Methodologies of Safety Assessment in Clinical Pharmacology

Werner Seiz

Contents

W. Seiz (✉)
R&D Translational Medicine and Clinical Pharmacology,
Sanofi-Aventis Deutschland GmbH, Frankfurt am Main,
Germany
e-mail: werner.seiz@sanofi.com

F. J. Hock, M. R. Gralinski (eds.), *Drug Discovery and Evaluation: Methods in Clinical Pharmacology*,
https://doi.org/10.1007/978-3-319-56637-5_30-1

Abstract

An important objective of clinical pharmacology is the early and ongoing assessment of the safety and tolerability of a new drug. This is done by assessing the type, frequency, and severity of side effects, assessing in which patient population these side effects may occur at which dose or exposure, for what duration and whether these side effects can be monitored and whether they are reversible. The terminology for the safety assessment of drugs has some specifics that need to be explained right at the beginning of this chapter.

Introduction/General Considerations

An important objective of clinical pharmacology is the early and ongoing assessment of the safety and tolerability of a new drug. This is done by assessing the type, frequency, and severity of side effects, assessing in which patient population these side effects may occur at which dose or exposure, for what duration and whether these side effects can be monitored and whether they are reversible. The terminology for the safety assessment of drugs has some specifics that need to be explained right at the beginning of this chapter.

Definition of Adverse Events as the Parameter to Assess Safety

The term "side effect" used for marketed drugs is replaced by the term "adverse event" (AE) in studies with investigational drugs. An adverse event is defined as any unfavorable and unintended sign, symptom, or disease temporally associated with the use of a drug, whether or not considered related to the drug.

In this chapter the term "adverse event" is explicitly also used for any abnormal laboratory value, as the consequence of abnormal values will be evaluated in the same scheme as for clinical adverse events.

The term "treatment related" is often added as a modifier in order to remove preexisting conditions from consideration. A further term "serious adverse event" is used to describe any untoward medical occurrence that, at any dose, results in death, is life-threatening, requires hospitalization or prolongation of an existing hospitalization, results in persistent or significant disability or incapacity, or is a congenital anomaly or a birth defect. Serious adverse events have to be reported to the health authorities in an expedited manner, typically.

The severity of an adverse or serious adverse event is classified as either mild, moderate, or severe. Standardized definitions for adverse events and classification of severity have been published by the National Cancer Institute (NCI) of the National Institute of Health (NIH) and are also used for clinical trials in nononcology indications (NCI 2017).

It is important to distinguish between the severity and the seriousness of an adverse event. A severe adverse event is not necessarily serious (e.g., severe abdominal cramps not causing hospitalization), and a serious adverse event is not necessarily of severe intensity (e.g., mild to moderate, prolonged dizziness of an outpatient causing hospitalization) (Herson 2000).

How to Manage the Safety Assessment of a Drug

One of the most critical steps in the development of a new drug is the first administration of a drug to humans, the first dose escalation, the first multiple dosing, and the first switch from healthy patients to the targeted patient population. In order to acquire the safety data in a responsible way, it is necessary to consider all of the following areas for each clinical study and to plan these items in advance:

- Expect, plan, and manage the occurrence of adverse events. This administrative part of the safety method includes the selection of the right preclinical animal models for the prediction of the target organ, the definition of the exposure to the drug at the no-observed-adverse-effect-level (NOAEL), the adequate calculation of the safe starting dose in humans, the decision about the dose escalation and when to stop it, the proper organization of the clinical trial, and the definition of the expected adverse event profile.
- Plan and manage the acquisition of adverse events data. This includes – based on the expected adverse event profile – the selection of the clinical, technical, and laboratory observations, by which the expected adverse events are to be monitored.
- Plan and manage the interpretation of the adverse events data and their impact on the subsequent development or study conduct. In order to avoid bias, the statistical analysis of the safety data obtained has to be predefined, using commonly accepted criteria. For each

parameter assessed, it should be clear prior to the analysis, which deviation is considered relevant and therefore an adverse event. This is usually done in the statistical analysis plan, which has to be finalized prior to closing the database of a study and prior to breaking the randomization code. In clinical pharmacology studies, data monitoring committees (DMC) are typically not included; however, in studies with adaptive designs DMCs might be installed very early in clinical development.

