CLINICAL INVESTIGATION

Managing safety in early phase trials: is there a gold standard?

"...it is time to encourage them to publish their best practices on what is considered acceptable in human clinical studies. The general term 'Thou shalt not harm' is in dire need of a scientific basis."

Keywords: healthy subjects • IMP • mode of action • Phase I • risk assessment • risk minimization • safety • toxicity

Anja Slikkerveer
Drug Discovery Science & Management, Drug Development Research, Astellas Pharma, Leiden, The Netherlands
anja.slikkerveer@astellas.com

Early clinical studies are intended to define safety and tolerability of a new investigational medicinal product (IMP). In such studies there is no benefit for the human subject participating and therefore the risk for harm for the subjects should be avoided. Since every change observed in clinical studies may currently be considered as harm, a framework for assessment needs to be defined. However, no definition of harm exists in the context of early clinical studies, so it is unknown what exactly we are preventing. If modest use of alcohol is considered acceptable, it may be inconsistent that a drug-induced increase in liver enzymes equaling one glass of wine is considered as harm. Equally, in a Phase I study, we may take precautions to prevent harm in the form of an increase in heart rate, which equals an effort comparable to walking up three flights of stairs or ten genuflexions. The human body is able to cope with this level of change as part of daily life, therefore it may be time to start taking this resilience into consideration. We should explore if the body can cope with limited effects of the IMP as well as it does with living. Understanding these processes is likely to improve early clinical safety characterization.

Before initiation of any clinical trial, the risks for the subject are balanced against the (potential) gains, such as improvements in the disease: the so-called risk–benefit assessment. Over many years a common practice has developed which allows the risk–benefit to be acceptable in the absence of benefit. There is not a gold standard for identifying and managing risk in early clinical trials, but there is a well-practiced sequence of steps which leads to adequate risk assessment and risk minimization for the subjects. This best practice is a forward-looking approach and starts with the interpretation of the results of nonclinical studies and knowledge with respect to the mode of action (MoA). A risk assessment does not only takes the study results into consideration but also has to take the weight of evidence into account. Although the sequence of steps involved in the risk assessment are the same for each IMP, the outcomes and the minimization measures may differ significantly.

The risk assessment

Data from multiple nonclinical disciplines need to be evaluated and integrated in order to obtain the best estimate of risk for the human subjects. The preclinical dataset includes an interpretation of the MoA of the IMP, the results of the classical toxicity package which contains the studies required by ICH M3(R2) [1] or ICH S9 [2] and the assessment of effects on vital functions (cardiovascular pharmacology and effects on the central nervous system). These are to be supported by the fundamental understanding of the fate of the IMP in the human body: its absorption, metabolism, distribution and elimination.

The assessment whether the pharmacology and the MoA can give rise to pharmacological responses which are not intended is an essential part of the risk assessment. This assessment is labor intensive and requires cooperation between pharmacologists, toxicologists

FUTURE SCIENCE part of fsg

and medically trained experts. The intended pharmacological effect ('intended effect') is the response under investigation during clinical studies. This does not mean that the intended effect is free of harm. The intended effect may be unwanted in healthy subjects in a Phase I trial (e.g., inhibition of immune response), while beneficial in a patient. Similarly, moderate pharmacological responses can be acceptable, while stronger responses may result in unacceptable unwanted effects (e.g., lowering of blood pressure or slowing of heart rate).

> "Presumptions are made that an adverse effect in animals would occur in humans at the same exposure level. This quantitative approach is a kind of double-edged sword: it can be useful but it can also blur the decision-making process..."

In addition to the assessment of the intended effect, receptor selectivity and species differences need to be considered. Complexity is added by modern drugs which are designed to be highly selective for the molecular targets. In the initial dose-escalation studies exploring dose ranges up to the highest tolerable dose, loss of receptor specificity is not rare. The possibility arises that the receptor (sub)type specificity is lower and the consequential events from interacting with other receptors may be adverse. Characterization of specificity of the receptor (subtype) in animals, can be complemented *in vitro* using human (subtype) receptors. This of course is only possible if the target is not too novel. It is an excellent common practice in pharmaceutical industry to screen IMP early for cross reactivity with other targets, either using a relevant panel of *in vitro* test or tissue (cross)reactivity assays. When the outcome of such screen is followed up properly, it can provide valuable clues on the ability of an IMP to trigger effects through other receptors than the target receptor. The number of identified receptors that may be useful to screen is ever-growing, and it is a challenge for companies to keep up.

Species differences with respect to pharmacological on-target and off-target responses need to be considered, since these are the main source of lack of predictivity [3,4]. The term species differences covers a wide range of receptor characteristics, such as receptor density, receptor distribution, receptor effect coupling and functionality of receptor subtypes, among others, in target tissue, as well as at the site where unwanted effects originate. The final pharmacological assessment is to evaluate if there is any evidence that the MoA of the IMP can be related to adverse responses, which have not been studied with the IMP itself. For these, an assessment is made to see if the weight of evidence is sufficient to warrant investigation and if data need to be generated prior to the start of human studies.

