

PK Data Explained

Timothy J Harrington, DataCeutics, Inc.

ABSTRACT

Pharmacokinetics, or PK, is the monitoring of the concentration level over time of an analyte within a human (or animal) body. The analyte is typically an active ingredient of a drug, or an element such as sodium or potassium, or a chemical compound which occurs naturally in blood or other body fluids, such as creatinine or bilirubin. PK data collection for an active drug compound is performed routinely in all clinical trials since it is a critical part of measuring the safety and efficacy of an experimental treatment and in determining the treatment dose amount, frequency, and time-release profile.

INTRODUCTION

The goals of PK analysis are to establish analyte concentration time lines and trends for individual or groups of patients and compare these with their dosing regimen and with other collected study data such as adverse events, ECG, or tumor assessment findings in Oncology studies. An analyte concentration trend can reveal the measure of a drug's effectiveness as a function of its concentration over time for a given dosing regimen. More importantly, certain trends in concentration levels may indicate the presence (or absence) of a safety issue.

An example of using PK analysis to monitor a patient's safety is the measurement of the patient's kidney function. A standard measure of kidney function is the Estimated Glomerular Filtration Rate (EGFR), which is calculated using the level of creatinine in a blood sample. (Creatinine is a waste product removed from blood by the kidneys). The benchmark value of EGFR is 60 mL/min per 1.73 m². A patient is considered to have renal failure, or End Stage Renal Disease (ESRD), if EGFR ≤ 15 mL/min per 1.73 m². Given that the goal here is to keep the patient from reaching ESRD, investigating Physicians need to determine if the patient is suffering a downward trend in EGFR (upward trend in the creatinine level). A patient with consistent EGFR values in the range 30 to 40 would be considered to have elevated creatinine but would be regarded as stable because there is no downward trend in EGFR. However, a patient with a first EGFR value of 60, followed by a value of 50 at a later test, then later still by 40, has a very much more serious problem, in that, if this trend continues unabated the patient will reach ESRD (which is fatal if untreated with dialysis or a transplant). Seeing this trend early in a study means preventative intervention can (and should) be taken in a timely manner.