Case Study

The importance of an adequate selection of animal models, assessing the significance of the preclinical data obtained for humans and planning adequately the study conduct in the first-in-human study has been shown quite dramatically a decade ago. The first dose step in the first-in-man study with the biotherapeutic TGN1412, a humanized agonistic anti-CD28 IgG4 monoclonal antibody (present on regulatory T-cells), induced a cytokine release syndrome in all six active-treated healthy volunteers, all of whom suffered from life-threatening, acute shock and subsequent multi-organ failure. At least in one of the participants of the TGN1412 first-in-man study, several fingers and toes were to be amputated finally. Obviously this severe and serious adverse events were not predicted by the animal studies conducted prior to human studies, the dose administered was obviously above the minimum active biological effect level (MABEL) for humans, and all volunteers were already dosed before the first dosed person suffered from the symptoms of the upcoming cytokine release syndrome, that is, within less than 90 min. Although a complete explanation of the event was never unanimously accepted (http://www.circare.org/foia5/clinicaltrialsuspension_interimreport.pdf), at least it appears that the drug was given too fast to each subject (3–6 min infusion time) and to too many subjects within too short a time (every 10 min the next subject was dosed) (Horvath and Milton 2009). As a consequence of this event with TGN1412, the regulators worldwide have changed several processes so that this should not happen again. The European Medicines Agency (EMA) has revised their guidance on first-in-human clinical trials to identify and mitigate risks for trial participants (EMA 2017) after a case of death of a human volunteer in a first-in-human clinical trial in 2016. In that case, a 49-year-old healthy subject who experienced neurological symptoms after the 5th out of 10 planned doses in the first multiple ascending dose study with a nonselective fatty acid amid hydrolase inhibitor was submitted to hospital and died about 7 days later (Brentano and Menard 2016). A major new request by the EMA is the treatment of a sentinel at least 24 h before subsequent subjects are to be treated.

It is self-evident that the evaluation and interpretation of the safety data obtained as a whole is of utmost importance to a drug development program; however, here in this chapter the topic will be the technical description of the most often used clinical, technical, and laboratory methods to acquire safety data and how this will influence decisions on dose escalation or termination of a study. No more thoughts are given to analyze the safety data as a whole and in the context of the already accumulated clinical safety data.

Categorization of Adverse Events for Decision Making

Purpose and Rationale

Adverse events should be categorized in the same way across studies so that the decisions based on these categories are consistent within a development program and across programs. The NCI-CTCAE v5.0 terminology is defining the following five grades of severity for each adverse event (AE):

Grade	Definition
1	Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated
2	Moderate; minimal, local, or noninvasive intervention indicated; limiting age-appropriate instrumental activities of daily living (ADL)
3	

(*continued*)

Grade	Definition
	Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self-care ADL
4	Life-threatening consequences; urgent intervention indicated
5	Death related to AE

Procedure

Each adverse event or finding is classified into one of five categories, where grade 1 indicates a deviation from the norm without an obvious relevance for the subject, grade 2 indicates a mild interference with daily activities for the subject but without need for treatment except non-opioid analgesics, for example. Occurrences of events of grade 2 have to be seen as an alert on reaching doses, where tolerability to the test compound decreases. Grade 3 indicates that the event or finding requires medical or other treatment or prevents daily activities of the subject. Grade 4 is reserved for definitely inacceptable adverse events, which typically, if not occurring in the placebo group, leads to a termination of the study at least of the dose, when the grade 4 event has been observed (e.g., rhabdomyolysis, angioedema). Grade 5 of course is an inacceptable consequence of treatment with an investigational drug and may even terminate a drug development project. If there is a rapid change in a parameter, this also might lead to an increase in grading. For laboratory parameters (chemistry, ECG), the grading is done based on the likelihood for further consequences or risks according to the categories above. In order to categorize laboratory values as abnormal, they have to be different from the normal range, which is specific to each laboratory. The numbers given here are suggestions and are based on published normal ranges (Kratz et al. 2004).

Evaluation

For each subject, the maximal adverse events' grading can easily be assessed. For gradings 3 or 4, unblinding is recommended; for grade 5, unblinding and immediate termination of the clinical trial is a must. An individual should not be further dosed if on active drug and grading 3 occurs. If only placebo-treated subjects suffer from an adverse event and if this event is not study-procedure related, it has no impact on further study conduct. If placebo- and active-treated patients suffer from grading 3, but with less than 50% of active-treated subjects, doses should be adapted (=lowered), the number of subjects treated at a time need to be reduced, and the time interval between subjects should be increased, in order to minimize the risk for treated subjects. If the dose step is well tolerated, additional subjects could be treated at the dose with the grade 3 events. Finally, if more than 50% of active-treated patients suffer from grade 3 adverse events at a given dose, the dose below is qualified as the maximal tolerated dose. At all grades, clinical judgment is needed based on the nature, reversibility, and monitorability of adverse events.

