POPULATION ENRICHMENT DESIGNS FOR ADAPTIVE CLINICAL TRIALS

An Aptiv Solutions White Paper
EXECUTIVE SUMMARY

The increasing pressure on governments caused by the spiraling healthcare costs is leading to a growing demand by payers for ‘value for money’ from new medicines reaching the market. This is widely recognized as a serious global issue and with an aging population, the demand for cost-effective healthcare will continue to rise. The pharmaceutical industry recognizes the impact this will have on pricing and reimbursement, and the challenge facing industry is to develop differentiated medicines, including novel targeted therapies, where effectiveness is clearly demonstrated at registration.

The real challenge in developing differentiated products lies in maximizing effectiveness and simultaneously the potential for commercial return. The traditional strategy of maximizing commercial return by targeting broad populations is fraught with high levels of risk in the face of emerging reimbursement challenges. For example, while it makes sense commercially to target a broad population should the drug work for all, if instead a drug works best only in a sub-population, targeting a broad population will have the effect of diluting the overall treatment effect. In the worst case this strategy may cause the trial to fail, resulting in costly rework, and a significant delay in getting the product to market. However, perhaps more importantly this represents a delay in bringing an effective medicine to those patients who would most benefit from the treatment. By contrast, it is equally unattractive, both ethically and commercially, to unnecessarily target a narrow patient population.

In the past, options for evaluating sub-populations were mostly undertaken as post hoc analyses after the trial was complete, an approach that was geared towards hypothesis generation and not acceptable in ensuring its validity by strong control of the FWER.

POPULATION ENRICHMENT DESIGNS

Enrichment designs are usually more efficient than selecting treatment arms, the enrichment design selection of new adaptive design methodologies. Reinforcement strategies for Clinical Trials to Support Proof-of-Concept, dose-ranging and confirmatory trials with adaptation. Vlad has played a major role in the design of real time learning in clinical drug development through the use of innovative approaches such as model-based drug development, and the design, implementation and execution of adaptive clinical trials. Prior to joining the Pharmaceutical Industry, Vlad was a Research Assistant Professor at the University of Rochester, NY. Before that, he has a 15-year record of distinguished service at prestigious research institutions in Moldova, Russia, Italy and Germany. Vlad received his Ph.D. in Probability Theory and Mathematical Statistics from the Steklov Mathematical Institute, Moscow in 1988.

GERNOT WASSMER: Senior Vice President, Software Architecture

Gernot Wassmer is currently Senior Vice President and Chief Software Architect in the Aptiv Solutions Innovation Center. He is also an adjunct Professor for Biostatistics at the Institute of Medical Statistics at the University of Cologne. Together with Reinhard Eisert Gernot created and developed the ADDPLAN® software. Gernot’s major statistical interests are in procedures for group sequential and adaptive designs in clinical trials, and he has authored over 80 methodological and application articles. Since 1995, Gernot has been developing statistical methods for confirmatory adaptive designs and many of these adaptive methods are now widely used in practice. Prior to working in industry, Gernot served as a Research Fellow at the Institute of Statistics, University of Munich, at the Institute for Epidemiology, GSF Neuherberg, and at the Institute of Medical Statistics, University of Cologne, and has worked as a statistical consultant for the pharmaceutical industry for over 20 years. Gernot received both his Masters in Statistics (1987) and his PhD (1993) from the University of Munich. In 1999, he earned his professorial standing from the University of Cologne.

BIBLIOGRAPHY


These exhibits represent only a few of the total number of population enrichment studies that can be designed using an adaptive approach. Such designs will impact the speed and consistency with which targeted therapies and personalized medicine are developed and approved.