PK SAMPLE COLLECTION

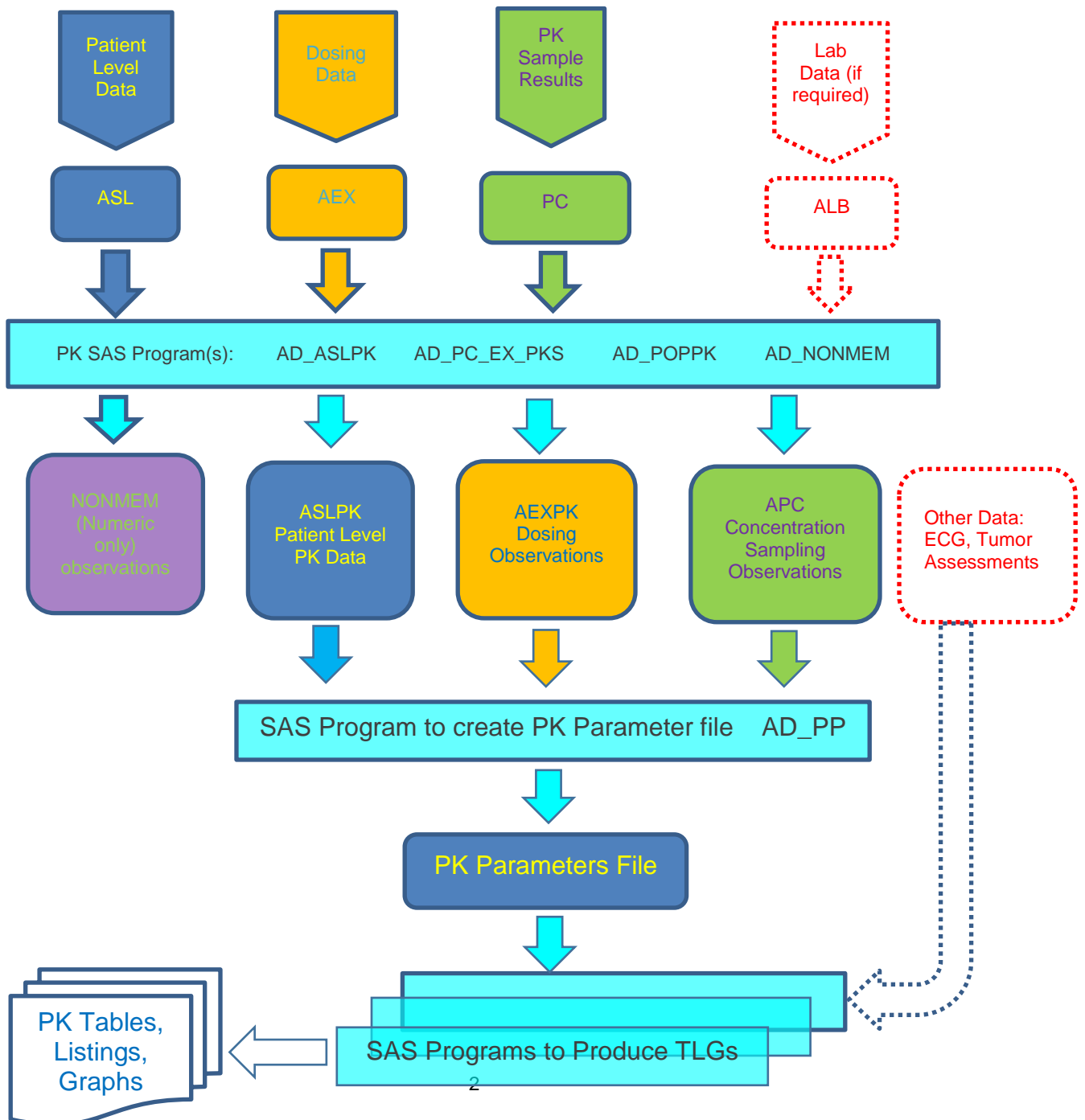
Any body fluid, secretion, or waste product can be a PK sample specimen (assay medium). The most common sampling medium is blood (or blood serum or plasma); an advantage of a blood draw is it is a snapshot of the analyte in the patient's body taken at a specific time. A series of blood samples provide a good live measure of an analyte's absorption, metabolism, distribution, and excretion. Urine is also a common assay medium but is less specific because the patient's bladder has a buffering effect and there are many other factors, such as liquid intake, sweating, and the use of non-investigative medication that affects the volume and concentration of urine produced. Specimens may be analyzed in house at a research center or medical facility, or at an independent laboratory. For FDA submissions, two samples of the same medium taken at the same time must be analyzed by two different independent laboratories. This ensures impartiality of the results and provides an error check since both sets of results should be very similar.

PK DATA SOURCE DATA SETS

PK data arrives from multiple sources, sometimes using different standards. So the results of the analysis must be standardized, that is stored as SAS® data sets using standard CDISC compliant variable names, formats, and informat. In some cases, units of measurement are different and affected results must be

converted to and labelled with the standard units. A common standard unit is 'mg/mL'. The SAS data sets created are PC for concentration results and ALB for general laboratory results. A 'concentration' result usually refers to a study medication active ingredient or metabolite, these observations are held in SDTM.PC. This PC data set and sometimes certain ALB (or SDTM.LB) observations, identified by the analyte of interest, are combined with patient or subject level data (ASL) and with the study medication treatment data (AEX) to produce the output PK data sets. These created PK data sets are ASLPK, the subject level PK data, AEXPK, the treatment, or dosing observations, and APC, the PK concentration results observations. More than one analyte may be present in the same data files. These three created files are then used to create a PK parameter file called PKPARAM. Sometimes a numeric data file NONMEM is also produced to be compatible with the NONMEM legacy analysis software. These data sets AEXPK, APC, PKPARAM, and NONMEM are used as source data files to produce the relevant tables, listings, and graphs.