Critical Assessment of the Method

The categorization of information leads to a loss of information and therefore has to be used with care. Everyone using this method needs to be aware that the full picture and information needs to be taken into account and not just the results from the categorization. The grading system suggested and described here is modified from (Sibille et al. 2010) and not approved by any authority but should be seen as way to consistently aggregate and interpret information. Grading adverse events is in use in oncology and vaccine studies already (Cancer therapy evaluation program 2009; FDA Guidance 2007). These systems use four or five gradings, where grade 5 is always "death" and grade 4 is mostly of life-threatening adverse events, which is not in contradiction with the grading used here.

Modifications of the Method

An alternative to formal categorization of adverse events for subsequent decision making on dose escalation or study progress is the repetitive assessment of the uncategorized clinical and laboratory data by the investigator, the sponsor, and additional experts to achieve a common understanding on how to proceed. The pre-categorization of events as described here in this chapter, however, does not prevent such an approach and has the advantage to provide a consistent assessment of the information across dose steps, studies, and compounds.

Decision Making on Dose Increase and to Stop the Study

Purpose and Rationale

The decision to stop dose escalation should be based on the observation of adverse events no longer tolerable (by frequency or severity) and by the observed exposure information.

Procedure

The grading of the adverse events and their frequency need to be assessed. As long as no adverse events are observed and the exposure is not above the exposure in the most sensitive species, dose escalation may go on as planned in the protocol. If the exposure is above the NOAEL exposure, careful further dose escalation may be reasonable to define the maximal tolerable dose. If the severity of adverse events is below grade 3 and the exposure of the NOAEL is not reached, dose escalation can proceed. No further dose escalation should be considered, if more than 50% of active-treated subjects suffer from adverse events of grade 3.

Evaluation

In case grade 3 or 4 adverse events do occur, treatment of ongoing subjects should be stopped immediately.

Clinical Adverse Events Monitoring (Report by Subjects)

Purpose and Rationale

Most drug-related adverse events are based on the spontaneous reporting of clinical signs by the clinical trial participants. Subjects participating in a clinical trial can realize these adverse events at any time.

Procedure

The subjects are asked to report any events, signs, or abnormal observations and feelings to the study personnel immediately. In addition, subjects should be asked direct questions from time to time, such as “Did you make any disagreeable or unexpected observations since you took the drug?” The information obtained need to be documented without interpretation at first. The (preliminary) diagnosis and decision about next steps (physical, and if indicated additional laboratory or technical examinations) will be based on the interpretation by the responsible MD on this report.

Evaluation

Categorization of adverse events reported by the subjects is to be done by experienced medical personnel taking into account possible differential diagnosis and their time course. The reference http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/Vaccines/ucm074775.htm gives an example of recommendations by the FDA.

Clinical Adverse Events Monitoring (Physical Examination by the Clinical Investigator)

Purpose and Rationale

Physical examination based on spontaneous reporting of adverse events will be conducted as needed.

Procedure

Physical examination can include auscultation, investigation of reflexes, or orientation, etc.

Evaluation

The investigator needs to decide on the classification of physical findings based on changes to baseline and their relevance.

Critical Assessment of the Method

Typically there will not be many clinical findings on physical examination. If there are some, this indicates already quite substantial effects (e.g., angioedema, rashes, ankle edema). Exceptions are findings in the vital signs of heart rate or blood pressure (see below), where effects are frequently seen.

Timing of Monitoring

Purpose and Rationale

Timing of clinical and laboratory assessments need to be in line with the timecourse of drug concentration over time.

Procedure

Standard monitoring needs to be done at baseline and repetitively after drug administration. For orally administered drugs, this is typically at baseline before drug administration – 30 min, 1, 2, 4, 8, 12, and 24 h after dosing for once daily drugs. Timing has to be tailored to the specific pharmacokinetic and pharmacodynamic profile of a drug.