ABOUT THE AUTHORS

REINHARD EISEBIT: Executive Vice-President and Head Innovation Center

Reinhard Eisebitt heads up the Aptiv Solutions Innovation Center which is responsible for development of new adaptive design methodologies. Reinhard was previously Managing Director of ClinResearch GmbH which became part of Aptiv Solutions in 2010. ClinResearch specialized in the design and execution of adaptive clinical trials and has conducted more than 100 adaptive studies. In 2001, Reinhard, together with Gernot Wassmer, founded ADDPLAN GmbH which created and launched the first commercial software package for the design, simulation, and analysis of adaptive clinical trials. Prior to ClinResearch and ADDPLAN, Reinhard spent nine years as Director Biostatistics in the German CRO industry. Reinhard holds a Master degree in Mathematics from Albertus-Magnus-University, Cologne, Germany.

DR. VLADIMIR DRAGALIN: Senior Vice President, Clinical Trial Innovation Strategies

Vlad Draganlin is a world-leading expert in adaptive trial design and execution and joined Aptiv Solutions in February 2011. Vlad is an elected member of PhRMA Biostatistics and Data Management Technical Group, and a Member of the American Statistical Association, the Drug Information Association, and an Associate Editor of Journal of Biopharmaceutical Statistics. Vlad was actively involved in the PhRMA Working Group on Adaptive Designs and the heterogeneous due to observable clinical characteristics and/or unobservable underlying genomic and epigenetic characteristics, or the experimental therapy is tailored to a specific mechanism of action. In relation to this, it is generally accepted that the development of biomarkers that define specific subsets of disease will enable a shift from empirical medicine to precision medicine. Thus, in the near future it is conceivable that there will be a move away from the concept of “one size fits all” and a shift to personalized medicine resulting in delivery of the right medicine, to the right patient, at the right dose, at the right time. The overall objective is an improvement in patient care and a reduction in healthcare cost.

Extension from the conventional single population design to a setting that encompasses several possible patient sub-populations will allow more informative evaluation of patients having different degrees of responsiveness to therapy. In adequate and well-controlled clinical trials, the incorporation of an adaptive design which involves a prospectively planned selection of patient sub-populations with a higher response to a particular therapy can generate an appealing conceptual framework from both the sponsor’s and patient’s perspective in addressing personalized medicine. These new adaptive designs, called patient enrichment or population enrichment designs, allow pre-planned mid-stream modifications of the study hypothesis including reallocation of patients and re-estimation of the sample size.

Factors used to limit the study population to patients believed more likely to benefit from the experimental therapy are termed enrichment factors. Enrichment factors may be predictive biomarkers, or they may be biomarkers, clinico-pathologic characteristics or demographic characteristics associated with a predictive biomarker or with the molecular target for a specific therapeutic agent.

The main purpose of using an enrichment biomarker in drug development is to improve the chance that the drug will show benefit in the tested subgroup of patients and hence more quickly establish that the drug is worth pursuing further. The lower the proportion of truly benefiting patients, the more advantageous it is to consider studying an enriched population. However, instead of limiting the enrollment only to the narrow sub-population of interest, prospectively specified adaptive designs may be used to consider the effect of the experimental treatment both in the wider patient population under investigation, and in various sub-populations.

STATISTICAL METHODOLOGY

The statistical methodology is based on multiple hypotheses testing within multi-stage adaptive designs. In order to control the Familywise Type I Error Rate (FWER) in a strong sense, the closed testing principle is used for defining the corresponding multiple testing procedures to test the treatment effect in pre-specified sub-populations. The flexible p-value combination test strategy is used to adjust for sub-population selection as well as other adaptations at the interim like early stopping or sample size re-estimation. This strategy was originally proposed by Bauer and Kühn (1994) in which the authors used the idea of combining the p-values for two-stage trials through the use of a predefined combination test. Lehmacher and Wassmer (1999) generalized the method for multi-stage designs with the use of the inverse normal method which uses the critical values of a specific group sequential test design. The resulting group sequential adaptive design is a powerful and flexible supplement to classical group sequential designs. In addition, the inverse normal method offers much broader types of adaptation, such as multiple comparisons and selection of treatment arms, selection of endpoints, and sub-group selection of patients within population enrichment designs. The approach proposed by Cui et al. (1999) turns out to be equivalent, though less general, than the inverse normal testing principle, since this approach was only proposed for sample size reassessment strategies. With the combination testing principle, the definition of confidence intervals and p-values is straightforward.
and can also be applied for the multiple hypotheses closed testing situation as proposed by Posch et al. (2005).

