NEXT-GENERATION DATA ANALYSIS SOFTWARE AIDS FASTER DRUG DEVELOPMENT

Researchers studying variability in drug concentrations in patient populations use S-PLUS software to accelerate data analysis to improve the understanding of drug safety and efficacy.

Faster and better drug development can reduce the time taken to make safe and effective drugs available to those who need them most. The pharmaceutical scientists in academia, the pharmaceutical industry, and regulatory agencies such as the Food and Drug Administration (FDA) have been greatly interested in the pharmacometric techniques to understand drug safety and efficacy.

Pharmacometrics is a specialty within clinical pharmacology that addresses the development and application of statistical and mathematical techniques to assess the relationship between drug dose and concentration [pharmacokinetics (PK)], PK and the effect of a drug on the body [pharmacodynamics (PD)], PK and safety/efficacy, and the efficient implementation of model-based strategies in drug development. Thus, pharmacometricians use mathematical and statistical techniques in PK/PD to promote rational drug development with rational pharmacotherapy (right dose for the right patient) as the end product).

Drug discovery and development can be time consuming and expensive. Each year, the global pharmaceutical industry spends at least $10 billion on clinical trials. To counter this expense, next-generation statistical software provides pharmaceutical scientists with the tools they need to discover hidden knowledge buried in data collected during clinical trials. The use of software tools with powerful graphical and modeling capabilities, such as S-PLUS, result in better and faster drug development. The S-PLUS software is based on S-language, which is specifically designed for data visualization and exploration, statistical modeling and programming with data.

Population pharmacokinetics is the study of the sources and correlates of variability in plasma drug concentrations among individuals who are preferably the target patient population receiving clinically relevant doses of a drug of interest. Patient demographic features or pathophysiological variables can be helpful in explaining alterations in the dose-concentration relationship. Population pharmacokinetics applies an exploratory approach to drug development. Integrated information on pharmacokinetics from concentration to time data collected during clinical trials can be gathered using pharmacokinetic techniques. Statistical tools allow the pharmaceutical scientist to explore and discover important hidden knowledge in clinical data sets efficiently.

Exploratory Data Analysis

Using graphical and modern statistical techniques to explore clinical data can isolate and reveal patterns and features in the data. Both of these techniques can uncover unexpected departures from hypothesized models (e.g., departure from a normal distribution). An important element of this exploratory approach is the software's flexibility to tailor the analysis to the data structure and to respond to patterns uncovered by successive analysis. Exploratory data analysis (EDA) is essential to population pharmacokinetic modeling, and new statistical software tools allow the data analyst to visualize and prototype data more efficiently.

According to Dr. Ene Ette, a leading pharmaceutical scientist with Vertex Pharmaceuticals, Inc., raw computer power, combined with sophisticated data analysis tools like S-PLUS, brings valuable findings to the drug development team more quickly with safety and efficacy in mind. "One of the difficult tasks for a pharmacometrician is to convey findings from pharmacometric analyses to clinicians and other members of the drug development team. Failure to communicate these findings successfully can hamper drug research and development. S-PLUS offers easy-to-read, attractive graphics (Trellis graphics) that can be used to communicate results effectively to the drug development team facilitating the development process," says Dr. Ette.

Dr. Ette heads a research team that integrates the population pharmacokinetics approach in drug development. Using state-of-the-art statistical software, his team captures, stores, and analyzes population pharmacokinetic and pharmacodynamic data sets to discover hidden knowledge about drugs in development. "We believe that data

analysis tools can enhance our capability to bring safe, effective drugs to the market sooner. S-PLUS provides the sophisticated statistical techniques we need to discover patterns and trends in clinical trial data sets and explore meaningful trends in the data we collect. With such information, better and smarter trials can be designed,' Dr. Ette remarks.

Dr. Ette's work involves using modern statistical techniques, such as those implemented in S-PLUS, to extract information hidden in population pharmacokinetic and pharmacodynamic data sets. "We selected S-PLUS because it offered powerful graphics combined with state-of-the art statistical techniques that would enable us make the most of the data. The package allowed us to effectively communicate the results to the drug development team," Dr. Ette remarks. He believes state-of-the-art data analysis software helps researchers to explore, isolate and analyze patterns in population pharmacokinetic and pharmacodynamic data sets, thereby aiding the discovery of information hidden in the data.