Vital Signs

Heart Rate

Purpose and Rationale

The heart rate is influenced by the sympathic and parasympathic system, which can be affected by drugs directly or indirectly. Heart rate as a vital parameter has to be quite stable as heart rate effects in patients with ischemic heart disease could lead to angina pectoris. In phase I studies, heart rate typically is most affected by increased vagal tone and subsequent bradycardia and occasional fainting.

Procedure

Continuous ECG monitoring by telemetry or Holter ECG is the method of choice to observe effects on heart rate.

Evaluation

Normal range: 50–80/min in supine position. Grade 3 definition: $<$45/min for bradycardia. Grade 1 definition: 100–115/min. Grade 2 definition: 116–130/min. Grade 3 definition: $>$130/min.

Critical Assessment of the Method

Basic method for safety and tolerability assessment.

Modifications of the Method

Holter monitoring allows continuous 24-h assessment of heart rate including analysis of cardiac arrhythmias. Holter monitoring should be used whenever there is evidence of proarrhythmic potential of a drug. Central analysis of Holter ECGs is recommended to also be able to compare acquired data to larger groups of subjects.

Vital Signs

Blood Pressure

Purpose and Rationale

Blood pressure is dependent on stroke volume, heart rate (stroke volume × heart rate = cardiac output), and peripheral resistance. Decrease in cardiac output and/or resistance decreases blood pressure and vice versa. A decrease in blood pressure is most often a result of either vasodilatation or decrease in heart rate, both of which can occur during increased vagal stimulation.

Procedure

Blood pressure can be measured manually or by a machine, in supine, sitting, or standing positions. For functional assessments the Schellong test is an easy to conduct procedure to measure the effect of a physical challenge on heart rate and blood pressure. After 10 min in supine position, the subject is asked to take a standing position. Heart rate and blood pressure are measured 2 min after end of supine position. Timepoints of blood pressure measurements need to be adapted based on the observed effects.

Evaluation

Normal range: 100–140 mmHg systolic in supine position, 50–85 mmHg diastolic in supine position. Grade 1 definition: 140–159 mmHg systolic in supine position and 90–99 mmHg diastolic in supine position for pressure increase; a quantity of 80–100 mmHg systolic in supine position for pressure decrease. Decrease in systolic blood pressure after 2 min standing by more than 20 mmHg together with increase in heart rate. Grade 2 definition: 160–179 mmHg systolic in supine position and 100–110 mmHg diastolic in supine position for pressure increase; a quantity of 70–80 mmHg systolic in supine position for pressure decrease. Cannot stay standing after 10 min of supine position for pressure decrease. Grade 3 definition: >180 mmHg systolic in supine position and >110 mmHg diastolic in supine position for pressure increase; below 70 mmHg systolic in supine position for pressure decrease or syncope during Schellong test.

Critical Assessment of the Method

Basic method for safety and tolerability assessment.

Modifications of the Method

Twenty-four-hour ambulatory blood pressure monitoring (ABPM) is the method of choice for any compound with known or suspected effect of blood pressure as the effect over time can be best followed by continuous monitoring. Some drugs affect the nocturnal decrease in blood pressure (so-called dipping); whenever there is evidence that a drug has such an effect, ABPM should be used early in clinical development. Blood pressure (and heart rate) will be measured every 15–20 min during day time (defined as 6 a.m. to 10 p.m.). During night time, the measurement intervals are 30 min. Full 24 h should be measured, if ABPM is used. ABPM allows to calculate precisely peak and trough effects and the duration of effect on blood pressure.

ECG Parameter

PR Interval

Purpose and Rationale

The PR interval in the ECG is the time during which the electrical excitation is conducted from the atria to the AV-node. Prolongation of the PR interval is a potential side effect of drugs affecting repolarization and bears the risk of AV blockade.

Procedure

Evaluators should be trained in ECG analysis. Automated analysis is frequent but needs to be validated in order to rely upon it.

Evaluation

Normal range: 120–200 ms. Grade 1 definition: <0.8-fold LLN or >1.1-fold ULN. Grade 2 definition: >250 ms. Grade 3 definition: AV-block 2nd degree or syncope.

Critical Assessment of the Method

Basic method for safety and tolerability assessment.

