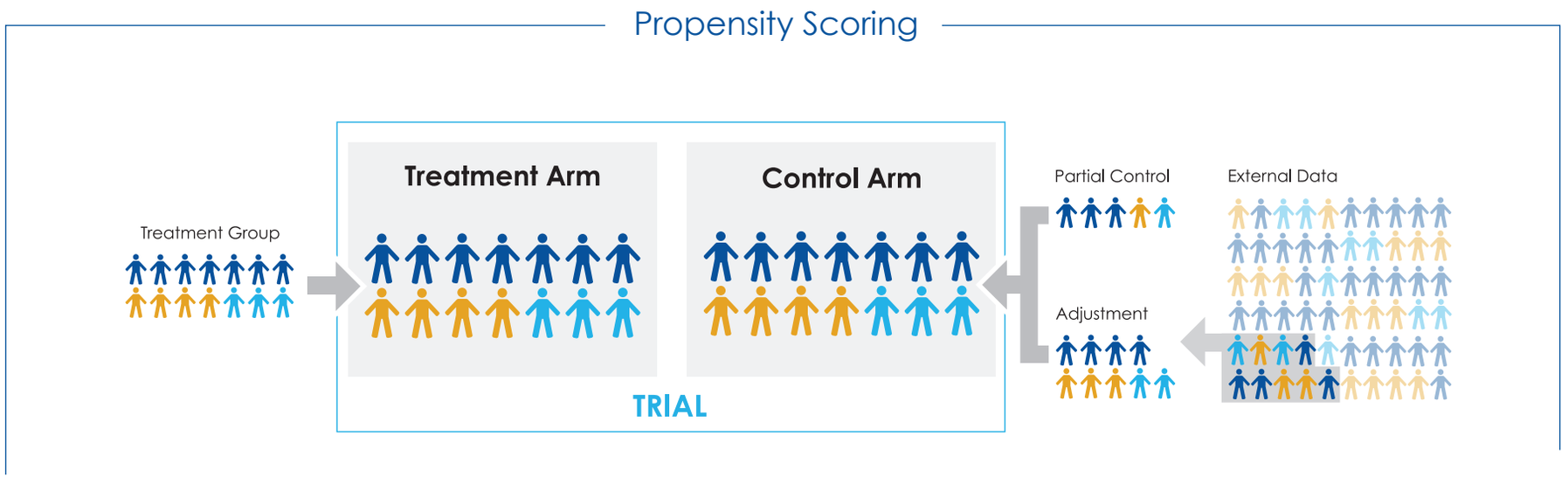


PROPENSITY SCORING METHODS

Unlike the experimental setting of the randomized clinical trial, the likelihood of a sampling bias is higher when using real world data. Such bias occurs when the population tested during an investigative study deviates from the demographics of the population intended for the new therapy.

Adjusting using propensity scores is a way to quantify and minimize such bias. The propensity score, as defined by Rosenbaum & Rubin, refers to the probability of enrollment into a given arm of a trial. In a traditional two-armed randomized clinical trial, there might be a propensity score 0.5 for each of the two arms. If a patient with a given trait (e.g. age, race, etc.), is more likely to enroll in one arm rather than the other, the propensity score deviates from 0.5. This reflects bias in a way that is quantitatively intuitive.

When working with external control data, propensity matching is typically accomplished by simply deriving propensity scores for how likely each individual is to be from the similar population as individuals receiving the treatment arm, and then weighting control individuals according to their propensity score.