Figure 1: PK Data Sources and Process Flow



PATIENT LEVEL DATA (ASLPK)

Patient level data includes information about the study, planned and actual treatment grouping (cohorts), unique patient identifier (USUBJID), patient demographics, baseline laboratory results, baseline disease characteristics, and any PK data specific flags. One of these flags is PKEVAL1, when this is 1 the patient's PK data is usable for PK analysis (evaluable), when this is 0 the patient is not PK evaluable. Reasons for a patient not being evaluable include missing dosing data or insufficient or inconclusive analyte concentration results, or a study protocol violation.

EVENTS AND TIMES

The basic dimensions of PK data are time and amount or concentration for each patient. A dose or a collected PC sample is called an 'event'. All event variables begin with EV. The highest-level event variable is the numeric flag EVID, event identifier, which indicates the observation type, dosing or concentration, and source (AEX or APC/ALB). This is set to 1 for dosing observations and to 0 for concentration observations. In some studies, EVID can have the value 2 when a PC observation does not have a valid result, for example if a 'Not Done' flag is set and such observations are not being dropped.

Three types of time variables are used, *nominal*, *actual*, and *imputed*. Most of these are prefixed with NOM, ACT, and IMP accordingly. Nominal time is the time indicated by Case Report Form (CRF) text such as the visit or study time-point. Actual time is the date and time recorded for an event such as a blood draw or the start of a dose being taken. Imputed time is the same as actual time unless the actual time is missing in which case the imputed time is the nominal time. In AEXPK event timing values are taken from the AEX dose start time EXSTDTC. In APC event times are taken from PCDTTC (LBDTTC in laboratory data). In all cases dates and date times are stored as character strings with a length of \$25, left justified in CDISC format as 'YYYY-MM-DDTHH:MM:SS'. The uppercase 'T' indicates time, for example, '2017-01-29T16:45:00'. The 'T' and the time part are omitted where a date but no time is recorded. All event times are actual recorded times as opposed to planned (CRF labeled) happenings.

NOMINAL TIMES

A nominal time is the named or planned dose administration or sample collection time, usually taken from the visit name, treatment cycle, day, and visit time-point printed on the CRF. The most common nominal time variables created for use in PK analysis are CYCLE and DAY, taken from the CRF visit text 'CYCLE xx DAY yy'. NOMTMFDS is the nominal time from the first dose of the corresponding study drug, and NOMTMMRD is the nominal time from the most recent dose of the corresponding drug. NOMTMMRD is calculated by subtracting the prior value of NOMTMFDS from the current value. Theoretically, NOMTMMRD could be taken from the applicable CRF visit or time-point label but an allowance must be made for the possibility of a missed dose. NOMTMFDS at the first dosing observation is zero and before the first dose is zero or negative because there is no prior dose. NOMTMMRD is the same as NOMTMFDS prior to the second dose. In some studies, the negative values of these are set to missing. Other nominal time values often used are NOMTMRDZ, the same as NOMTMMRD except negative values are imputed to zero, NOMTMRDE, the nominal time from the end time of the most recent dose, this usually applies to IV infusions, which may take 90 minutes or more to administer. NOMTMRDP is the same as NOMTMMRD but with pre-dose sample values set to the time before the dose about to occur instead of the prior dose when the pre-dose sample is less than 24 hours before the time of the dose, and NOMTMRDZ is the same as NOMTMRDP but with the pre-dose negative values set to zero. There is also NOMTRDEP, which is the same as NOMTMRDE except, like with NOMTMRDP, if the observation is a pre-dose sample less than 24 hours before the dose the time is measured from the end of that dose which is about to occur. Nominal times are usually measured in days or hours, this unit of measurement is specified as 'day' or 'hr' in the character variable ETMU, the Elapsed Time Unit. These variables are derived from the CRF visit time text, this is usually VISIT in both AEX PC/ALB and also PCTPT in PC (LBPTPT in ALB) which is the sample time-point text, examples are 'PREDOSE', '30MIN POST', and '4HR POST'. NOMTMFDS is normally calculated by multiplying the cycle number less 1 by the cycle length in days and adding the day number minus one. If ETMU='hr' the result is multiplied by 24. PCTPT is then used, where it has a time text, to add the applicable fraction of a day or number of hours accordingly. NOMTMMRD is now calculated by subtracting the NOMTMFDS of the current PC observation from the