The classic toxicity studies are typically run in two animal species, one rodent and one nonrodent, because it is believed that between these two species a potential risk for humans would emerge. It needs to be ensured that anticipated major human metabolites are present in at least one of these animal species. Importantly, these studies are designed to show adverse effects, meaning that the doses should be so high that toxicity/adverse effects occur. Consequently, one should not be too worried observing toxicity, but rather assess the result for the target organs affected by the IMP when the doses reach a level where the effects can no longer be compensated. It is the compounds that do not show an effect at the highest feasible nonclinical doses which should raise concern, since no IMP specific precautions can be made.

The package for initiation of single and multiple dose studies of limited duration [1] provides information on the genotoxic potential of the IMP, its target organ profile, including effect on the reproductive organs and the respective no observed adverse event level (NOAEL). The target organ profile consists of knowledge of which organs show adverse responses, the nature of these responses, the dose response and the associated systemic exposures. In recent years, quantification of the exposure levels (C_{max} and/or AUC) as a way to predict when human side effects may occur, has gained a lot of attention. Presumptions are made that an adverse effect in animals would occur in humans at the same exposure level. This quantitative approach is a kind of double-edged sword: it can be useful but it can also blur the decision-making process, unnecessarily limiting the dose escalation in Phase I studies.

> "The presence or absence of risk minimization options drives the decision to proceed to clinical studies or the selection of the subpopulations in which the study can be performed."

How systematic the above approach may seem, each individual decision on a risk, is made conservatively. Over the years the risk assessment for Phase I studies has grown very conservative, leading to inherent questions on the current ability to separate signal and noise. In case one cannot be certain, any observation tends to be classified as a signal. This dilemma should trigger a debate on how to separate signal from noise, both in nonclinical and early clinical studies. Overall, there is a need for a better approach to judge the nonclinical safety data. Grouping of observations and findings into pathophysiological entities, an approach which has been proven its merits in the field of medicine, may be a possible way forward.

The risk management/risk minimization

After the risk assessment, the process continues with the assessment how the recognized potential adverse effects can be detected and managed in humans (risk minimization). The presence or absence of risk minimization options drives the decision to proceed to clinical studies or the selection of the subpopulations in which the study can be performed. Commonly the subjects in Phase I studies are healthy and young. Nevertheless specific effects should rather not be tested in healthy young subjects, for example, an IMP with an effect on female reproductive organs, will generally be tested in postmenopausal or surgically sterilized women, while a mutagenic compound may be dosed for the first studies already in patients in which a positive risk–benefit could be present in the case the IMP was effective.

The nature of the potential adverse effect provides guidance on the appropriate measurement tools (e.g., biomarkers and functional tests) and other risk minimization measures (e.g., continuous cardiac monitoring). The nature of the effect also indicates if the adverse response can be expected to be detectable, serious/life-threatening, treatable or reversible. At the end of the process, the balance can be made between the potential human risks, the available risk minimization and the potential benefits.

Despite the above IMP-specific assessment, Phase I protocols of single ascending dose studies and multiple ascending dose studies are very similar since the majority of assessments are standardized. On occasion the risk assessment finds a target which requires additional parameters to be added to the already extensive standard measurements. The most impactful risk minimization measure is not explicitly mentioned: Phase I studies are performed under close clinical observation – the subject is typically staying in a special unit for the duration of the study.

The quintessential risk minimization measure is not recognized in a Phase I protocol: it is the skill to identify for which IMP the human safety cannot be guaranteed and to stop those compounds from proceeding to clinical development.

Companies have been developing extensive internal approaches to assess and minimize risk in early clinical studies, I think it is time to encourage them to publish their best practices on what is considered acceptable in human clinical studies. The general term 'Thou shalt not harm' is in dire need of a scientific basis.

Disclaimer

This editorial reflects the author's personal opinions and does not necessarily reflect the opinion of Astellas.

Financial & competing interests disclosure

A Slikkerveer is an employee of Astellas Pharma Europe located in The Netherlands, where she was involved in the subject matter for many years. This allowed her to develop the ideas reflected in this paper. The author has no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

1 ICH Expert Working Group. *Harmonised Tripartite Guideline: Guidance on Non-clinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals M3(R2).* ICH. Geneva, Switzerland (2009).

2 ICH Expert Working Group. *Harmonised Tripartite Guideline: Nonclinical Evaluation for Anticancer Pharmaceuticals S9.* ICH, Geneva, Switzerland (2009).

3 Lounkine E, Keiser MJ, Whitebread S *et al.* Large-scale prediction and testing of drug activity on side effect targets. *Nature* 486(7403), 361–367 (2012).

4 Bowes, J Brown AJ, Harmon J *et al.* Reducing safety-related drug attrition: the use of *in vitro* pharmacological profiling. *Nat. Rev. Drug Discov.* 11(12), 909–922 (2012).