ICH Guidelines: Inception, Revision, and Implications for Drug Development

Tao Wang,*,1 David Jacobson-Kram,† Anne M. Pilaro,‡ Daniel Lapadula,† Abigail Jacobs,† Paul Brown,† John Lipscomb,§ and William David McGuinn†

Since the inception of the International Conference on Harmonization of Technical Requirements for the Registration of Pharmaceuticals for Human Use (ICH) in 1990, six-party Expert Working Groups (EWG) have developed and revised numerous guidelines on preclinical safety evaluation (Table 1). The six parties to ICH represent the regulatory bodies and pharmaceutical manufacturing organizations in the three regions: Europe, Japan, and the United States, where the majority of new medicines are developed. The development of ICH guidelines is a stepwise process. In step 1, the EWG prepares a “final harmonized draft.” Step 2 entails forwarding the draft to the Steering Committee for signature, which signifies acceptance for consultation. Step 3 is a process of consultation and discussion with regulatory organizations within the three regions and usually takes about 6 months or longer. Under step 3, the step 2 guideline is also released for public comments. Step 4 generates the Experts Document as a result of the consultation and is then submitted to the Steering Committee. Step 4 is reached when the Steering Committee endorses the adoption of the guideline by the regulatory bodies of the three regions. In the final step (step 5), the guidelines are incorporated into national or regional internal procedures, and the process is completed.

The experience gained over the past decades, combined with the refined approaches for the generation of new drugs, has led to the recent ICH initiatives for developing new guidelines and revisiting some of the existing guidelines. The purpose of these initiatives is to optimize the scientific and technical aspects of the drug development process, ultimately accelerating the development of safe and effective medicines, while reducing animal use. The major ICH initiatives (bold in Table 1) include a new guideline, ICH S9, for preclinical evaluation of anticancer pharmaceuticals, an addendum to ICH S6 regarding preclinical safety evaluation of biologics, revision of ICH M3 which addresses the timing of preclinical studies in relation to various stages of clinical development, and a proposed new guideline on genotoxicity testing (ICH S2) that replaces and combines the ICH S2A and S2B. In this paper, we present the rationale behind these initiatives at ICH and some interpretation of the revised and new guidelines. We also provide commentary and perspective on the potential impact these new guidelines may have on preclinical safety evaluation programs.

ICH S2(R1): PROPOSED REVISIONS ON GENOTOXICITY TESTING

Before the initiation of the first clinical trial of a new drug candidate, regulatory agencies across three regions require a battery of tests to ensure the safety of clinical trial participants. This battery includes tests to determine whether the drug under study is potentially genotoxic. These studies are described in the ICH S2A and S2B originally published in 1995 and 1997, respectively. It is well known and accepted that gene mutations and chromosomal rearrangements are often the seminal events in the generation of human cancers (Bos, 1989; Yunis, 1983). Prior to their approval, most drugs also undergo carcinogenicity testing. These studies are usually done in parallel with late-stage clinical trials. Hundreds or thousands of patients can be treated for many months with pharmacological drug doses without the knowledge of potential for carcinogenicity. The Food and Drug Administration (FDA) uses the results of genotoxicity studies as a placeholder until the results of the carcinogenicity tests are available. When phase 1 studies are performed in healthy volunteers, there is clearly no risk/benefit paradigm for this population, only potential for risk. A positive result in
a genotoxic assay can be the basis for a clinical hold or at the very least the basis for additional testing to demonstrate a lack of risk for the clinical trial participants.

The current ICH battery calls for a bacterial mutation assay (Ames test), an in vitro mammalian cell assay (metaphase analysis for chromosome breakage or a mouse lymphoma thymidine kinase gene mutation assay), and an in vivo rodent bone marrow assay for chromosomal damage (rodent micro-nucleus assay). The two in vitro assays are required prior to phase 1, whereas the in vivo assay is required prior to phase 2. In practice, the majority of sponsors perform all three assays simultaneously. A positive response in the in vivo assay is a high hurdle for continued development, and consequently sponsors want this information early on. The guideline also provides direction on what sponsors should do if an in vitro assay gives positive results: “positive result in vitro is followed up by a second in vivo study—using tissue other than bone marrow.” The nature of this second study has evolved over the years. The EWG for the original guideline most likely had the rat unscheduled DNA synthesis (UDS) assay in mind as the follow-up test because it uses a different tissue and also a different endpoint. Moreover, a large database existed for the liver UDS assay. Although this was a reasonable suggestion at the time, experience over the last 15 years has shown that this assay almost always gives negative results with pharmaceuticals. UDS detects a DNA excision repair process used to remove bulky adducts (Mirsalis and Butterworth, 1980). In practice, few if any new drugs possess such properties. Once the insensitivity of this assay became clear, Center for Drug Evaluation and Research (CDER) began requesting the Syrian hamster embryo (SHE) cell transformation assay as a follow-up to in vitro positive results. This assay also has limitations. First, it is obviously not an in vivo assay and does not directly assess genotoxicity. Second, whereas the SHE assay has been shown by some investigators to have a good correlation with the rodent bioassay (Pienta et al., 1977), other studies suggest that it poorly predicts human carcinogens (Mauthe et al., 2001). Finally, the SHE cell assay suffers from many technical challenges, such as qualification of cell and serum lots and overlapping frequencies of negative and positive controls. Most recently, CDER has been accepting the results of in vivo “comet” assay as a follow-up to in vitro positives. Although the 2006 CDER guidance on genotoxicity discusses in vivo options to follow up in vitro positive results, the spectrum of choices is limited (http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm079257.pdf). In vivo mutation assays with transgenic mice are time and resource intensive and have also been found to relatively insensitive. The in vivo comet allows for assessment of DNA damage in a wide variety of tissues, and established protocols and procedures have been developed in many laboratories (Rothfuss et al., 2010).