NOMTMFDS of the prior dose. All of the variables with a NOM prefix except NOMTMFDS are usually left as zero for dosing (AEXPK) observations. There are often special cases of visits such as 'END OF STUDY', 'EARLY TERMINATION', and 'UNSCHEDULED'. Typically, these situations are handled by using study specific numeric values such as 8888, 9998, and 9999. Such values will cause the end of study observations to be the last for the patient when sorted by USUBJID NOMTMFDS.

ACTUAL TIMES

Actual times are the date and time, date, or time of an event. An actual time is stored in AEX or PC in a \$25 character string as described above. Dose times are taken from EXSTDTC, the dose start time. The most important of these is the date and time of the first dose of each study drug because the relative actual time from first dose is measured from this point. The date and time of first dose may be present in ASL as a numeric datetime, for example, TRTSDTM, otherwise it has to be taken as the earliest non-missing valid dose observation in AEXPK for the given drug. Sometimes the date and time of the end of the dose, EXENDTC, is needed to evaluate the dose duration, this mainly applies to IV infusions. A concentration sample time, in PC, is held in PCDTC, sometimes the date and time of the dose to which the sample refers is held in a reference date PCRFTDTC. The derived variables, which correspond to nominal times are prefixed with ACT and are measured in the same ETMU units. ACTTMFDS is the actual time from first dose, and is calculated by subtracting the start time of the first dose from the start time of the current dose of the same medication. ACTTMFDS is zero for the first dose. ACTTMMRD is the actual time since the most recent dose and is evaluated by subtracting the time of the prior dose from the time of the current dose for the same medication. Note that ACTTMMRD will be the same as ACTTMFDS until the second dose is taken. ACTTMMRZ is the same as ACTTMMRD but with negative values imputed as zero. ACTTMRDE is the difference between the start time of the current dose and the end time of the prior dose. ACTTMRDP is the same as ACTTMMRD but with pre-dose times occurring less than 24 hours before the next dose imputed to the sample time minus the time of that dose which is about to occur, hence this result is negative. ACTTMRDZ is the same as ACTTMRDP but the negative values are imputed to zero. ACTMRDEP is the same as ACTTMRDE but pre-dose samples less than 24 hours before the next dose are imputed to the sample time minus the time of that dose which is about to occur, hence this result is negative. Figure 2 below illustrates these timing values.

EVENT TIMES

Event times are actual recorded times, which have an EV instead of ACT prefix. The most useful of these is the event start datetime, EVNTSDTM, which is the numeric datetime of EXSTDTC or PCDTC (or LBDTC) accordingly. Sorting the PK data with USUBJID and EVNTSDTM as the BY variables places the whole data in chronological order. Other timing variables derived from EVNTSDTM are EVNTDT, the event date as a SAS internal numeric date, EVNTSTTM, the event time as a SAS internal numeric time, EVNTDTC, the character event date formatted as 'YYYY-MM-DD', and EVNTSTTC, the character event time formatted as 'HH:MM:SS'. All character event time variables have a length of \$25 and use the applicable CDISC format. In practice problems occasionally occur due to erroneous input, for example incomplete dates such as '2018-01', or times which are improperly sequenced such as a pre-dose time which is later than, or equal to, the corresponding dose time. Date imputation rules need to be followed in these situations, one such rule is that if the pre-dose time is missing (or is 'T00:00:00') this time is imputed from the time of the dose about to occur less an estimate of the time the pre-dose is taken before the dose is administered. This estimated time is usually five minutes but should be defined in the study protocol. When the time is missing for a post dose observation an estimate is added to the dose time based on the CRF time text in PCTPT or the actual times of the other post dose samples.