The statistical methodology for enrichment designs is very similar to the statistical methodology of seamless Phase II/III designs incorporated in the adaptive design methodology of ADDPLAN® MC. While in a seamless Phase II/III design the adaptation pertains to selection of treatment arms, the enrichment design continues seamlessly either in the sub-population(s) of patients or in the whole population on the basis of data obtained in the beginning of the trial. In a two-stage design, at the end of the trial, the data from both stages are combined in the final analysis to assess the efficacy of the selected sub-population(s), preserving its validity by strong control of the FWER. Enrichment designs are usually more efficient than separate Phase II and Phase III programs in that fewer patients are required to achieve a given program-level power. The benefit arises from the inclusion of the first stage data on the selected sub-populations, suitably adjusted for multiplicity, in the final analysis at the end of the trial.

The new opportunities, scope and challenges of enrichment strategies in drug development have been presented by Temple in 1994, and the statistical methodology was further developed by Simon and his colleagues in a series of publications. Recent improvements in the efficiency of the enrichment design have been achieved by incorporating the correlation between test statistics based on the known prevalence of the sub-population.

These initial developments have been focused on enrichment design with a fixed sample size. Methodology of adaptive enrichment design combining the flexible p-value combination tests with the closed testing principle was initiated by Kieser et al. (1999), and further developed by Brannath et al. (2009), Wang et al. (2007, 2009), and Wassmer (2011). The methodology has been extended to the case when the sub-population selection is based on a short-term endpoint (e.g. progression free survival) while the final primary analysis is based on a correlated long-term endpoint (e.g. overall survival) by Jenkins et al. (2011) and Friede et al. (2012). Application of adaptive enrichment designs in different therapeutic areas have been published recently (see Mehta et al. 2009, 2011, Tournoux et al. 2011), but many more examples are expected, especially after the release of the Draft Guidance for Industry on Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products by the US Food and Drug Administration.

IMPLEMENTATION OF ENRICHMENT DESIGNS

The adaptive design software ADDPLAN® PE provides an extension of the adaptive options of an enrichment design allowing for multiple stages, sample size re-estimation, early stopping for efficacy or futility using the adaptive group sequential approach. Inverse normal and Fisher’s p-value combination testing strategies, several multiple testing procedures and many alternative sub-population(s) selection rules have been considered.

Enrichment designs have several elements that can be used for design adaptation: number of stages, number of pre-specified sub-populations, the multiple comparison testing strategy, the interaction test, the sequential stopping rule, the sub-population selection rule, and sample size re-estimation rule. Generally, in ADDPLAN® PE, an enrichment design can be assessed through the use of a powerful simulation tool, or undertaken via a real data analysis. The simulation tool allows combining different versions of these elements and creating different variants of enrichment designs. The powerful output capabilities facilitate the comparison of these variants under a set of simulation scenarios using many generated graphs and tables for their operating characteristics. The tool can also be used for fine-tuning the parameters of a given design to achieve the best performance, for example by modifying the threshold condition of the selection rule or the target conditional power for the next stage in the sample size re-estimation rule.

Several Selection Procedures have been implemented:

- Select the sub-population with the largest effect
- Select the r sets with largest effect
- Select sets with effect compared to best not worse than $\epsilon$ estimated, where $\epsilon$ is some real number
- Select sets with effect compared to the effect in the full population not worse than $\epsilon$
- Fix the $i^{th}$ sub-population to be selected
- Deselect sets for which effect is smaller than $\epsilon$
- The p-q-selecion rule (originally suggested by Bretz and Maurer for seamless Phase II/III design of treatment selection)

Generally, the effect can be measured either as estimated treatment difference or as test statistic.