Because of the many developments in the science of genetic toxicology and in the practices of drug development in the nearly 15 years since the original guidelines were published, a proposal to update the guidelines was made to and accepted by the ICH Steering Committee. The rationale behind the proposal was that (1) new assays and scientific knowledge should be taken into account to improve genotoxicity risk assessment, (2) high frequencies of positive results are consistently seen in the in vivo mammalian cell assays that are not confirmed in the in vivo genotoxicity assays or carcinogenicity studies, and (3) there was a need to consider the reduction of animal use (without affecting patient risk). An EWG consisting of genetic toxicologists from pharmaceutical industries and regulatory agencies in the United States, Europe, and Japan began revising the S2 guidelines in 2006. In March of 2008, a draft revision was published on the European Medicines Agency (EMA) Web site

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Code</th>
<th>Title</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1A</td>
<td>Need for carcinogenicity studies of pharmaceuticals</td>
<td>Step 5, 1995</td>
</tr>
<tr>
<td>S1B</td>
<td>Testing for carcinogenicity of pharmaceuticals</td>
<td>Step 5, 1995</td>
</tr>
<tr>
<td>S1C(R2)</td>
<td>Dose selection for carcinogenicity studies of pharmaceuticals</td>
<td>Step 5, 2008</td>
</tr>
<tr>
<td>S2(R1)</td>
<td>Guidance on genotoxicity testing and data interpretation for pharmaceuticals intended for human use</td>
<td>Step 3, 2008</td>
</tr>
<tr>
<td>S3A</td>
<td>Note for guidance on toxicokinetics: the assessment of systemic exposure in toxicity studies</td>
<td>Step 5, 1994</td>
</tr>
<tr>
<td>S3B</td>
<td>PK: guidance for repeated dose tissue distribution studies</td>
<td>Step 5, 1994</td>
</tr>
<tr>
<td>S4</td>
<td>Single-dose toxicity tests</td>
<td>Agreement, 1991</td>
</tr>
<tr>
<td>S4</td>
<td>Duration of chronic toxicity testing in animals (rodent and nonrodent toxicity test)</td>
<td>Step 5, 1998</td>
</tr>
<tr>
<td>S5(R2)</td>
<td>New title: detection of toxicity to reproduction for medicinal products and toxicity to male fertility</td>
<td>Step 5, 1995</td>
</tr>
<tr>
<td>S6</td>
<td>Preclinical safety evaluation of biotechnology-derived pharmaceuticals</td>
<td>Step 5, 1997</td>
</tr>
<tr>
<td>S6(R1)</td>
<td>Addendum to ICH S6: preclinical safety evaluation of biotechnology-derived pharmaceuticals</td>
<td>Step 3, 2009</td>
</tr>
<tr>
<td>S7A</td>
<td>Safety pharmacology studies for human pharmaceuticals</td>
<td>Step 5, 2000</td>
</tr>
<tr>
<td>S7B</td>
<td>The preclinical evaluation of the potential for delayed ventricular repolarization (QT interval prolongation) by human pharmaceuticals</td>
<td>Step 5, 2005</td>
</tr>
<tr>
<td>S8</td>
<td>Immunotoxicity studies for human pharmaceuticals</td>
<td>Step 5, 2005</td>
</tr>
<tr>
<td>S9</td>
<td>Preclinical evaluation for anticancer pharmaceuticals</td>
<td>Step 5, 2009</td>
</tr>
<tr>
<td>M3(R2)</td>
<td>Guidance on preclinical safety studies for the conduct of human clinical trials and marketing authorization for pharmaceuticals</td>
<td>Step 5, 2009</td>
</tr>
</tbody>
</table>
and in the United States Federal Register. Comments on the draft were received and responded to, and the guideline was modified and prepared for final signature. Major changes to the guideline include the following:

(1) The in vitro micronucleus test is listed as an option for an in vitro mammalian cell assay for chromosomal breakage.
(2) In the absence of toxicity and precipitation, the limit dose in vitro mammalian cell assay for chromosomal breakage.
(3) Limit of solubility is defined as the first dose resulting in a visible precipitate.
(4) Limits of toxicity are reinforced, 50% for in vitro chromosomal aberrations and in vitro micronucleus assay and 80% relative total growth in the mouse lymphoma assay.
(5) Repeat assay not required for adequately performed negative bacterial mutation assay.
(6) Positive control is not required with every in vivo assay. Positive control slides are to be included when the assay is scored and periodic qualification with a positive control is proposed.
(7) A second battery of tests for assessing genotoxicity is offered. The second option includes a bacterial mutation assay (Ames test) and an in vivo test with two endpoints, e.g., bone marrow micronucleus and assay for DNA strand breakage in liver cells (e.g., comet assay).
(8) The revised guidance also encourages sponsors to include genotoxicity endpoints in routine repeat dose toxicology studies whenever reasonably possible instead of doing separate acute studies.

Many aspects of the revised guideline will result in the reduction of animal use for genotoxicity testing. Perhaps, the biggest impact will result from a reduction in false-positive results often generated in in vitro mammalian assays with resulting in vivo follow-up testing. The guideline currently remains at step 3 awaiting final signature.

ADDENDUM TO ICH S6: PRECLINICAL SAFETY EVALUATION OF BIOTECHNOLOGY-DERIVED PHARMACEUTICALS

Since its implementation in 1997, the ICH S6 has provided a description of the appropriate timing and framework for conducting preclinical safety studies for protein therapeutics derived using recombinant DNA technology. However, the experience over the past 15 years combined with the technological advances and refinements in the generation of new biopharmaceuticals have led to identification of areas within the ICH S6 that require clarification or expansion to harmonize implementation while retaining the “state of the science” approach to safety testing. The EWG agreed that there were five specific topics that required priority for updating, including the following: (1) species selection, (2) study design, (3) immunogenicity assessments, (4) developmental and reproductive toxicity testing, and (5) carcinogenicity testing. These five topics were discussed and approved by the ICH Steering Committee for incorporation as an addendum to the original ICH S6. The ICH Steering Committee was adamant that the language in the original ICH S6 would stand as is and that the addendum would provide clarification and expansion of areas in the original guidance that were either problematic in implementation across the three regions or would address areas where the science had significantly advanced since the original ICH S6 was adopted in 1997. The addendum is hereafter referred to as ICH S6(R1).