ECG Parameter

QT Interval

Purpose and Rationale

The QT interval in the ECG is the time during which the electrical excitation and repolarization of the ventricles takes place. Prolongation of the QT time and especially the QTc time (QT time corrected for effect of heart rate) is a risk factor for torsades des point, a ventricular arrhythmia associated with an increased incidence of drug-induced sudden cardiac death. QT prolongation of drugs is one of the most frequent reasons for termination of a drug development program.

Procedure

Evaluators should be trained in ECG analysis. Automated analysis is frequent but needs to be validated in order to rely upon it.

Evaluation

Normal range for QTc: 360–425 ms for men, 380–445 for women, increase below 40 ms. Grade 1 definition: Increase above 40 ms and QTc below 475 ms. Grade 2 definition: 476–499 ms and increase below 60 ms. Grade 3 definition: Above 500 ms or increase exceeding 60 ms.

Critical Assessment of the Method

Basic method for safety and tolerability assessment.

Laboratory Parameter

Glucose

Purpose and Rationale

A sufficient glucose concentration in blood ($>$2 mmol/L or $>$40 mg/dL at minimum) is essential for all life processes. Whenever there are signs of decreased consciousness, this vital parameter has to be assessed immediately.

Procedure

Blood glucose should be measured from capillary or venous blood at predefined timepoints and in addition in cases of suspected hypoglycemia or impaired consciousness.

Evaluation

Normal range 3.8–6.4 mmol/L. Grade 1 definition for hypoglycemia: 3.5–3.8 mmol/L. Grade 2 definition: 2.2–3.4 mmol/L. Grade 3 definition: 1.7–2.1 mmol/L.

Critical Assessment of the Method

Basic method for safety and tolerability assessment.

Laboratory Parameter

Potassium

Purpose and Rationale

Potassium concentration in cells is 25-fold higher than in blood. In all cases were potassium is released into the peripheral blood (e.g., during and after hypoxic events) or a decrease in renal excretion occurs, potassium increases will have the potential for cardiac bradyarrhythmias. Hypokalemia can lead to ventricular tachyarrhythmias. Therefore close monitoring of potassium concentration in serum is very important in early phases of development as long as the effect on its concentration in serum is not yet known.

Procedure

Potassium values are measured from serum taken from peripheral veins at predefined timepoints.

Evaluation

Normal range 3.5–5.0 mmol/L. Grade 1 definition: 3.1–3.4 for hypokalemia and 5.1–6.0 for hyperkalemia. Grade 2 definition: 2.5–3.0 mmol/L for hypokalemia and 6.1–6.5 mmol/L for hyperkalemia. Grade 3 definition: 2.0–2.4 mmol/L for hypokalemia and 6.6–7.0 mmol/L for hyperkalemia.

Critical Assessment of the Method
Basic method for safety and tolerability assessment.

Laboratory Parameter

Alanine Aminotransferase (ALT)

Purpose and Rationale
Hepatic damage is one of the most frequent drug-related adverse events and needs to be monitored in every clinical pharmacology study. Transaminases (SGPT/ALT and SGOT/AST), alkaline phosphatase, and total and conjugated bilirubin are the serum assays to detect liver damage.

Procedure
ALT, AST alkaline phosphatase, and bilirubin are taken from serum of peripheral blood at predetermined timepoints and more frequently, if any increases are seen. If increase of ALT is above threefold upper limit of normal (ULN), ALT needs to be followed until normalization (below ULN) or until no further decrease of ALT after termination of treatment is observed.

Evaluation
Any transaminase elevation above the upper limit of normal should be considered as an indicator for hepatic damage. ALT increase is the enzyme specific for liver damage. Normal range is 0–60 IU/L, strongly dependent on lab. Grade 1 definition: Increase >1.2-fold ULN. Grade 2 definition: Increase 2.5- to 5-fold. Grade 3 definition: >5- to 10-fold.

Critical Assessment of the Method
Basic method for safety and tolerability assessment. Increases of ALT are very specific to the liver. Alkaline phosphatase can be increased in other diseases as well, for example, bone disease. Depending on preclinical data of potential liver toxicity and upcoming clinical data early in development, the reporting thresholds for increases in liver enzymes should be adapted specifically.

Laboratory Parameter

Aspartate Aminotransferase (AST)

Evaluation
Normal range is 0–40 IU/L, strongly dependent on lab. Grade 1 definition: Increase >1.2-fold ULN. Grade 2 definition: Increase 2.5- to 5-fold. Grade 3 definition: >5- to 10-fold.