The sample size reassessment method is based on conditional power. The corresponding conditional power is calculated either using the observed treatment effect (ML estimate) or the assumed treatment effect. For testing means, the assumed treatment effect is the standardized treatment effect in the full population. For testing rates, the assumed rates in the full population for the control and experimental treatment are required. For the survival case, the assumed hazard ratio in the full population is used. Users can specify the minimum relative (to the predefined value) reduction and maximum relative increase of sample size per stage for the experimental treatment arm. These two values determine the minimum and the maximum number of patients in the trial.

For the analysis module, summary data (means, rates, log rank statistics, etc.) are entered either stage wise or in a cumulative way. The calculated output shows the test decision together with suitable summary statistics like treatment effect estimates, confidence intervals, and overall p-values and sample size calculations based on conditional power are provided. Graphical illustration of the conditional power for a specified range of parameter values together with the likelihood function for the unknown parameter provides a convenient way for assessing study results at interim stages.

EXAMPLES OF ENRICHMENT DESIGNS

Exhibits 1 to 4 provide examples of the use of enrichment designs.

Exhibit 1 (two sub-populations of interest) considers a case in which the primary efficacy analysis is pre-planned on all randomized patients (full population, F), on a subgroup of patients defined by a baseline characteristic (sub-population, S1), and on a gene signature derived subgroup of patients (sub-population, S2).

However, the objective of the trial might also include a test of the treatment effect in the sub-population with both the desired baseline characteristic and good gene signature (see Exhibit 2 – ‘three sub-populations of interest’). Such a case can be considered in ADDPLAN® PE using the following four analysis sets: S1 - the sub-population with the desired baseline characteristic, S2 - the sub-population with the good gene signature, S3 - the sub-population with both the desired baseline characteristic and good gene signature, and F - the full population.

A further example is shown in Exhibit 3 (‘two non-overlapping sub-populations’) for a global clinical trial with pre-planned efficacy analyses for both patients from US only (sub-population, S1) and patients from Japan only (sub-population, S2), besides the efficacy analysis for all patients (full population, F). The fourth example is a case of a trial in HER2 negative breast cancer patients (Exhibit 4 - ‘two nested sub-populations of interest’). There is a sub-population of triple negative breast cancer patients (sub-population, S1) that may benefit better from a new drug and among the triple negative patients those with a specific gene signature (sub-population, S2) may benefit even more.
Methodology of adaptive enrichment design combin-
These exhibits represent only a few of the total number of population enrichment studies that can be designed using an adaptive approach. Such designs will impact the speed and consistency with which targeted therapies and personalized medicine are developed and approved.

ABOUT THE AUTHORS

REINHARD EISEBIT: Executive Vice-President and Head Innovation Center

Reinhard Eisebit heads up the Aptiv Solutions Innovation Center which is responsible for development of new adaptive design methodologies. Reinhard was previously Managing Director of ClinResearch GmbH which became part of Aptiv Solutions in 2010. ClinResearch specialized in the design and execution of adaptive clinical trials and has conducted more than 100 adaptive studies. In 2001, Reinhard, together with Gernot Wassmer, founded ADDPLAN GmbH which created and launched the first commercial software package for the design, simulation, and analysis of adaptive clinical trials. Prior to ClinResearch and ADDPLAN, Reinhard spent nine years as Director Biostatistics in the German CRO industry. Reinhard holds a Master degree in Mathematics from Albertus-Magnus-University, Cologne, Germany.

DR. VLADIMIR DRAGALIN: Senior Vice President, Clinical Trial Innovation Strategies

Vlad Dragalin is a world-leading expert in adaptive trial design and execution and joined Aptiv Solutions in February 2011. Vlad is an elected member of PhRMA Biostatistics and Data Management Technical Group, and a Member of the American Statistical Association, the Drug Information Association, and an Associate Editor of Journal of Biopharmaceutical Statistics. Vlad was actively involved in the PhRMA Working Group on Adaptive Designs and the heterogeneous due to observable clinical characteristics and/or unobservable underlying genomic and epigenomic characteristics, or the experimental therapy is tailored to a specific mechanism of action. In relation to this, it is generally accepted that the development of biomarkers that define specific subsets of disease will enable a shift from empirical medicine to precision medicine. Thus, in the near future it is conceivable that there will be a move away from the concept of “one size fits all” and a shift to personalized medicine resulting in delivery of the right medicine, to the right patient, at the right dose, at the right time. The overall objective is an improvement in patient care and a reduction in healthcare cost.

