

How PK and PD analyses drive dataset designs

David B. Radtke and Brian A. Moser, Eli Lilly and Company

ABSTRACT

Historically, the format of pharmacokinetic analysis datasets has been driven by the nature of time-based evaluations and the pharmacokinetic analysis programs used for these analyses. The format of these analysis datasets differs, sometimes remarkably, from those used for common statistical analyses of clinical data. Perhaps not surprisingly, attempts to standardize data collection and reporting have focused on meeting the clinical safety and efficacy analysis needs of statisticians, and have largely overlooked the corresponding structure and format of data for pharmacokinetic, pharmacodynamic, and pharmacometric (PK, PD, and PMx) evaluations. Given the increasing importance of PK, PD, and PMx analyses for dose selection and optimization en route to product registration, the need to develop an understanding of data collection, compilation, and reporting for PK, PD, and PMx analyses has never been greater. Therefore, efforts to standardize data collection and develop PK, PD, and PMx dataset standards not only hold promise for streamlining the pharmaceutical research and development paradigm, but also optimizing the regulatory review process.

APPROACHING THE CHALLENGE

Pharmacokinetic analysis and data standards are largely foreign to those outside the collective PK/PD/PMx discipline. This, in turn, has led to a highly variable and specialized approach, making it difficult for non-PK/PD/PMx personnel to support dataset construction or to adhere to a particular set of dataset standards. Many companies have groups of special programmers dedicated solely to creating analysis datasets for PK/PD/PMx analysis, and attempts to include these results within the CDISC model have met with only limited success. Consequently, there is a growing realization that data handling and data storage formats, so often driven by statistical analysis and regulatory reporting, are not conducive or even directly translatable to the PK/PD/PMx world. It's therefore valuable to highlight the differences between analytical approaches, the historical context surrounding those differences, and the underlying contrast in dataset formats. And just as any component details of an analytical method are inextricably linked to the purpose of the corresponding analysis, any discussion of data handling and differences in dataset formats also requires an understanding of the underlying science of PK, PD and PMx driving that analysis.

WHY THE DIFFERENCE?

So why is the handling, structure, and analysis of PK, PD, and PMx data so different from those of typical statistical approaches? To answer that, we need to recognize that within the world of PK/PD/PMx there are multiple types of analyses addressing similar but different goals. Those differing approaches and goals all stem from the actual purpose of PK/PD/PMx: optimizing the dose. Moreover, the goals of these analyses evolve along with the stage of drug development, culminating in the ultimate deliverable of dose optimization. To this end, pharmacokinetics focuses on the movement of a chemical substance (drug) through the body and the effects the body conditions have on that drug, its success of reaching the target of interest, and its functionality. Pharmacodynamics focuses on the subsequent body response to the presence of the drug. Pharmacometrics is a much broader field encompassing PK and PD relationships, along with other factors of interest. It is useful to note that common statistical approaches also evaluate drug-response and drug-safety relationships. The primary difference between these analyses is the focus on continuous chronological sequencing that that occurs in the PK/PD/PMx world and the snapshot, time window, or threshold evaluations that occur in the statistical world. These approaches are complementary as they apply to the development and safe evaluation of drugs, but differ greatly in their approaches, data structures, and data collection requirements.

CONVENTIONAL VS. POPULATION ANALYSIS: THE TREND TOWARD GREATER COMPLEXITY

Dataset handling and structure in PK, PD, and PMx analyses fall primarily into two main types. The first focuses on evaluation of data within individual subjects, the primary goal of which is understanding dose level as it relates to exposure and hence tolerability and safety in humans. This is typically performed in early phase clinical pharmacology and biopharmaceutics studies, and is commonly referred to as “conventional” PK/PD analysis. The second type focuses on the holistic evaluation of PK, PD, and PMx data and patient characteristics across populations of individuals, and is therefore known as “population” analysis. The population of interest is typically the target patient population, but models developed to this data can be used to extrapolate to other subpopulations or disease states. This type of analysis is most often performed in later-phase dose selection and confirmatory studies within the target population.

In addition to these basic types of analyses, historical context plays an important role in understanding the structure and evaluation of time-based data. In the early 1990’s the predominant focus of PK and PD analyses was conventional analysis in early phase studies. Very little work was being performed in the population realm, due in large part to the costs and limitations of data storage and computational power available at that time. Only a few years before my introduction to PK/PD/PMx, my department was performing calculations on the device shown in Figure 1, and some of the first personal computers purchased at my company were utilized in PK and PD analysis. In relative terms, this represented a remarkable leap forward, but paled in comparison to the distributed computing schemes and client-server approaches of today.