Species Selection

ICH S6(R1) discusses the need for a second test animal species and provides clarification that although testing in two pharmacologically responsive or “relevant” species is expected, there is no need to create either a transgenic animal expressing the human target or a homologous protein to the rodent orthologous target when the clinical candidate is only pharmacologically active in nonhuman primates and man. The point here is that homolog should not be created solely for the purpose of generating preclinical safety data in a second species. Nevertheless, ICH S6(R1) does provide clarification on when these models might be useful for hazard identification purposes, e.g., to obtain information to inform patients and physicians of reproductive and developmental risk when the clinical candidate is only active in humans or chimpanzees. Furthermore, the addendum proposes that when similarities in pharmacodynamic response and short-term toxicity (i.e., up to 4 weeks of repeat dose testing) are noted between two pharmacologically relevant species for a specific clinical candidate, further toxicology testing in a single species may be acceptable, provided that appropriate scientific justification for the species selected for testing is conveyed in support of this approach.

The original ICH S6 directs testing of a monoclonal antibody clinical candidate using tissue cross-reactivity studies with a panel of tissue sections from human and different test animal species to identify the relevant species for further toxicology testing. Discussion among the EWG members at the ICH meetings and their respective home parties revealed that this method no longer should be the sole criterion for selecting a pharmacologically responsive species, and language to this effect was included in the addendum. The basis for this position by the EWG was that these assays generally are not sensitive enough to detect low-level binding that may still be important for safety and pharmacologic response. Other assays including binding affinity to the target protein or flow cytometry or in vitro functional assays using cells from humans and test animal species provide more sensitive and quantitative evaluations of the comparable responses of the test species and should be used whenever feasible for species selection. This position was incorporated into the step 2 document for ICH S6(R1) and has proven to be one of the more controversial aspects of the addendum.
Study Design

One of the most difficult areas for implementation of the original ICH S6 across the three regions was selection of the highest dose for repeat dose toxicity studies. The EWG discussed whether to recommend the use of fixed multiples of the highest anticipated clinical dose; however, it was quickly recognized that the same pharmaceutical may be used across a number of different indications, e.g., multiple sclerosis, rheumatoid arthritis, and cancer, and that the highest dose for each of the indications might differ by orders of magnitude. Although unable to agree upon a “one size fits all” approach for high-dose selection, the EWG agreed that for a biopharmaceutical whose clearance is mediated by the target protein or receptor, once maximal pharmacodynamic effect or receptor saturation has occurred, there is no reason to explore higher doses. Once all targets are engaged, it is expected that the maximum pharmacologic effect and any toxicities that might be extensions of that effect would occur. Increasing the dose, e.g., for a monoclonal antibody, would not be expected to increase the response but instead would result in prolonged exposure, as the clearance of most antibodies is mediated by their targets. Once all targets are engaged, additional clearance can only occur if the target is either reexpressed or new cells expressing the target (e.g., hematopoietic cells) are generated. Therefore, the EWG recommended in the step 2 guidance that the maximal dose selected for repeat dose toxicity studies should be the highest of either (1) a dose that gives the maximum intended pharmacological effect in the preclinical test species (i.e., taking into account any differences in affinity or pharmacological response of the test species) or (2) a dose that gives up to a 10-fold exposure multiple over the maximum, estimated exposure to be achieved in the clinic, following corrections for differences in target-binding affinity or potency between the human and the test animal target.

ICH S6(R1) addendum also addresses the duration of chronic toxicity studies and inclusion of recovery periods. The EWG concurred that based on the data available to date (Clarke et al., 2008), a 6-month toxicity study in a pharmacologically responsive species was sufficient to achieve long-term exposure and assess the potential for toxicity. Increasing the duration of the study, especially when it includes at least one dose level at which the maximal pharmacodynamic effect was achieved, was not likely to provide any additional useful information. Additionally, although reversibility of toxicity reported in animal studies should be understood, the ICH S6(R1) states that complete recovery is not essential during the course of the study as long as the reversibility of an adverse effect is well understood or an adequate margin of safety between the animal doses and the clinical dose levels can be achieved.

Immunogenicity

The original ICH S6 states that development of antidrug antibodies (ADA) is expected when human biopharmaceuticals are introduced in test animals and infers that measurement of ADA should be performed during repeat dose toxicity studies. The ICH S6 guidance also states that the ADA responses should be characterized in terms of antibody isotype, titer produced, and whether they are neutralizing or nonneutralizing of the function of the target in order to aid in interpretation of the study results. However, experience over the last 15 years has shown that frequent detection of ADA may not be feasible or may be confounded when high levels of the biopharmaceutical are still present in serum or plasma from the test species (Ponce et al., 2009). In these cases, ICH S6(R1) recommends that measurement of ADA need not necessarily be conducted, especially if other results, e.g., sustained pharmacodynamic effect, no evidence of vasculitis, or other immune-mediated toxicity, suggest that characterization of an ADA response is not needed in order to interpret the findings from the repeat dose toxicity study.