Critical Assessment of the Method
Supporting parameter for ALT analysis.

Laboratory Parameter

Phosphatase

Evaluation
Normal range is 30–120 IU/L, strongly dependent on lab. Grade 1 definition: Increase >1.1-fold ULN. Grade 2 definition: Increase two- to threefold. Grade 3 definition: three- to tenfold.

Critical Assessment of the Method
Supporting parameter for ALT analysis.

Laboratory Parameter

Bilirubin

Purpose and Rationale
Bilirubin assessment together with ALT measurement is used to identify potential risks of hepatic toxicity.

Evaluation
Normal range is 5–27 μmol/L. Grade 1 definition: Increase >1.3-fold ULN.

Hy's law (FDA 2009) is a prognostic indicator that a drug-induced liver injury leading to jaundice has a case fatality rate of 10–50%. Hy's law cases have the three following components:

- The drug causes hepatocellular injury, generally shown by more frequent threefold or

greater elevations above the ULN of ALT or AST.
- Among subjects showing such ALT/AST elevations, often much greater than 3 × ULN, some subjects also show elevation of serum bilirubin to >2 × ULN, without initial findings of cholestasis (serum alkaline phosphatase [ALP] activity >2 × ULN).
- No other reason can be found to explain the combination of increased transaminases and bilirubin, such as hepatitis A, B, or C, preexisting or acute liver disease, or another drug capable of causing the observed injury.

Critical Assessment of the Method
Together with ALT, a very powerful parameter to identify true drug-related hepatic damage.

Laboratory Parameter

Creatinine

Purpose and Rationale
Creatinine is solely excreted by the kidney, primarily by glomerular filtration, and therefore is a good marker of renal perfusion and filtration. Drugs affecting renal perfusion or filtration lead to an increase in creatinine. Increases in creatinine only occur if there is already a significant decrease in renal glomerular filtration rate.

Procedure
Creatinine concentration needs to be measured in plasma and urine. Together with the urine production per time (either within 24 h, or time overnight sampling; for example, 1,500 ml excreted between 8 pm and 7 am), the glomerular filtration rate can easily be calculated.

Evaluation
Normal range: 80–130 μmol/L. Grade 1 definition: >1.1-fold ULN. Grade 2 definition: >1.5-fold ULN. Grade 3 definition: >1.9- to 3.4-fold ULN.

Critical Assessment of the Method
Serum creatinine levels are not very sensitive to large changes in GFR as long as the GFR is still above 60 ml/min/m^2, but then a rapid increase will be observed. A more sensitive method for renal function is the GFR or the fractional excretion of electrolytes.

Laboratory Parameter

Albumin in Urine

Purpose and Rationale
Presence of albumin in urine is an indicator of glomerular damage.

Procedure
Albumin is measured from morning urine.

Evaluation
Normally no albumin is excreted via urine. Any finding of albumin above 300 mg/24 h in urine is indicative of a renal issue that needs to be further evaluated (if prior to treatment with investigational drug the value was negative).

Critical Assessment of the Method
Albumin in urine is always a pathological sign, which needs further analysis.

Laboratory Parameter

Creatinphosphokinase (CPK)

Purpose and Rationale
CPK is released during damage of skeletal muscle, a frequent side effect of lipid lowering compounds like statins.

Procedure
CPK is taken from serum of peripheral blood at predetermined timepoints and more frequently, if any increases are seen. If increase of CPK is above threefold ULN, CPK needs to be followed until normalization (below ULN) or until no further

decrease of CPK after termination of treatment is observed.

Evaluation

Normal range: 50–400 IU/L. Grade 1 definition: 480–1,000 IU/L. Grade 2 definition: 1,000–2,000 IU/L. Grade 3 definition: 2,000–5,000 IU/L.

Laboratory Parameter

Hemoglobin (Male Subjects)

Purpose and Rationale

Hemoglobin can be affected by acute bleeding, by chronic suppression of erythrogenesis, or by dilution/concentration due to changes in the intravasal volume.

Procedure

Hemoglobin is assessed from whole blood taken from peripheral veins at predetermined timepoints.

Evaluation

Normal range for males: 13.5–17.5 g/dL. Grade 1 definition: 12.0–12.5 g/dL and decrease >1.5 g/dL. Grade 2 definition: 10.0–11.9 g/dL. Grade 3 definition: <10.0 g/dL. Normal range for females: 12.5–15.5 g/dL. Grade 1 definition: 11.0–12.0 g/dL and decrease >1.5 g/dL. Grade 2 definition: 9.5–10.9 g/dL. Grade 3 definition: <9.5 g/dL.