BAYESIAN DYNAMIC BORROWING

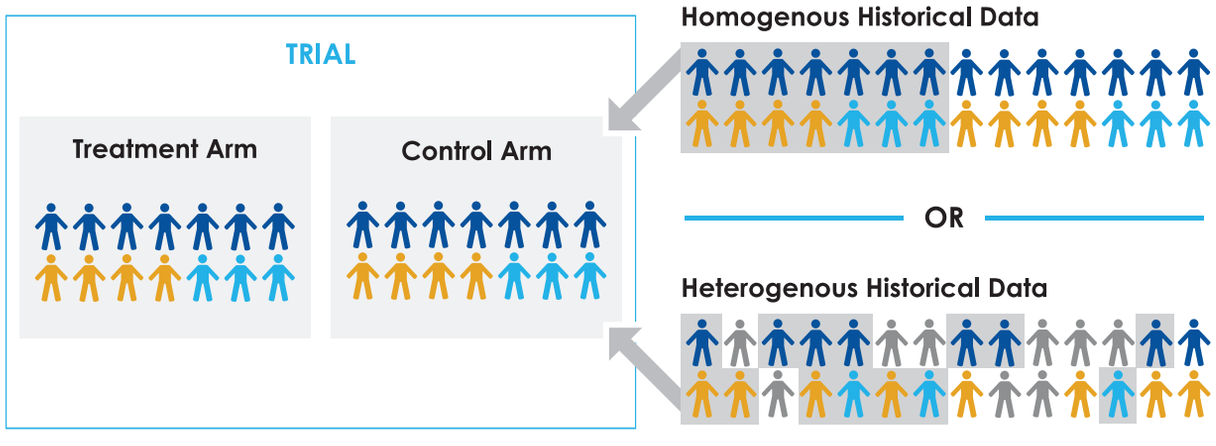
When combining historical datasets to create an SCA, researchers might discover that the historical data involve patients who are either homogenous or heterogenous to those enrolling in the new trial. Homogeneity and heterogeneity describe how similar external data are to the needs of the current trial. The Ibrance study discussed on Page 15 is an instance where external data reflected similarities in clinical endpoint, population demographic and other variables (i.e. highly homogenous to the data that would have been collected had there been a placebo-controlled study). On the other hand, sometimes external data reflect a slightly different population sample (i.e. different age, sex or race), or perhaps identify somewhat different clinical endpoints. In these instances of heterogeneity, further statistical adjustment must be incorporated. Note that while homogenous data contribute less bias than a heterogenous dataset might have, both are susceptible to some bias for which the analysis must account.

BAYESIAN DYNAMIC BORROWING

When real world data reflect a patient population homogenous to those currently enrolling, these data are easier to combine and the comparison can proceed as any other comparative analysis. When there is heterogeneity, the historical data might have to be weighted to reflect the fact that the historical data are biased. The solution, in such cases, is to add weights to the historical datasets, using a **Bayesian hierarchical model**.

Suppose we have two datasets, one which reflects the desired population and endpoint, the other which reflects the desired endpoint but not the population. The conclusions of the investigative study must account for the fact that the first dataset is giving us information more suited to the study's needs and statisticians would not treat these two datasets the same way. Weighting is therefore a technique which takes into greater account information from the data set better suited for the study.

Homogenous & heterogenous historical data can be combined to create the control arm

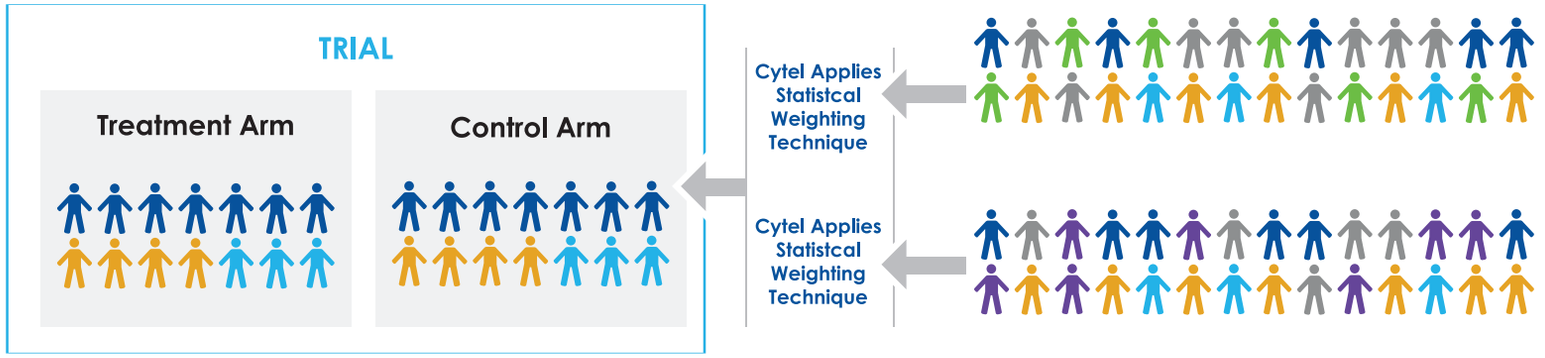


BAYESIAN DYNAMIC BORROWING

Bayesian models offer a flexible way of incorporating historical controls in the analysis of trial data (whether single arm and randomized). In the context of SCAs, one popular utilization of Bayesian models is **Bayesian Dynamic Borrowing**. **This approach can particularly create efficiencies in clinical trials as fewer patients are randomized to control than the experimental treatment**, but to make up for the sparse control information, external controls supplement the strength of the concurrent control to provide the same statistical efficiency. Here the historical or external control datasets are incorporated via Bayesian prior, and the final control arm is constructed by enrolling some new patients into the study (see next case study).