Developmental and Reproductive Toxicity Testing

By far, the most controversial topic discussed by the EWG and included in the ICH S6(R1) addendum was the assessment of developmental or reproductive toxicity, particularly for those biopharmaceuticals where the clinical candidate is only biologically active in humans and nonhuman primates. The original ICH S6 does neither provide specific direction on which test animal species to use nor the design of the studies or numbers of animals to use per dose group and thus required further explanation in the addendum for clarity and consistency. In order to clarify the language in the original guidance, the ICH S6(R1) states that when the clinical candidate is pharmacologically active in rodents and rabbits, these species should be used in place of nonhuman primates whenever possible, appropriate scientific justification should be provided to support the choice of test species, and the study designs should follow the guidance set forth by ICH S5(R2). However, for products that are only active in humans and nonhuman primates and therefore necessitate testing in monkeys, alternative study designs to the harmonized ICH S5(R2) stages C–D (embryo-fetal developmental toxicity study) and D–E (peri- and postnatal developmental toxicity studies) are discussed. Specifically, testing of the clinical candidate in the nonhuman primate is preferred over testing an homologous protein in the rodent; however, in some cases, e.g., when the biology of the target suggests an adverse effect on pregnancy outcome, ICH S6(R1) suggests that use of alternative approaches to testing in nonhuman primates may provide sufficient hazard identification to communicate risk in product labeling. For products where the clinical candidate necessitates testing in the nonhuman primate because of species specificity of its interaction with the target, the addendum also discusses obtaining fertility endpoints such as gonad weights and histopathology, sperm counts, motility and viability, and male or female reproductive hormones in chronic, repeat dose toxicity studies in place of conducting stand-alone, harmonized ICH S5(R2) stages A–B reproductive toxicity studies.
One of the key pieces of information learned since the original ICH S6 was implemented in 1997 is that transplacental transfer of therapeutic proteins, specifically monoclonal antibodies, occurs by specific mechanisms in humans and nonhuman primates and that expression of the receptor required for this transfer does not occur until near the end of the organogenesis period (Pentsuk and van der Laan, 2009; Schlamowitz, 1976; Simister, 2003). Therefore, the developing embryo or fetus may not receive significant exposure during the period when organogenesis occurs, potentially resulting in a negative study if a strict embryo-fetal developmental study design (i.e., ICH S5(R2) stages C–D) is employed. For this class of biopharmaceuticals, ICH S6(R1) recommends an alternative, combined embryo-fetal and peri- and postnatal study design (i.e., ICH S5(R2) stages C–E) for testing of the clinical candidate in nonhuman primates. Treatment of the pregnant females would initiate at gestational day 20 and continue up until parturition in order to maximize exposure to the developing offspring. This study design would not obtain nonhuman primate fetuses at gestation day 100 for traditional teratology evaluations but would essentially use the same endpoints as available for clinical follow-up of fetal outcome, specifically live birth and neonatal survival, external malformations, skeletal effect by x-ray evaluation, and visceral morphology and histopathology at necropsy. The ICH S6(R1) suggests that these studies be used for hazard identification purposes only because it is not practical to statistically power a developmental study in nonhuman primates for true assessment of risk. Furthermore, the addendum suggests that because these studies are conducted only for identification of hazard, it may be possible to conduct the harmonized ICH S5(2R) stages C–E study in nonhuman primates using only a control and a single-dose group (e.g., at maximal pharmacodynamic activity) to obtain adequate evidence of hazard to embryo-fetal and peri- and postnatal development (Jarvis et al., 2010). Although the study designs recommended in ICH S6(R1) clearly represent significant advances in the understanding of how large molecular weight biopharmaceuticals cross the placenta and subsequent embryo-fetal exposure and identification of risk, there is still considerable controversy about the number of animals per test group, the number of dose levels to be tested, and whether or not these design principles can be applied to biopharmaceuticals other than monoclonal antibodies. These issues remain to be resolved.

Carcinogenicity Assessment

The original ICH S6 states that “standard carcinogenicity bioassays are generally inappropriate for biotechnology-derived pharmaceuticals. However, product-specific assessment of carcinogenic potential may still be needed . . . . When there is a concern about carcinogenic potential, a variety of approaches may be considered to evaluate risk.” The original ICH S6 also states that when the product is biologically active and not limited by development of ADA in rodents and when the available information regarding the biology of the target, the patient population, and the duration of treatment indicates a need to communicate risk of tumorigenesis in labeling, carcinogenicity testing in a single rodent species may be warranted. This language needed further clarification, specifically regarding the types of information needed to provide an adequate assessment of carcinogenic potential for a new biopharmaceutical. The ICH S6(R1) document provides clarification of what types of information may be acceptable to provide appropriate, scientific evidence of the potential for a biopharmaceutical to induce or stimulate tumor growth, with the goal of communicating risk to patients and their prescribing physicians, in addition to informing any potential risk management plans. The addendum discusses several different ways to obtain this information when sufficient information may be considered in place of conducting animal studies and when rodent bioassays might provide useful information for labeling, as well as when they might not. For example, if the weight of evidence from the literature, any mechanistic studies, and the chronic toxicity studies do not suggest carcinogenic potential, ICH S6(R1) does not recommend a rodent bioassay. Similarly, if the weight of evidence strongly suggests that there is a concern, e.g., that the product is severely immunosuppressive and therefore expected to result in viral recrudescence and subsequent tumorigenesis, ICH S6(R1) recommends that the risk be communicated through labeling and that rodent bioassays may not provide additional, useful information. These points were discussed extensively among the EWG members and their respective regional member parties, and as a result, this section appears to be generally acceptable to all regions.

In summary, the original ICH S6 guidance provides a basic and flexible framework for designing preclinical safety studies for biopharmaceuticals, recognizing that there is no “one size fits all” approach that can be applied for all therapeutic proteins. The ICH S6(R1) addendum was drafted to provide updates to the original ICH S6 guidance based on advances in the science. The guidelines are intended to be used together, with ICH S6(R1) supplementing, and in some cases superseding the recommendations provided in the ICH S6. When completed, ICH S6 will be reissued together with the addendum as the step 4 ICH S6(R1) guidance document.