Figure 1. An early device for PK analysis.

A basic evaluation of time course data can provide an approximation of drug exposure (an estimate of what the body actually “sees” as opposed to the strength of the pill swallowed or the solution injected). This basic analysis uses a simplistic summation of geometrical subsections over the time course of measured drug concentrations within a given subject. Importantly, time courses such as these form the basis of all PK, PD, and PMx analysis. Figure 2 shows a typical plasma concentration over time curve,

along with a visual overlay showing the geometrical subsection used to estimate systemic exposure. Each subsection represents a time window into the subject's exposure to the drug. It is the assessment of these individual snapshots and measurements at any one time point or slice that drives the layout of data in most conventional analysis datasets.

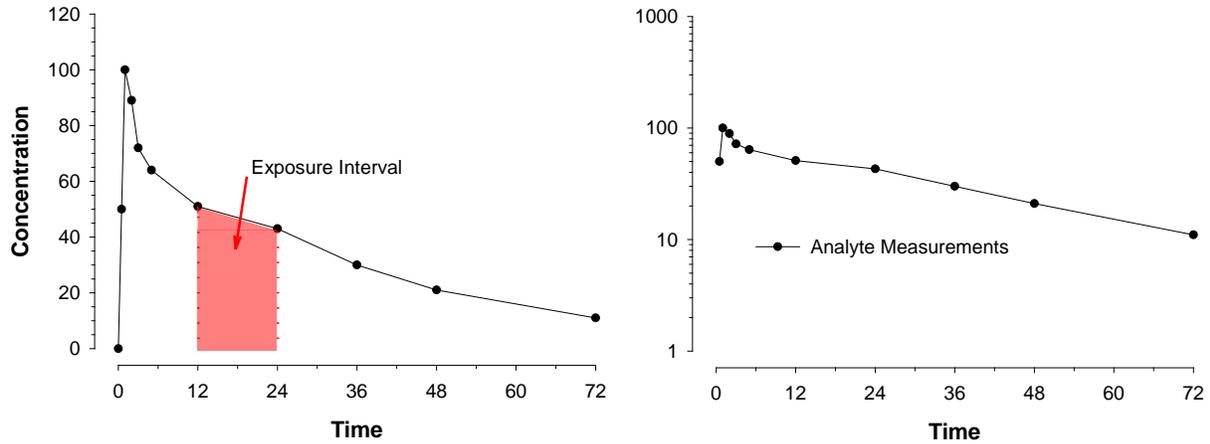


Figure 2. The typical concentration vs. time profile.

In most conventional analysis datasets, each individual time point is a separate row or record in the dataset. Figure 3 illustrates this layout, highlighting the different variables and record types. This layout makes sense when you consider that conventional analysis was originally performed on paper, long before the evolution to calculators and eventually to computer spreadsheets. Spreadsheet programs like Lotus 1-2-3™ and Microsoft Excel™ allowed scientists to organize data the way they visualized it, in slices or windows. Eventually, programmatic solutions to this common analysis type were developed around this common data layout. This was driven not by data science specialists with no knowledge or experience in PK/PD/PMx, but instead by the people performing the analysis.

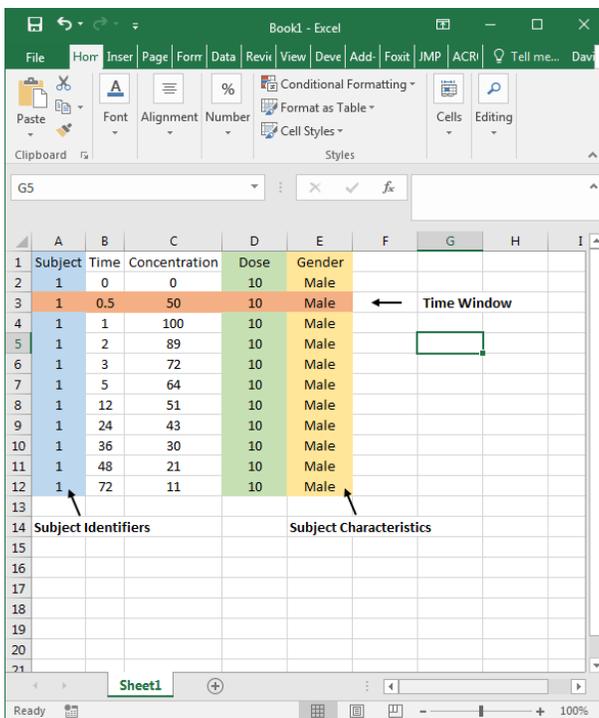


Figure 3. A simple time and concentration dataset.