Critical Assessment of the Method

Basic method for safety and tolerability assessment.

Laboratory Parameter

Polymorphonuclear Leucocytes (PMN)

Purpose and Rationale

Immunotoxic effects of drugs on white bloods cells are not uncommon and need to be detected early on in development.

Procedure

PMN count is assessed from whole blood taken from peripheral veins at predetermined timepoints.

Evaluation

Normal range: 1.7–7.5 G/L. Grade 1 definition: <0.7-fold LLN or >1.3-fold ULN. Grade 2 definition:1.0–1.3 G/L. Grade 3 definition: <1.0 G/L.

Critical Assessment of the Method

Basic method for safety and tolerability assessment.

Laboratory Parameter

Platelets

Purpose and Rationale

Immunotoxic effects of drugs on platelets are not uncommon and need to be detected early on in development.

Procedure

Platelet count is assessed from whole blood taken from peripheral veins at predetermined timepoints.

Evaluation

Normal range: 150–350 G/L. Grade 1 definition: <0.85 LLN. Grade 2 definition: 50–125 G/L. Grade 3 definition: <50 G/L.

Critical Assessment of the Method

Basic method for safety and tolerability assessment.

Coagulation Parameter

Activated Partial Thromboplastin Time (aPTT)

Purpose and Rationale

Effects on aPPT are seen in cases of decreased hepatic protein synthesis rate.

Procedure
aPTT is assessed from plasma.

Evaluation
Normal range: 22–43 s. Grade 1 definition: 1.1- to 1.3-fold ULN. Grade 2 definition: 1.3- to 1.5-fold ULN. Grade 3 definition: >1.5-fold ULN until minor bleeds. Grade 4 definition: Major bleeds.

Critical Assessment of the Method
Basic method for safety and tolerability assessment.

Laboratory Parameter

Kidney Injury Molecule-1 (KIM-1)

Purpose and Rationale
KIM-1 is a rather new biomarker indicating renal toxicity at the tubular level. KIM-1 has been preclinically qualified as an excellent marker for drug-related renal toxicity. If there is preclinical evidence for renal toxicity at this region and if KIM-1 has been used in nonclinical toxicity studies, this parameter should be monitored.

Procedure
KIM-1 can be measured using commercially available kits.

Evaluation
Look for significant changes from baseline and if those occur, stop treatment, follow KIM-1 until normalization.

Critical Assessment of the Method
There is limited experience with KIM-1 in healthy subjects and in clinical pharmacology studies. The marker is not well established in its performance in nondisease states so far. Therefore descriptive analysis of the marker and analysis of traditional parameters such as serum creatinine or BUN together with KIM-1 in order to get more experience with the marker is advised.

Modifications of the Method
There are several additional bimarkers for assessment of renal toxicity like alpha-GST or NGAL. They are also well qualified in nonclinical toxicity studies with nephrotoxicants. There is only limited information about the normal ranges and the spontaneous variations available currently.

Visual Analogue Scale for Semiquantitatively Assessing Pain and Other Subjective Factors

Purpose and Rationale

A visual analogue scale (VAS) is a psychometric response scale, which can be used in questionnaires. It is a measurement instrument for subjective characteristics or attitudes that cannot be directly measured, for example, pain or subjective assessment of the effectiveness of a treatment.

For the quantification of these subjective factors, the VAS is an instrument that tries to measure the severity across a continuum from none to an extreme amount of the characteristic. For example, the spectrum of pain to a subjective suffering from it appears to be continuous and does not take discrete jumps, as the typical categorization of none, mild, moderate, and severe suggests. In order to reflect this idea of an underlying continuum the VAS was introduced.

Procedure

Operationally, a VAS is usually a horizontal line, 100 mm in length, anchored by word descriptors at each end, for example, “no pain” and “maximum pain.” When responding to a VAS, subjects are asked to indicate their level of agreement to a statement by indicating a position along a continuous line between the two end points mentioned. This continuous (or “analogue”) aspect of the scale differentiates it from discrete scales (e.g., “none-mild-moderate-severe” or “A to F”).

Evaluation

The VAS score is determined by measuring in millimeters from the left-hand end of the line to the point that the patient marks.