NEW ICH S9: PRECLINICAL DEVELOPMENT OF ONCOLOGY THERAPEUTICS

The international pharmaceutical community has long recognized that the development of new cancer therapies needs to be efficient and expeditious because of the life-threatening nature of neoplastic disease. The world wants and needs new cancer treatments brought to market quickly and with a minimum use of experimental animals. But until recently, this universal demand has led to independent and sometimes disparate
approaches to the preclinical development of oncology drug products in Europe, the United States, and Japan. The EMA released a guideline on preclinical development of anticancer medicine in 1999 (http://www.ema.europa.eu), focusing primarily on anticancer drugs that were cytotoxic or cytostatic. In the United States, DeGeorge et al. (1998) at the FDA published a paper entitled “Regulatory Consideration for Preclinical Development of Anticancer Drugs.” This important paper and the invaluable flowchart published on the FDA Web site describing how to properly select the starting dose for experimental cancer therapies have been used extensively to guide oncology drug development. The DeGeorge’s paper also described the core preclinical studies needed to support the development of cancer therapies, but this paper and the EMA document focused only on therapies based on small molecules. They did not address the development of biological compounds. As a result, pharmaceutical companies have been without specific guidance on the development of this class of compounds.
In the last decade, the Ministry of Health, Labor, and Welfare (MHLW) in Japan also worked toward the promulgation of a guidance to address the need for preclinical studies to support the marketing of new anticancer therapies. The absence of ICH guidance had led to a hodgepodge of requests from regulatory authorities to industry wherever an anticancer therapy was developed, with ICH M3 being used as the basis for many requests for preclinical studies. The FDA, EMA, and MHLW all recognized that the release of independent guidelines would lead to widely varying standards and inefficient drug development. In addition, a guiding principle of ICH is to reduce, refine, and replace the use of animals, whenever possible. It is recognized that the disparate guidance available and nonharmonization had led to a greater use of animals. This recognition has led to the formulation of a coordinated guidance within the established aegis of ICH S9. Herein we present the key principles of ICH S9 and rationale behind these principles.

The new ICH S9 applies to small molecule and biotechnology-derived pharmaceuticals. Its scope focuses on the initial treatment of cancer in patients with late stage or advanced disease. These are patients whose disease has failed to respond to standard of care or have progressive disease, disease for which there are no remaining standard treatment options, disease with limited life expectancy, or disease where current therapy is considered ineffective. ICH S9 is intended to provide an outline of the minimum preclinical information needed to initiate phase 1 studies in this population. The early preclinical data and early (phase 1) clinical data could be used to start phase 2 not only in this population but also in first and second line cancer therapy in patients with advanced cancer. The document then describes the further information needed to support continued development and marketing of drug candidates for patients with life-threatening disease. ICH S9 is not intended as guidance for the development of therapies intended for proliferative diseases with a good prognosis or expected prolonged survival, cancer prevention, the treatment of symptoms or side effects of other chemotherapies, studies in healthy volunteers, vaccines, cellular or gene therapy, or radiopharmaceuticals. ICH M3, S6, or other guidance discuss the development of such therapies.

ICH S9 outlines the reasons for obtaining a minimum of pharmacology information prior to the initiation of clinical development. These reasons include the need to provide preclinical proof of principle, to estimate parameters that will influence schedule dependency and dose escalation, to provide information on species selection, and to guide starting dose selection and selection of investigational biomarkers. During the development of the guidance, there was a great deal of discussion around the level of detail that the document should provide. The guidance does not provide a comprehensive list of studies that must be done prior to entry of the drug into phase 1 development but rather discusses the sound scientific principles that guide the selection of the studies that should be done in particular situations. The document does specify the safety pharmacology parameters that should be defined to support phase 1 development, including cardiovascular, respiratory, and central nervous system effects. These assessments can be made within the conduct of the general toxicity studies to minimize the unnecessary use of animals. Stand-alone safety pharmacology studies are usually not needed, with the caveat that the regulatory agency may request such studies if a significant toxicity is identified that could put patients at additional risk. In that case, studies described in ICH S7A or S7B may be requested. The document specifies that the initiation of phase 1 studies requires a minimum of pharmacokinetic (PK) information and stresses that this information may be critical to the determination of dosing schedule and dose escalation. But the document also states that the determination of PK parameters within the context of supporting toxicology studies is efficient and usually sufficient.