As the realm of population-based techniques developed, conventional dataset construction approaches carried forward to the development of tools for population analysis. The first programs for population analysis— in which a mathematical model is fit to a series of data— were not commercial products, but were instead developed in academia and at individual institutions. Often these programs were created by the same scientists who performed conventional analysis, so it makes sense they would follow a similar data format. In fact, some of the early programs developed for population-based model development are still the accepted, present-day standards in both industrial and academic settings.

These programs read in data as individual slices or windows in time, and combine the data from multiple subjects into one unified “population” of data points used to fit a base mathematical model representing the population median time course of that drug and/or its metabolites. Further refinement can be achieved by assessing the effect of patient cofactors or covariates on the difference between an individual patient’s profile, or time-based exposure, to that of the population average. The effect of patient characteristics, some of which are themselves time-based, may also be incorporated into the mathematical equations representing the data, thus minimizing the overall inter- and intra-subject variability with respect to the population average. Figure 4 shows a hypothetical example of a population model fit to individual data (panel A) and population model predictions (curves) contrasting the differences in exposure vs time profiles of an individual with physiological deficiencies to that of the population average (panel B). If adjustments for this difference(s) are then incorporated into the mathematical model, the model is better suited to generate exposure estimates for the represented group of patients. This iterative evaluation and representation of data that changes over time serves as the primary driver for the structure for PK, PD, and PMx datasets.

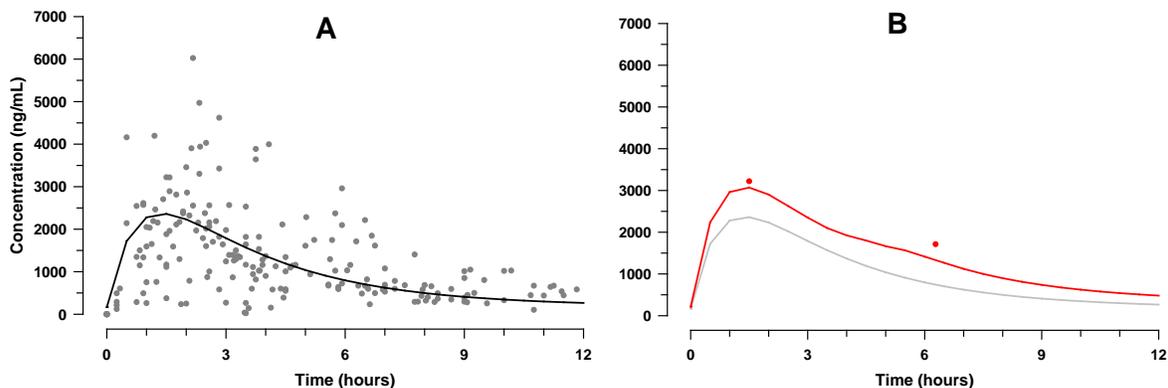


Figure 4. A simple population analysis. Concentration data (dots) and associated model fit (line) are shown in panel A, and the effect of inter-individual variability or patient cofactors on model predictions are shown in panel B.

THE IMPACT OF ANALYSIS TYPE ON DATASET COMPLEXITY: FUNCTION FOLLOWS FORM

Conventional analysis has a more focused endpoint, with few variations in the way the analysis is approached. In fact, most computer programs that perform conventional analysis handle the basic equations in the same way, which means most scientists performing conventional analysis organize their datasets in a very similar manner, irrespective of the program they choose. Exceptions may result in cases where the formulas for conventional analysis require modification for use in other analysis software like SAS®, resulting in the need for a different data layout. However, data layout choices may be driven as much by experience and familiarity as true need. This similarity in approach for analysis would seem to be an obvious opportunity for the PK, PD, and PMx world to standardize on an analysis dataset format for NCA. But it is only recently that the task of creating an analysis data standard for NCA has been attempted: a draft analysis data model (ADaM) standard is being proposed, following the unique time-based structure common to NCA analysis, and is currently in the stages of review and refinement.