IMPUTED TIMES

These variables apply to the timing variables with NOM and ACT prefixes and use the IMP prefix. If there is a non-missing actual time the imputed time is set to it. If the actual time is missing the imputed time is set to the nominal time. The imputed time is often used for ordering observations chronologically because the nominal time is usually close to the actual time. Unscheduled visits have an actual time but do not have a nominal time, and termination visits have an actual time but an imputed nominal time. Common imputation values are 9999 and 8888 respectively. These large numbers ensure termination observations

protocol. This is typically longer for a first IV infusion than for subsequent IV infusions. When the dosing method is tablets taken orally these duration variables are usually missing. FRMUL describes the dosing formulation, examples are 'INTRAVENOUS' or 'CAPSULES'. ROUTE describes the route of dosing administration, such as 'IV' or 'PILLS ORAL'. The dosing regimen is stored as a character text in the variable REGIMEN in the standard pharmaceutical notation; examples are 'BID' and 'QD'. This information comes from the study protocol or EXDOSFRQ. FEDSTATE is a text describing the patient's fed status at the time of dosing, typically 'FED' indicates the stomach is full and hence absorption of orally taken medication is expected to be slower than if 'FASTED'. Occasionally problems occur during dosing, for example, with orally taken medicines, difficulty swallowing or vomiting results in a missed or incomplete dose. Sometimes the patient forgets to take the pills, takes them at the wrong time or takes an incorrect amount. To handle these situations there is a text variable named TXOK, when this is 'OK' there were no reported dosing problems, otherwise an applicable text indicates the nature of the problem such as 'Adverse Event'. When dosing is administered by IV, if there is an interruption, this is recorded in TXOK together with the reason for the interruption. An example is "Problem: Adverse Event: Infection at IV site". The AEX data usually contains a flag, such as EXVDFL, indicating if the dose was a valid dose. Normally observations where EXVDFL is not 'Y' ('N' or missing) are dropped.

Note: Although the variables included above apply to dosing (AEXPK, EVID=1) observations, some of them such as ACTAMT, NOMAMT, and IMPAMT, are also required to be present on sample (APC, EVID=0) observations. Temporary variables are needed to store values such as the date of first dose so they may be used for deriving in both AEXPK and PC.

CONCENTRATION DATA (APC)