Extension from the conventional single population design to a setting that encompasses several possible patient sub-populations will allow more informative evaluation of patients having different degrees of responsiveness to therapy. In adequate and well-controlled clinical trials, the incorporation of an adaptive design which involves a prospectively planned selection of patient sub-populations with a higher response to a particular therapy can generate an appealing conceptual framework from both the sponsor’s and patient’s perspective in addressing personalized medicine. These new adaptive designs, called patient enrichment or population enrichment designs, allow pre-planned mid-stream modifications of the study hypothesis including reallocation of patients and re-estimation of the sample size.

Factors used to limit the study population to patients believed more likely to benefit from the experimental therapy are termed enrichment factors. Enrichment factors may be predictive biomarkers, or they may be biomarkers, clinicopathologic characteristics or demographic characteristics associated with a predictive biomarker or with the molecular target for a specific therapeutic agent.

The main purpose of using an enrichment biomarker in drug development is to improve the chance that the drug will show benefit in the tested subgroup of patients and hence more quickly establish that the drug is worth pursuing further. The lower the proportion of truly benefiting patients, the more advantageous it is to consider studying an enriched population. However, instead of limiting the enrollment only to the narrow sub-population of interest, prospectively specified adaptive designs may be used to consider the effect of the experimental treatment both in the wider patient population under investigation, and in various sub-populations.

STATISTICAL METHODOLOGY

The statistical methodology is based on multiple hypotheses testing within multi-stage adaptive designs. In order to control the Familywise Type I Error Rate (FWER) in a strong sense, the closed testing principle is used for defining the corresponding multiple testing procedures to test the treatment effect in pre-specified sub-populations. The flexible p-value combination test strategy is used to adjust for sub-population selection as well as other adaptations at the interim like early stopping or sample size re-estimation. This strategy was originally proposed by Bauer and Köhne (1994) in which the authors used the idea of combining the p-values for two-stage trials through the use of a predefined combination test. Lehmann and Wassmer (1999) generalized the method for multi-stage designs with the use of the inverse normal method which uses the critical values of a specific group sequential test design. The resulting group sequential adaptive design is a powerful and flexible supplement to classical group sequential designs. In addition, the inverse normal method offers much broader types of adaptation, such as multiple comparisons and selection of treatment arms, selection of endpoints, and sub-group selection of patients within population enrichment designs. The approach proposed by Cui et al. (1999) turns out to be equivalent, though less general, than the inverse normal testing principle, since this approach was only proposed for sample size reassessment strategies. With the combination testing principle, the definition of confidence intervals and p-values is straightforward.
The increasing pressure on governments caused by the spiraling healthcare costs is leading to a growing demand by payers for ‘value for money’ from new medicines reaching the market. This is widely recognized as a serious global issue and with an aging population, the demand for cost effective healthcare will continue to rise. The pharmaceutical industry recognizes the impact this will have on pricing and reimbursement, and the challenge facing industry is to develop differentiated medicines, including novel targeted therapies, where effectiveness is clearly demonstrated at registration.

The real challenge in developing differentiated products lies in maximizing effectiveness and simultaneously the potential for commercial return. The traditional strategy of maximizing commercial return by targeting broad populations is fraught with high levels of risk in the face of emerging reimbursement challenges. For example, while it makes sense commercially to target a broad population should the drug work for all, if instead a drug works best only in a sub-population, targeting a broad population will have the effect of diluting the overall treatment effect. In the worst case this strategy may cause the trial to fail, resulting in costly rework, and a significant delay in getting the product to market. However, perhaps more importantly this represents a delay in bringing an effective medicine to those patients who would most benefit from the treatment. By contrast, it is equally unattractive, both ethically and commercially, to unnecessarily target a narrow patient population.