Because clinical trials in advanced cancer investigate a patient population with a limited life expectancy, it is critical to give these patients hope that the therapy may be of some benefit. The EWG recognized that this requires an approach to specifying a clinical starting dose that is quite different from that used for most drugs. Optimally, the clinical oncologists should be able to achieve a significant systemic exposure without compromising patient safety. It is also desirable to specify a starting dose that is high enough to diminish the number of dose escalation steps needed to reach the maximum tolerated dose (MTD) or phase 2 dose, the dose that is most likely to be effective for many oncology drugs. Therefore, supporting toxicology studies need not necessarily identify a no-observed-effect level or no-observed-adverse-effect level (NOAEL) because toxicity is usually an expected consequence of oncological therapy. Oncologists rarely use most cancer drugs at or below the NOAEL. One of the approaches to identifying a starting dose for clinical trials is to use a well-designed investigational new drug (IND)-enabling study to identify a dose that causes severe toxicity in 10% of the rodents (STD10) or the highest nonsignificantly toxic doses (HNSTD) in nonrodents as originally specified by DeGeorge et al. (1998). If such a study also provides some indication of the steepness of the toxic dose response curves, this information can be used to guide clinical dose escalation. A variety of different toxicities can define an STD10 in rodents depending on the mechanism of the drug. An HNSTD in nonrodents is the dose above which lethality, life-threatening toxicities, or irreversible toxicities occur. Nevertheless, the guidance points out that all available data should be used in determining the starting dose, and it was clear that the EWG did not intend this to be prescriptive. Other approaches can, and should, be used, as justified by not only the toxicity data but also all the data including pharmacodynamic and PK information. The EWG specifically discussed the information needed to support the development of biologically derived therapeutics and in particular monoclonal antibodies, based on the tragedy that occurred with TGN1412 (Suntharalingam et al., 2006). TGN1412, a humanized monoclonal antibody, was developed as an immunomodulatory drug.
In its first human clinical trials in March 2006, it caused catastrophic systemic organ failure in at least four out of six volunteers, despite being administered at a subclinical dose of 0.1 mg/kg, which was approximately 500 times lower than the dose found safe in animals. The EWG did not want this one case to unduly influence the guidance for the development of monoclonal antibodies in cancer therapy and therefore recommended limiting the use of minimally anticipated biologic effect level as a basis for determining starting doses only to protein therapeutics with immune agonistic properties. This allows for the toxicity data to be used as the basis of initial dosing with antibodies and, in all other cases, allows for the main principle to be achieved, i.e., having a pharmacological effect from the first dose and being reasonably safe to use. One other principle that the document implicitly encourages and explicitly allows is the use of limited animal data, in essence 4-week preclinical data, to support continued clinical dosing beyond 4 weeks in phase 1 and 2 clinical trials. This is a radical departure from drug development in other indications as specified in ICH M3, where there is usually an expectation that the duration of preclinical studies used to support clinical development will be as long as or longer than the duration of clinical treatment. The reason for this, and the discussion in the EWG centered on, is ethical considerations for patients in this desperate population. It would clearly be unethical to stop a patient from continuing treatment if they are seeing a benefit. If there were no preclinical data past 4 weeks duration, patients were not permitted to continue on the experimental drug past that time point under some previous regulatory guidance. This was clearly a problem in patient recruitment because few of any patients would start a trial where they could not continue receiving a therapy that was providing objective or perceived clinical benefit. Previous requirements for full support of the clinical schedule also sometimes encouraged the unnecessary use of animals, as some pharmaceutical companies would start multiple animal studies (4-, 13-, and 26-week studies) in parallel in order to accumulate preclinical safety data ahead of progressing clinical trials. By allowing 4-week studies to support continued development in phases 1 and 2, animal use will be curtailed, and only drugs showing clinical promise will need to have further animal testing. ICH S9 also provides an outline of studies needed to support clinical studies with various schedules and provides a table of examples. The examples are not meant to be an exhaustive listing but only some of the more common paradigms. The document states that the duration and schedule of supporting preclinical studies should reflect the anticipated clinical schedule but also emphasizes the use of science in selecting the animal dosing schedule, with the use of toxicokinetic data playing a key role in the design of these schedules. It is more important to approximate the expected clinical exposure than the exact dosing schedule. Animals may need to be dosed at a more aggressive schedule than that of the clinical trial when the drug is expected to have a long half-life in patients and shorter one in animals. For example, if a half-life for a monoclonal antibody is 1 week in nonhuman primates and the clinical schedule is monthly dosing with the estimated half-life of the drug in patients of 1 month, instead of mimicking clinical monthly dosing schedule, a more frequent dosing in the preclinical studies might be needed. The document emphasizes that the drug formulation and route of administration should be comparable to those proposed for the clinical studies. Therefore, when designing a toxicology study to support a clinical trial, toxicologists should work closely with clinicians, clinical pharmacologists, and chemists to identify the appropriate dosing schedule, route, and formulation prior to the initiation of IND-enabling preclinical studies. Significant variation of the clinical schedule and route or formulation may result in the need for new or additional supporting toxicology studies. Most cancer drugs cause significant toxicity, and treating physicians are accustomed to providing supportive care to aid the patient’s recovery. Knowledge about the ability to recover from a particular toxicity is important; nevertheless, many of the toxicities that occur clinically have very well-known pattern of recovery (e.g., neutropenia). There was extensive discussion in the EWG over the need for recovery arms in preclinical studies. The guidance was therefore worded in a very specific way to describe the mechanisms by which recovery might be assessed. The document does not specify a need for a recovery group in all preclinical studies but rather the need to assess the likelihood of recovery. This assessment needs to be justified by the science around the particular toxicity. Nevertheless, if a severe toxicity is observed at the projected clinical exposure or if recovery cannot be predicted, inclusion of recovery groups in toxicology studies would most likely be needed. If recovery groups are included in the study, it is not necessary to observe a complete recovery of the lesion; it is only necessary to demonstrate that recovery is ongoing.

The preclinical studies done to support phase 1 development and clinical data collected during the phase 1 are usually sufficient to support initiation of phase 2 trials not only in advanced cancer but also to support further development in first or second line therapy. In order to define what toxicity studies will be needed to support further clinical development and marketing, the EWG was asked a key question: “Were any new toxicities found in chronic toxicity studies that altered the clinical development of cancer therapies?” It was not “Were there any new toxicities found?”, but rather their impact on the clinical development that was questioned. The EWG assessed preclinical study information collected from all regions and determined that within the available data, studies of 3-month duration identified all toxicities relevant to continued clinical development. There were no cases of new toxicities found in chronic toxicity studies that affected clinical development that were not identified in 3-month studies. Therefore, ICH S9 specifies that repeat dose studies of 3-month duration in rodents and nonrodents are usually adequate to support the initiation of phase 3 studies.
ICH S9 states that embryo-fetal development toxicity studies in a rodent and nonrodent species should be available in the marketing approval package of small molecules. Nevertheless, the document does provide exceptions for pharmaceuticals that are genotoxic and target rapidly dividing cells. The wording in this case is really referring to cytotoxic drugs. The EWG struggled with the use of the words “cytotoxic” or “cytostatic,” and could not come to a common understanding. Instead, it was decided to explicitly state the mechanism of action for drugs for which these studies would not be required. Another class of drugs where these studies would not be required is those in a class known to cause developmental toxicity. Of course, in both of these cases, the labeling for those drugs would state that there is an expectation of embryotoxicity, teratogenicity, or both. If a sponsor has a drug in those categories and would not want them labeled as teratogenic or fetotoxic, they would then have to run the appropriate preclinical studies and demonstrate that they are not. ICH S9 also states that if a drug causes significant embryo-fetal lethality or teratogenicity in an initial embryo-fetal developmental study, further study is usually not warranted. Although not explicitly stated, this would usually include such observations in pilot studies. Damage to reproductive organs is assessed during the general toxicity studies; therefore, separate fertility studies are usually not needed. Likewise peri- and postnatal toxicology studies are usually not needed to support clinical development in this patient population. Genotoxicity studies are usually not needed prior to entering phase 1 clinical trials but should be included with the marketing approval information package. Carcinogenicity studies are usually not needed to support marketing for anticancer drugs intended to treat patients with advanced cancer, as was specified previously in ICH S1A. Immunotoxicity is similar to safety pharmacology where assessment in general toxicology studies are sufficient to evaluate immunotoxic potential, but additional endpoints may be needed if the mechanism of action involves immunomodulation.