Values in sample concentration (EVID=0) observations are derived from PC (and/or ALB). A given PC observation may be excluded if it does not have a valid result or is marked as 'NOT DONE', or, such an observation may be included but with EVID set to 2. The type of analyte is taken from the PC test name PCTEST, or the test code PCTESTCD (PCTEST and PCTESTCD may be PARAM and PARAMCD). The corresponding analyte name is stored in the character string ASTYP. Note: The concentration analyte ASTYP must be associated with the corresponding DRUG dosing; usually a temporary variable is used to link the DRUG with the corresponding ASTYP. The specimen material type, for example 'PLASMA' or 'SERUM' is taken from PCSPEC or LBSPEC and is stored in ASMED (Assay medium). ASCM is a character string, which is usually blank but may contain any relevant comments about the assay. Every sample is identified with a unique ID number (as a bar code on the sample container) held in PCSPID or LBSPID. This is kept as a character string in SMPID. Any additional comments 'other comments' about the sample are stored in OTHCND. The test result is taken from PCSTRESN, the numeric result, PCSTRESC, the character representation of the result, and PCSTRESU, the units of measurement of the result. These values are stored in AVAL, AVALC, and AVALU respectively. Typically, when reading the PC or LB data any observations where PCSTRESN is missing and PCSTRESC is blank are excluded. If PCSTRESN is not missing PCSTRESC will have the corresponding value as a character string, i.e. when PCSTRESN is 1.23, PCSTRESC would be '1.23'. However, some results have such low levels of concentration that they are not measurable, in this situation PCSTRESC may be 'LTR' (Less Than Recordable), 'BQL' (Below Quantification), 'QNS' (Quantity Not Sufficient), or may contain a '<' symbol such as '<1.5' where 1.5 is the lowest limit of quantification the laboratory can measure for the particular analyte. When the result is less than quantifiable PCSTRESN is missing. This situation usually occurs with pre-dose samples before the patient takes their first dose or when there is no measurable analyte remaining from a prior dose. When a result is below quantification (e.g. PCSTRESC='LTR'), methods of imputation may be employed to impute such a missing numeric quantity as a Summary Concentration, or SUMMCONC, value. The numeric lowest level of quantification is taken from PCLLOQ and is stored in LLOQ for use for imputing less than quantifiable concentration values. There are three main derived concentration variables. The first of these is the character variable RAWCONC, the raw concentration, which is the same as AVALC, but may have numeric digits rounded to a specified decimal precision, and have any less than quantifiable AVALC result set as one such result text. For example 'LTR', 'BLQ', 'QNS', and if AVALC contains '<' are all mapped to RAWCONC='LTR'. The second derived concentration variable is GRAFCONC, or graph concentration, which has all 'LTR' values imputed to half of LLOQ, except for a 'LTR' value when NOMTMFDS<=0 (At or before the first dose), in which case GRAFCONC is set to zero. GRAFCONC, as its name implies, is used for producing graphic plots of the concentration

with time. The third variable is SUMMCONC, or the Summary Concentration. SUMMCONC is evaluated by imputing numeric results for 'LTR' values based on the proportion of the results which are 'LTR' and the LLOQ. The most common such imputation rule is the 'Rule of Thirds'. This is performed by sorting the concentration (APC/EVID=0) observations by analyte (ASTYP) and nominal time (NOMTMFDS), then, when NOMTMFDS is greater than zero, if less than one third of the RAWCONC values are 'LTR', SUMMCONC is imputed as half of LLOQ in those observations. If one third or more observations have RAWCONC='LTR' SUMMCONC is set to missing in those observations. In both cases the observations where RAWCONC is not 'LTR' have SUMMCONC set to AVAL. When NOMTMFDS is zero or less SUMMCONC is set to zero. SUMMCONC is now ready for use in producing tables and charts.

NONMEM DATA

Nonmem data is a specific collection of the PK data. All variables are numeric and all variable names are four or less characters. Often the NONMEM column name is a shortened form of the regular PK name, for example STUDYID becomes STUD, PATNUM becomes PTNM, ACTTMFDS becomes TIME, ACTTMMRD becomes TMRD, and NOMTMFDS becomes NOMT. NONMEM is an acronym for NON-linear Mixed Effects Modeling and is an industry standard software package, which compensates for random variability between individual patients in patient groups.

PK PARAMETERS

PK parameters are measurements pertaining to the concentration and duration of the given analyte in the patient's blood. The parameters are measured for each corresponding dose for each patient. For medications administered by IV infusion the concentration response in the blood is immediate and climbs steadily during the infusion reaching its maximum at the end of the infusion. The concentration then slowly and exponentially decreases over time due to excretion. Post dose blood draws are taken at predetermined intervals (post dose nominal times) after dosing is complete and the analyte concentration is measured and recorded. For orally administered medications the increase in concentration in the blood is relatively slow due to the gastro-intestinal delay, this is partially dependent on whether the patient has recently consumed food, hence the need for knowing the patient's fed state ('Fed' or 'Fasted'). Slower absorption of the analyte causes the maximum concentration to be reached after a longer time compared to IV administration. A pre-dose blood draw is taken just before each dose administration to verify the concentration just prior to dosing (CMIN).