Most population analysis procedures are based on explicit data assumptions. The validity of data analysis depends upon the validity of the assumptions. Exploratory data analysis provides powerful diagnostic tools for confirming assumptions or, when the assumptions are not met, for suggesting corrective actions.

Acceleration of Population Pharmacokinetic Analysis

Population pharmacokinetic analysis is time consuming, and software tools can reduce the time spent exploring and analyzing data.

Analytic results are important for designing studies and for filing new drug applications (NDA). "The use of EDA as part of population pharmacokinetic analysis reduces the time spent in performing more detailed analysis. It provides the analyst the information needed for 'a straight to the point analysis,' so they don't have to test all of the possible relationships between pharmacokinetic parameters and possible predictor variables (covariates). S-PLUS has proven to be a good resource for prototyping pharmacokinetic experimental designs," says Dr. Ette.

For example, when Dr. Ette explored pharmacokinetic design in a recent study he pooled data from 138 subjects in six studies involving both sexes, approximately 16 percent of which were renal impaired. They ranged in age from pediatrics to the elderly. These subjects were treated with a test drug in different studies and the data was used to develop a model characterizing the disposition of the drug (i.e., time course of drug) in the patient population. "S-PLUS provided us with the tools (data visualization and modern regression techniques) we needed for the initial exploratory examination and analyses of the data to discover patterns and relationships that set the stage for the analyses procedure," says Dr. Ette.

Further, an exploratory examination of the raw drug concentration data revealed the probable presence of two subpopulations (see figure 1). Dr. Ette then fit a model to the data to characterize the dose-concentration relationship - in the population of subjects studied. With this characterization, individual parameter estimates were generated. Dr. Ette explored his data using modeling techniques in S-PLUS to determine possible covariates that could be used to explain variability in dose-concentration relationship observed in the patients.

Researchers analyzed the distribution of the rate at which the drug was cleared from the body per unit time (termed clearance) in the study population using a probability density plot. [Drug clearance determines drug levels on multiple dosing.] This revealed the existence of two subpopulations (figure 2) as was suggested by the concentration-time profile plot. Multiple linear regression models, generalized additive models [(GAM), a modern regression technique], and tree-based models implemented in S-PLUS were very useful for selecting co-variables that could be used to explain the variability in the dose-concentration relationship. [Generalized additive models are a group of models that are as tractable as the linear model but do not force the data into unnatural scales. Separate functions are introduced to allow for nonlinearity and heterogeneous variances. GAM is closer to a reparameterization of the model than a re-expression of the response.] The existence of two subpopulations and covariates that could be used to explain this variability in dose-concentration relationship, was captured with these exploratory modeling (data fitting) techniques.

The inadequacy of the linear model is clearly seen in figure 3. The index of renal function (creatinine clearance) was found to be the major determinant of the variability in the dose-concentration relationship and the covariate explaining

the subpopulation differences. The contribution of age was judged to be of minimal importance. The dendrogram from the tree based model with the associated histogram drove home this point (see figure 4). The tree-based model confirmed the inadequacy of trying to use linear models to explain the relationship between creatinine clearance (CLCR) and clearance (CL) and confirmed the uncovered structure in the data (i.e., the existence of two subpopulations). "The whole population analysis was done in a much shorter time than is traditionally the case," Dr. Ette remarks.

Visualization of Data Analysis Results

Data analysis software such as S-PLUS, with its data visualization capability, provides scientists with a powerful tool for communicating results. Failure to deliver these findings successfully can endanger all data analyses efforts, despite the quality of the work. The efficient and timely delivery of information is critical to efficient drug development. S-PLUS conditioning plots and Trellis graphics can be very useful to scientists in communicating their results effectively to their associates. Conditioning plots can be very useful for visualizing and presenting the results of a clinical trial simulation study in which various study designs are investigated to determine the design most likely to produce the desired outcome. "S-PLUS offers interoperability with leading business reporting packages such as Microsoft Office, Excel and Power Point, providing the perfect platform for the presentation of results," says Dr. Ette.

Next Steps

Efficiency in drug development can be achieved by understanding the sources of variability and accounting through better study design, improved data analysis and better study monitoring. S-PLUS offers researchers a powerful tool for discovering knowledge hidden in study data and understanding the consequences of different study designs. This next-generation software offers scientists the flexibility and power to analyze data faster for early decision making, thereby saving time and money. The use of this software would aid faster drug development, making it possible for safer and more effective drugs to be brought to the market sooner, rather than later.

Source: Insightful Corp., 2007