The guidance specifies that a separate safety evaluation of metabolites that have been identified in humans but have not been qualified in animal studies is usually not necessary but could be requested by the regulatory agency. It was generally agreed in the EWG that this would be rare because by the time significant metabolites are identified in patients who were not present in animals, there would be a considerable amount of clinical data supporting the continued use of the drug. EMA guideline on “The Limits of Genotoxic Impurities” (http://www.ema.europa.eu/pdfs/human/swp/519902en.pdf) discusses safety testing of genotoxic impurities in drug products based on potential increase in lifetime risk of cancer. This type of testing may not be appropriate for oncology therapies intended to treat patients with short life expectancy. Therefore, exceeding established limits for impurities identified in the EMA guideline and ICH Q3A/Q3B may be acceptable. The developer should provide justification for higher limits based on the indication, patient population, nature of parent drug (pharmacology, genotoxicity, etc), and duration of treatment. If the impurities are also metabolites in animal or human studies, they are considered qualified.

**ICH M3(R2): HIGHLIGHTS OF REVISIONS IN FUNDAMENTALS OF PRECLINICAL DRUG DEVELOPMENT**

ICH M3(R2) is a major revision of the original ICH M3, which was implemented in 1997. Extensive revisions were undertaken and new sections were added to address some areas in which harmonization had not been achieved in the original ICH M3 and to incorporate some new areas not previously addressed. For some of the topics, regional guidance had been developed in the interim. The ICH M3(R2) supersedes any regional guidance in areas in which they may differ. The major revisions are explained here.

**Limit Dose in Toxicity Studies**

A dose limit of 1000 mg/kg/day for rodents and nonrodents is considered appropriate for general toxicity studies if the human dose does not exceed 1 g/day (based on Organization for Economic Cooperation and Development guidelines) and the preclinical exposure exceeds the clinical exposure by 10-fold (usually based on group mean area under the curve [AUC]). Alternatively, doses providing a 50-fold margin of exposure relative to the clinical systemic exposure usually are also considered acceptable as the maximum dose for acute and repeated dose toxicity studies in any species. For the United States, if giving a dose that yields 50 times the maximum human exposure does not cause dose-limiting toxicity, then one study of at least 1 month may be recommended to support initiation of a phase 3 clinical study. The dose in this study should be 1000 mg/kg, the maximum feasible dose (MFD) or the MTD, whichever is lowest.

**Duration of Repeated Dose Studies in Nonrodents**

The ICH EWG evaluated a review of accumulated preclinical data (rodent and nonrodent) for about 150 compounds developed for diverse indications from European Union (EU) countries, the United States, and Japan from the period between 1999 and 2006. The working group then assessed 6-month, 9-month, and 12-month toxicology studies to determine the added benefit of studies longer than 6 months, i.e., do such studies discover toxicities not seen in the 6-month studies, and will any new toxicity observed in longer term studies affect clinical decisions. The overriding criterion was whether clinical decisions would have changed based on new toxicity uncovered in longer term studies. The EWG concluded that 6-month toxicology studies in nonrodents (primarily dogs) are usually but not always sufficient to support marketing.
approval. The available data did not show major differences between studies of 9- or 12-month duration; thus, 9-month nonrodent chronic studies should be adequate to support chronic use of small molecule drugs in humans without exception.

### Exploratory INDs

The most significant addition to ICH M3(R2) is a new section describing criteria for exploratory clinical studies (exploratory INDs in the United States). These are studies designed to help decision making for further clinical development of a new drug or biological therapy. The exploratory studies involve limited human exposure (determinations of MTD are allowed only in a few specific situations), have no diagnostic or therapeutic intent, and usually evaluate PK. The requirements for preclinical studies needed to support such limited clinical objectives are also less than those needed to support developmental clinical studies, the goal of which is to explore tolerability and toxicity during short-term use. The guideline describes five examples of exploratory clinical study approaches. The recommended preclinical study battery that supports each of these approaches is usually not adequate for further development. The five approaches include two microdose (≤ 100 μg) approaches, single-dose subtherapeutic studies, and two repeated dose exploratory studies. Both of the repeated dose studies allow dosing up to 14 days and dose escalation into the therapeutic range. In both cases, the ratio between the exposures determined at the NOAELs in the rodent and nonrodent studies and those observed clinically limits dose escalation. The approaches described are examples. Similar approaches may be proposed.

#### Microdose Approach No. 1

Microdose approach no. 1 may be useful to investigate target receptor binding or tissue distribution in a positron emission tomography (PET) study or to assess PK with or without the use of radiolabeled agents. In this approach, the total clinical dose is limited to 100 μg, which can be administered all at once or as divided doses without interdose interval limitations. The clinical dose is limited to 1/100th the NOAEL or 1/100th the pharmacologically active dose, scaled across species. This approach is supported by an extended single-dose toxicity study in one species by the intended route and PK. Generally, the extended single-dose toxicity studies should be designed to evaluate hematology, clinical chemistry, gross pathology, and histopathology data after a single administration, with further evaluations conducted 2 weeks later to assess delayed toxicity and recovery. The maximum dose in the preclinical study can be 1000-fold greater than clinical dose, scaled across species. The in vitro pharmacodynamic profile is needed, but genotoxicity testing is not needed.

Microdose Approach No. 2

Microdose approach no. 2 may be useful to investigate target receptor binding or tissue distribution in a PET study with less active PET ligands or to assess PK with or without the use of radiolabeled agents. In this approach, the total dose is limited to 500 μg, with up to five administrations of up to 100 micrograms per administration, and a washout between administrations. The clinical dose is no greater than 1/100th the NOAEL and 1/100th the pharmacologically active dose, scaled across species. This approach is supported by a 7-day repeat dose toxicity study in one species by intended route with PK data. The maximum dose in the preclinical study can be 1000-fold greater than clinical dose, scaled across species. The pharmacodynamic profile in vitro is needed, but genotoxicity testing is not.

#### Approach No. 3

**Single-dose subtherapeutic studies.** This approach can be used to determine clinical PK with nonradiolabeled drug at or near the predicted pharmacologically active dose but below the MTD. The clinical starting dose is based on the toxicity observed in the most sensitive species, the pharmacologically active dose, and regional guidance. Clinical dose escalation is limited to a dose that yields an exposure one-half that of the exposure at the NOAEL in the more sensitive species. Any relevant toxicity observed in animals should be anticipated to be monitorable and reversible in humans. Preclinical assessment should include the in vitro and in vivo characterization of the pharmacodynamic effect and the core battery of safety pharmacology studies. The supporting toxicity studies are extended single-dose studies in both a rodent and nonrodent by intended route with PK. The top dose should be the MTD, MFD, or the limit dose. An Ames assay or equivalent assay should be conducted to evaluate mutagenicity.