There are three possible phases to a dosing regimen. The first is *titration*, when a large enough dose is administered frequently enough so there is still an expected level of the analyte in the patient's blood at the pre-dose sample taken just before the next dose (CMIN). Over a number of doses, the effect is to raise the average dose level to a predetermined steady state level. The second phase is *steady state* where the mean dose level is maintained for the duration of the phase or epoch of the study. This may involve the dose level being allowed to fall to a certain minimum before the next dose or to less than recordable (LTR) a certain time before the next dose. The third phase is the *withdrawal* phase where the dose level is gradually reduced or dosing is less frequent until a lower level of dosing regimen is achieved or treatment is stopped. Not all clinical studies have titration or withdrawal phases.

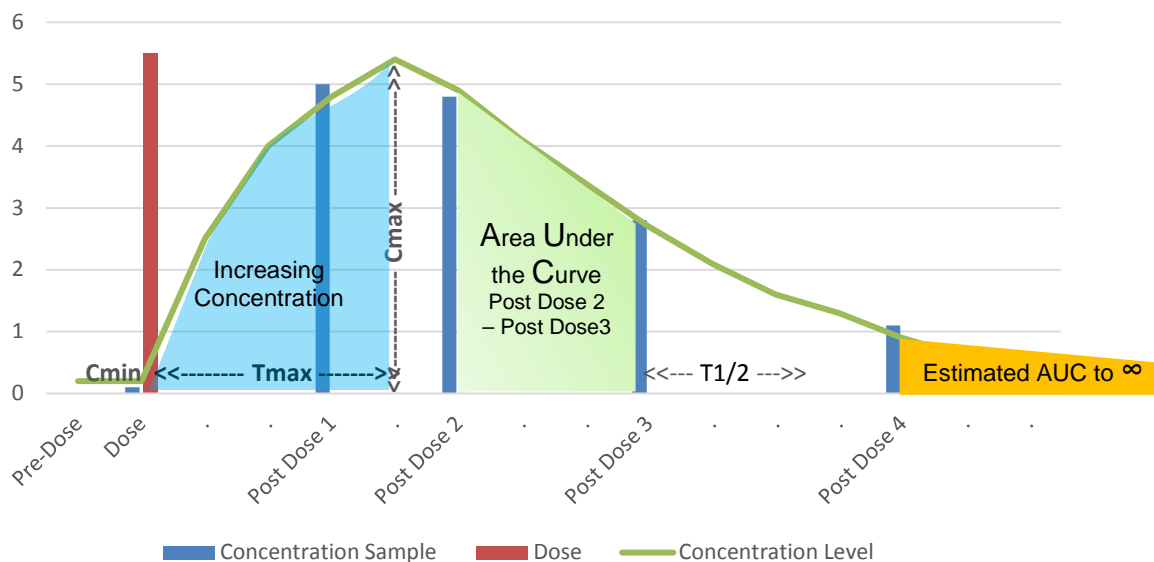
The PK parameters are calculated for each separate dose administration. The most important of these are CMAX, the maximum level of concentration, calculated from the post dose readings, and CMIN, the lowest level of concentration or pre-dose blood draw trough. The units of measurement of these two parameters are the mass (or molarity) of the analyte per unit volume of sample material. The most common unit is mg/L. CMIN should be less than recordable (LTR) before the very first dose and at a predetermined minimum (which can also be LTR) for subsequent pre-doses depending on the phase. TMAX is the time after dose commencement when CMAX occurs. TMAX is usually measured in hours or fractions of a day. Since CMAX is unlikely to occur at an exact post dose interval a good estimate has to be calculated using certain statistical models and the available concentration results and sample times. T1/2 is the half-life of the analyte concentration, that is the time between a starting time-point, at TMAX or later, and the time-point where the concentration has fallen to 50% of this starting time amount. This is useful for determining future dosing frequency. AUC is the Area Under the Curve, that is the area under the concentration level curve between two given time-points within a dosing cycle and is measured in