In the past, options for evaluating sub-populations were mostly undertaken as post hoc analyses after the trial was complete, an approach that was geared towards hypothesis generation and not acceptable in establishing a body of evidence considered acceptable for registration. However an approach to trial design called population enrichment is available and offers a potential solution. This is an adaptive design approach that specifically targets the issue of population selection, and while this has application in early phase trials, it is also appropriate for registration studies. It is generally accepted that adaptive design trials dramatically improve the efficiency of the clinical development process, and the ability to design an adaptive trial that studies the effectiveness of novel and existing treatments in pre-defined sub-populations of patients. This represents a significant step forward in delivering on the promise of personalized medicine. The adaptive approach to population enrichment is discussed in more detail in this white paper.

Importantly, the population enrichment methodologies described in this paper have been embodied in a fully validated adaptive design software called ADDPLAN® PE, which is available to both academic and industry groups. The application of this novel population enrichment methodology in the clinical development process will dramatically enhance the ability of both regulators and industry to realize the true potential of targeted therapy with the ultimate objective of improving patient care and reducing healthcare spending.

**BACKGROUND – ENRICHMENT DESIGNS AND ADAPTIVE CLINICAL TRIALS**

A recent report from the National Academy of Sciences calls for precision medicine - the use of genomic, epigenomic, exposure, and other data to define individual patterns of disease, potentially leading to better individual treatment. Precision medicine couples established clinical-pathological indexes with state of the art molecular profiling to create diagnostic, prognostic and therapeutic strategies tailored for specific groups of patients.

The notion of “one size fits all” surrounding the conventional design of clinical trials has been challenged, particularly when the disease is considered

PhRMA Working Group on Adaptive Dose Ranging Studies. Prior to joining Aptiv Solutions, Vlad worked in both the CRO and pharmaceutical industry for more than 12 years with companies such as Wyeth Research and GSK. Vlad has significant experience in the implementation of innovative approaches to drug development in “Learn and Confirm” settings including Proof-of-Concept, dose-ranging and confirmatory trials with adaptation. Vlad has played a major role in the application of real time learning in clinical drug development through the use of innovative approaches such as model-based drug development, and the design, implementation and execution of adaptive clinical trials. Prior to joining the Pharmaceutical Industry, Vlad was a Research Assistant Professor at the University of Rochester, NY. Before that, he has a 15-year record of distinguished service at prestigious research institutions in Moldova, Russia, Italy and Germany. Vlad received his Ph.D. in Probability Theory and Mathematical Statistics from the Steklov Mathematical Institute, Moscow in 1988.

**GERNOT WASSMER: Senior Vice President, Software Architecture**

Gernot Wassmer is currently Senior Vice President and Chief Software Architect in the Aptiv Solutions Innovation Center. He is also an adjunct Professor for Biostatistics at the Institute of Medical Statistics at the University of Cologne. Together with Reinhard Einsele Gernot created and developed the ADDPLAN® software. Gernot’s major statistical interests are in procedures for group sequential and adaptive designs in clinical trials, and he has authored over 80 methodological and application articles. Since 1995, Gernot has been developing statistical methods for confirmatory adaptive designs and many of these adaptive methods are now widely used in practice. Prior to working in industry, Gernot served as a Research Fellow at the Institute of Statistics, University of Munich, at the Institute for Epidemiology, GSF Neuhberg, and at the Institute of Medical Statistics, University of Cologne, and has worked as a statistical consultant for the pharmaceutical industry for over 20 years. Gernot received both his Masters in Statistics (1987) and his PhD (1993) from the University of Munich. In 1999, he earned his professorial standing from the University of Cologne.

**BIBLIOGRAPHY**


POPULATION ENRICHMENT DESIGNS FOR ADAPTIVE CLINICAL TRIALS

An Aptiv Solutions White Paper