It should be noted that no statistical model can offset poor data quality. One of the key first steps, therefore, when considering building an SCA to support a single arm submission is the evaluation of the available data. Datasets should not be combined if a source of data for the SCA does not meet the needs of the comparison with the single arm trial results. A continuous dialogue is needed between RWD, statistical and clinical development experts at the time of planning this part of the clinical development program.

Statistical Weighting Technique



NON-SMALL CELL LUNG CANCER

An SCA is sometimes constructed by combining data from newly enrolled patients in the control arm of the trial, with historical ones using Bayesian methods. This enables fewer patients to be enrolled into the control group and optimizes the use of data already collected.

Bayesian Dynamic Borrowing

What is the ideal ratio for new to old patients, when constructing a synthetic control group? Using Bayesian Dynamic Borrowing, Dron et al., re-analysed a non-small cell lung cancer trial to discover that changes in ratio of new to historical patients have little effect on the statistical rigor of the trial. The original trial recruited 734 patients to the control arm. Dron et al.'s findings reveal that the trial could have enrolled 440 new patients, and still achieved similar results. While situations with heterogeneous historical datasets require more patients, even in what they viewed as high-risk combinations of historical and new enrollments, new enrollments decreased significantly.

Are you considering a choice between placebo controlled and single arm trial?

The use of single arm trials and SCAs requires careful planning and consideration. Here we provide a few typical questions to ask when determining if to explore this new approach.

1. Is your trial population well defined? For example are you in a rare disease space, or perhaps are trying to stratify certain subpopulations within a larger therapeutic group using a well-recognized biomarker?

2. If you are running a single arm trial, have you decided what type of additional evidence to provide as part of your regulatory submission? Are there existing data readily available that match your trial population? Considerations might include whether data were collected the same way, how homogeneous the existing data are to the single arm trial population, or the type of statistical adjustments necessary to correct for biases.

3. Are there ethical reasons not to enroll patients into a control arm?

4. Have the clinical outcomes in the new single arm trial been measured by other trials?

REFERENCES

1) Tibau, Ariadna, et al. "Magnitude of clinical benefit of cancer drugs approved by the US Food and Drug Administration." JNCI: Journal of the National Cancer Institute 110.5 (2018): 486-492.

2) Collins, Rory, et al. "The Magic of Randomization versus the Myth of Real-World Evidence." (2020): 674-678.

3) Baumfeld Andre, Elodie, et al. "Trial designs using real-world data: The changing landscape of the regulatory approval process." Pharmacoepidemiology and Drug Safety (2019).

4) Soria, Jean-Charles, et al. "Osimertinib in untreated EGFR-mutated advanced non–small-cell lung cancer." New England journal of medicine 378.2 (2018): 113-125.

5)https://www.astrazeneca.com/media-centre/press-releases/2018/us-fda-approves-tagrisso-as-1st-line-treatment-for-EGFR-mutated-non-small-cell-lung-cancer.html

6) Drilon, Alexander, et al. "Efficacy of larotrectinib in TRK fusion–positive cancers in adults and children." New England Journal of Medicine 378.8 (2018): 731-739.

7)https://www.fda.gov/drugs/fda-approves-larotrectinib-solid-tumors-ntrk-gene-fusions-0

8)https://www.pfizer.com/news/press-release/press-release-detail/u_s_fda_approves_ibrance_palbociclib_for_the_treatment_of_men_with_hr_her2_metastatic_breast_cancer

9) https://www.avexis.com/content/pdf/press_release.pdf

10)https://www.fda.gov/news-events/press-announcements/fda-approves-first-treatment-form-batten-disease

11) Rosenbaum, Paul R., and Donald B. Rubin. "The central role of the propensity score in observational studies for causal effects." Biometrika 70.1 (1983): 41-55.

12) Dron L, Golchi S, Hsu G, Thorlund K. Minimizing control group allocation in randomized trials using dynamic borrowing of external control data - An application to second line therapy for non-small cell lung cancer. Contemporary Clinical Trials Communications 2019; 16: 100446.

ABOUT CYTEL

As a pioneer in evidence generation, with deep expertise in advanced analytical solutions, Cytel is uniquely equipped to unlock the value from increasingly complex data. Life sciences companies count on Cytel to deliver exceptional insight, minimize trial risk, and accelerate the development of promising new medicines that improve human life. Cytel has specialized teams with expertise in analyzing real-world and novel datasets and creating tailored real-world evidence designs to meet the disparate needs of clinical, medical affairs and market access audiences.

For more information on Cytel, visit www.cytel.com

www.cytel.com

Cytel