mass per unit volume per unit time (e.g. mg/L/hr). This provides a good estimate of the bioavailability of the drug between any two sampling points. Figure 3 below shows the AUC between the 'Post Dose 2' and 'Post Dose 3' time-points. The total AUC is the area under the entire curve, $AUC_{0-\infty}$. In Figure 3 this is the shaded 'Increasing Concentration' area, $AUC_{0 \text{ to } T_{max}}$, plus AUC from T_{max} to 'Post Dose 4', plus an estimate of the AUC following 'Post Dose 4'. This is the shaded area after post-dose 4 in Figure 3 and is estimated assuming exponential decay. The example shown in Figure 3 is for an orally administered drug. For a drug administered intravenously, T_{max} is much shorter, and since there is no delay in the drug entering the bloodstream, T_{max} is the infusion duration. In this model the increase in drug concentration can be assumed to be linear and hence the AUC prior to T_{max} is calculated as $0.5 \cdot C_{max} \cdot T_{max}$.

These parameters provide useful information on the absorption distribution, metabolism, and excretion processes pertaining to the drug and hence its bioavailability. This information helps investigators determine the most beneficial dosing regimens and formulation. PK Parameters are a good bench-mark for comparing the performance of two different drugs, such as generic and non-generic variants. Mapping the timeline of PK parameter values with other data timelines from the same patient population enables any correlation between the parameter values and other happenings such as selected adverse events, change in tumor size in oncology studies, or ECG data, to be tested for significance particularly when there are a small number of doses.

Greater accuracy is achieved when collecting data from a larger number of patients in each treatment group. Grouped statistics such as the mean, standard deviation, median, minimum, and maximum can now be calculated and used for producing tables and graphs: Such statistics are commonly grouped by patient demographics such as age, gender, race, and baseline body mass index

Figure 3: Illustration of the main PK Parameters (Concentration v. Time)



CONCLUSIONS

The information gained from the results of ongoing PK analysis during a clinical trial is invaluable for both monitoring and for compiling FDA submission reports on both the safety and efficacy of very many experimental drugs. This information is critical for determining the medication formulation and frequency and amount of dose to be administered, and hence the long term dosing regimen for future patients.

ACKNOWLEDGMENTS

The author would like to express gratitude to the following for their assistance in writing and reviewing this paper:

Jeffrey Dickinson, Sharon Hall, and Kathy Greer at DataCeutics, Inc. 215 West Philadelphia Avenue, Boyertown, PA, 19512 | [610.970.2333](tel:610.970.2333) | info@dataceutics.com.

Paul Slagle, Bradford Danner, and Xiaohui Wang PharmaSUG 2018 Academic Chair, at AcademicChair@PharmaSUG.org

REFERENCES AND SUGGESTED FURTHER READING

Hannah Twitchett and Paul Grimsey 2012. "An Introduction to Pharmacokinetics". Proceedings of the Annual PhUSE Conference, Budapest, Hungary. www.lexjansen.com/phuse/2012/is/IS05.pdf.

Smitha Mullapudi. 2011, "An Introduction to PK/PD Data Structures and Process Optimization to get Analysis Ready Data". Annual Northeast SAS Users' Conference, Portland, Maine www.lexjansen.com/nesug/nesug11/cc/cc26.pdf.

Bolton, Sanford. 1990. "Pharmaceutical Statistics, Practical and Clinical Applications, Second Edition, Revised and Expanded". St. John's University, Jamaica, New York. Marcel Dekker, inc. New York and Basel.

Dr Graham Lappan Visiting Professor of Pharmacology, University of Lincoln, UK. March 6, 2017. Pharmacokinetics: video-1 foundations (series of 7 videos) <https://www.youtube.com/watch?v=qwVDYO46KIE>

CONTACT INFORMATION

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

Timothy J. Harrington
DataCeutics, Inc.,
215 West Philadelphia Avenue,
Boyertown,
PA 19512
(610) 970 2333
harringt@dataceutics.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.