In contrast, population-based ADaM standards are more difficult to develop. Unlike conventional analysis, many different approaches to model-fitting and evaluation have been developed, leading to similar but varied dataset formatting requirements. It is this variation in programming approaches and dataset formats that has made efforts to define an industry standard challenging at best. In fact, it may be that the development of a true ADaM data model, in the spirit of “analysis ready,” may be impossible due to the many popular, and passionately adopted, approaches to population analysis. Efforts by the International Society of Pharmacometrics (ISoP) and American Statistical Association (ASA) community of pharmacometricians are currently underway to arrive at a dataset format for population analysis, but customization is still required based on individual approaches. It’s these different approaches that make standardization of population datasets more challenging relative to the complementary efforts currently underway for conventional approaches. Because a population dataset can potentially be modified to accommodate various types of analyses, it does not fit within the “analysis ready” paradigm of ADaM. As a result, an ADaM standard may not be achieved, but a commonly adopted precursor could be defined as a “best practice.”

ACCEPTING THE DIFFERENCES

As an industry, we need to accept that analysis methods for PK, PD, and PMx have developed over time in a way that differs from common statistical approaches. Those differing methodologies require structural data layouts and data collection techniques that differ from existing ADaM models. The fact that established ADaM models were developed to support common statistical evaluations means they are not generally flexible enough to be utilized for PK, PD, and PMx analysis. And while there are some efforts

underway to establish a new ADaM standard for more straightforward conventional analysis, the challenges of adapting a constrained set of data definitions, standards, and formats to an increasingly diverse, complex, and evolving set of analysis techniques — many of which were not under consideration at the time of standards development — is not a trivial task. It's a reality that many industrial organizational structures center data management around the format and supply of information for statistical approaches. This "statistical analysis" focus often means that translation of data into formats that support PK, PD, and PMx analyses may defy efficiency, require specialization in programming staff, and potentially impact our ability to support drug development and registration.

SPECIALIZATION IS STILL NECESSARY

Understanding the analysis, the need for supporting information (e.g., specific dates and times of any endpoints to be considered for modeling), and how data collection impacts the success, variability, and accuracy of the modeling process, is paramount to the success of present-day drug development. Many institutions have recognized that a special skill set beyond basic data coding is required for assembly of PK, PD, and PMx datasets. Effectively merging separate data domains into a holistic, integrated dataset for PK/PD/PMx analysis requires an understanding of not only the rules dictated by individual analysis programs, but an understanding of how the data itself is used in the analysis. Differences between and among patient populations and disease states, different approaches to large vs small molecules, and variation in endpoints and study designs all contribute to special formatting requirements that extend beyond simple data layout models. The importance of this specialized skill set is only magnified when considering the need to understand and account for inherent variability in individual patient data, sample collection deviations, dose data entry errors, human nature, and the increasingly numerous and complex procedures in clinical protocols. Understanding how that variability impacts study conduct, the development of PK, PD, and PMx analysis datasets, and the ensuing analyses is key to effectively assembling PK/PD/PMx datasets that are fit-for-purpose.

CONCLUSION

PK, PD, and PMx analysis is different enough from standard statistical evaluations that current standards do not lend themselves to these types of analyses. A better understanding of the analysis techniques and the associated data requirements for PK, PD, and PMx is critical to successfully creating analysis datasets, and standardization efforts have the potential to improve communications and alignment between the Statistics and PK, PD, and PMx disciplines. This, in turn, can lead to more efficient study designs and a more unified approach to the integration of results obtained from each group's analyses of clinical outcomes. Establishing industry data standards, driven by the requirements of the analyses, therefore has the potential to improve the collective understanding of PK, PD, and PMx approaches, aid in regulatory reviews, and impact the quality of clinical trial data.

RECOMMENDED READING

- *J. Gabrielsson and D. Weiner (1997), Pharmacokinetic and Pharmacodynamic Data Analysis, Concepts and Applications, Second Edition, Stockholm, Sweden, Swedish Pharmaceutical Press*
- *M Rowland and T Tozer (1995), Clinical Pharmacokinetics, Concepts and Applications, Third Edition, Media, PA, Williams and Wilkins*

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

David B. Radtke
Eli Lilly and Company
radtke_david_b@lilly.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.