#### Approach No. 4

**Multiple dose subtherapeutic studies.** The clinical starting dose and maximum dose depend upon whether there is toxicity in one or both preclinical species tested. Here M3 refers to regional guidance for selecting the first dose in humans. If toxicity is not seen in either species or is seen only in one species, the clinical starting dose should be one that gives a predicted clinical AUC value (based on either interspecies PK modeling or milligrams per square meter conversion) that is approximately 1/50th of the AUC at the NOAEL from the species yielding the lower exposure. Standard 14-day toxicology studies in rodents and nonrodents are needed to support this approach. If dosing does not cause toxicity in either species, the maximum clinical dose should not exceed that which yields 1/10th the lower exposure in either species at the highest dose tested in the animals. When dosing causes toxicity in only one species, the maximum clinical dose should not be higher than the NOAEL in the species showing toxicity or that...
which yields an AUC one-half that determined at the highest
dose tested in the species not showing toxicity, whichever is
lower. With toxicity in both species, the maximum clinical
dose should be based on standard risk assessment approaches
and, in this specific case, the clinical MTD can be explored.

**Approach No. 5**

**Multiple dose subtherapeutic studies.** In this approach, the
clinical starting dose should not exceed 1/50th the NOAEL in
the more sensitive species on a milligrams per square meter
basis. Toxicity studies should include a standard 2-week
repeated dose toxicity studies in rodents (with justification of
the rodent as an appropriate species). The top dose in this study
should be the MTD, MFD, or limit dose. A confirmatory study
in nonrodents \( n = 3 \) at the anticipated NOAEL exposure in
rodent with duration of a minimum of 3 days and at least the
intended clinical study duration is recommended. Alternatively,
an escalating dose study in the nonrodent with duration of
a minimum of 3 days and at least the intended clinical study
duration at the anticipated NOAEL exposure in the rodent may
be conducted. The maximum exposure in humans should not
be higher than the AUC at the NOAEL in the nonrodent
species or 1/2 the AUC at the NOAEL in the rodent species,
whichever is lower. In the absence of adverse effects in the
clinical trial, escalation above this AUC can be appropriate if
the findings in the toxicity studies are anticipated to be
monitorable, reversible, and of low severity in humans.

**Reproduction Toxicity Studies**

This section considers the nature and timing of reproductive
toxicity studies to support the conduct of different phases of
clinical trials. To support this section, EWG reviewed data
from dose ranging and definitive studies in rats and rabbits for
several hundred drugs developed for diverse indications from
EU, the United States, and Japan during 1999–2006. The
working group tried to determine how well dose-ranging
studies predicted the results of definitive studies and how any
differences between these studies affect clinical decisions or
product labeling. They concluded that dose-ranging studies
with visceral/external examinations have good predictivity for
the outcome of definitive studies. Dose-ranging studies could
be used to support clinical studies before completion of
definitive reproductive toxicity studies if such studies are of
3 months duration or less and enroll only up to 150 women of
child-bearing potential (WOCBP) all of whom are using
precautions to prevent pregnancy. Nevertheless, the FDA
currently allows and will continue to allow the conduct of such
clinical trials without dose-ranging studies. In the EU and
Japan, definitive studies are usually required to support
inclusion of WOCBP in clinical studies, but the regulatory
agencies in these regions allow some exceptions. These include
short duration clinical trials (such as 2 weeks) with intensive
control of pregnancy risk. It was noted for monoclonal
antibodies for which embryo-fetal exposure during organogen-
esis is understood to be low in humans, the developmental
toxicity studies can be conducted during phase 3 (rather than
before phase 3) and the completed reports should be submitted
with the marketing application.

**CONCLUSION**

These new and newly revised guidelines are expected to
promote continued rational, scientifically designed approaches
to preclinical testing based on the best available methods and
technology to date while fostering careful study design and
minimizing use of animals. The recommendations in these
guidelines are drawn from the scientific and regulatory
information available at the time of its publication. A highlight
of the major changes discussed here is summarized in Table 2.
As these new guidelines or revisions are implemented,
feedback through either regulatory authorities or industry
groups will determine in what circumstances the guidance is
not efficient or effective; this will prompt consideration of
additional clarifications to the guidance. As with all guidelines,
modifications can be expected to be made in the future as
a result of advances in any of the biomedical disciplines
associated with pharmaceutical development.

**REFERENCES**

4682–4689.

Clarke, J., Hurst, C., Martin, P., Vahle, J., Ponce, R., Mounho, B., Heidel, S.,
toxicity studies for biotechnology-derived pharmaceuticals: is 6 months still

Goh, M. A., Lee-Ham, D. Y., McGuinn, W. D., Schmidt, W., Sun, C. J.,
et al. (1998). Regulatory considerations for preclinical development of

Greenblatt, M. S., Bennett, W., Hollstein, M., and Harris, C. C. (1994).
Mutations in the p53 tumor suppressor gene: clues to cancer etiology and


Jarvis, P., Srivastav, S., Vogel, E., Stewart, J., Mitchard, T., and
Weihbauer, G. F. (2010). The cynomolgus monkey as a model for
developmental toxicity studies: variability of pregnancy losses, statistical
power estimates, and group size considerations. _Birth Defects Res. B. Dev.

Mauther, R. J., Gibson, D. P., Bunch, R. T., and Custer, L. (2001). The Syrian
hamster embryo (SHE) cell transformation assay: review of the methods and

synthesis in hepatocytes isolated from rats treated with genotoxic agents: an
in vivo—in vitro assay for potential carcinogens and mutagens. _Carcinogen-
ecies_ 1, 621–625.

placental antibody transfer: new insights into developmental toxicity testing