Critical Assessment of the Method

As such an assessment is clearly highly subjective, these scales are of most value when looking at change within individuals, and are of less value for comparing across a group of individuals at one time point. It could be argued that a VAS is trying to produce interval/ratio data out of subjective values that are at best ordinal. Thus, some caution is required in handling such data. Many researchers prefer to use a method of analysis that is based on the rank ordering of scores rather than their exact values, to avoid reading too much into the precise VAS score.

However, a VAS is extremely simple to use, easy to teach and understand. Therefore, bias introduced by complexity can be ignored.

For efficacy studies in patients, where pain is a primary or secondary outcome parameter, the VAS is only of limited value.

In practice, computer-analyzed VAS responses may be measured using discrete values due to the discrete nature of computer displays.

The VAS can be compared to other linear scales such as the Likert scale or Borg scale. The sensitivity and reproducibility of the results are broadly very similar, although the VAS may outperform the other scales in some cases [1].

Modifications of the Method

Due to the limitations mentioned above, several additional tools for pain assessment have been developed and validated, such as the McGill pain questionnaire, where several dimensions of pain are assessed. As for all questionnaires, it is very important to have the questionnaire available in the validated version of the native language. Otherwise the outcome of the questionnaires from different languages cannot be compared. These complicated and often patent protected questionnaires do not have a major place in clinical pharmacology studies.

Summary

It should be kept in mind that during the first clinical studies, there is practically no information about the safety and tolerability of a new drug as compared to the knowledge accumulated later on. Nevertheless, only during these initial studies the administration of the drug occurs under such kinds of secure conditions concerning the ability to handle side effects that dose escalation should not be stopped too early. It has to be kept in mind that during phase II and III studies and even more during marketing of a drug, the exposure of the drug to patients might occasionally – due to overdose, poor metabolization, or other causes of accumulation – be much higher than intended. Especially for these cases the company developing a drug should know, which kind of side effects would be expected.

References and Further Reading

Cancer Therapy Evaluation Program (2009) Common terminology criteria for adverse events (CTCAE) v4.0. Accessed 29 May 2009

Brentano CF, Menard J (2016) The BIAL/Bootrial case of death of a human volunteer in the first-in-human study of BIA 10-2474: are we missing the fundamentals. Presse Med 9:719–722

EMA (2017) Strategies to identify and mitigate risks for first-in-human and early clinical trials with investigational medicinal products. EMEA/CHMP/SWO/28367/07 Rev. 1

FDA Draft Guidance (2007) Guidance on industry drug induced liver injury: premarketing clinical evaluation. CDER, Rockville, MD, USA

FDA Guidance (2007) Toxicity grading scale for healthy adult and adolescent volunteers enrolled in preventive vaccine clinical trials, Rockville, MD, USA

FDA Guidance (2009) Drug-induced liver injury: premarketing clinical evaluation, Rockville, MD, USA

Herson J (2000) Data and safety monitoring committees in clinical trials. Chapman & Hall, Boca Raton, p 48

Horvath CJ, Milton MN (2009) The TeGenero incident and the Duff report conclusions: a series of unfortunate events or an avoidable event? Toxicol Pathol 37(3):372–383

http://www.circare.org/foia5/clinicaltrialsuspension_interimreport.pdf. Accessed 19 Jul 2010

http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/Vaccines/ucm074775.htm. Accessed 19 Aug 2010

ICH (2005) The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential of non-anti-arrhythmic drugs. ICH harmonized tripartite guideline E14, final version

Kratz A, Ferraro M, Sluss P, Lewandroski K (2004) Laboratory reference values. N Engl J Med 351:1548–1563

Lewis JH (2006) "Hy's law", the "Rezulin rule", and other predictors of severe drug-induced hepatotoxicity: putting risk benefit into perspective. Pharmacoepidemiol Drug Saf 15:221–229

Lock EA (2010) Sensitive and early markers of renal injury: where are we and what is the way forward. Toxicol Sci 116:1–4

NCI (2017) https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_8.5x11.pdf. Accessed 23 Nov 2018

Sibille M, Patat A, Caplain H, Donazzolo (2010) A safety grading scale to support dose escalation and define stopping rules for healthy subject first-entry-into-man studies. Br J Clin Pharmacol. https://doi.org/10.1111/1365-2125.2010.03741.x