



Regulators Taking Steps to Add Fresh Perspectives to the ICH S6 (R1) Guidance

The International Conference for Harmonisation (ICH) S6 guidance document has been an indispensable reference for developers of biopharmaceuticals and is the starting point for the design of any safety study

for biotechnology-derived pharmaceuticals¹. The guidance has been in place since 1996 and has provided a framework for standardization of safety testing and a reduction in animal studies by harmonizing the expectations of regulatory groups in the United States, Europe, and Japan. The document was authored in an era when recombinant proteins and monoclonal antibodies were the almost exclusive types of biopharmaceuticals submitted to regulatory authorities. Since then, a variety of novel biologics have been engineered, including pegylated proteins, oligonucleotides, bi-functional peptides, and spiegelmers, to name only a few. Over time, the application of this novel biotechnology gave rise to regulatory and development questions that were either not addressed in the original ICH S6 guidance or were addressed in language too general, which resulted in different understandings or interpretations among Sponsors and regulatory groups.

As a remedy, in 2008 the International Conference for Harmonisation convened an Expert Working Group (EWG), composed of drug regulatory authorities and pharmaceutical experts, to clarify the ICH S6 language. The topics to be addressed included species selection, study design, reproductive and developmental toxicity, immunogenicity, and carcinogenicity. The draft recommendations made by the EWG were released to the public as a Step 3 ICH S6 (R1) (formal ICH process) Addendum in November 2009 and were open for public comment until February 2010². Comments were received from industry scientific associations, contract laboratories, consultants, pharmaceutical companies, ICH regulatory authorities (FDA, CBER, CDER, EMEA), and regulatory authorities in Canada, China, Korea, and Singapore. The document will become a finalized or Step 4 guidance in the ICH process after the EWG incorporates changes or revisions to the current Addendum. The EWG was scheduled to make its final recommendation in November 2010, but finalization is pending as the EWG debates remaining differences of opinion among its members.

Overall the draft Addendum is a pertinent update that was necessary to incorporate technical advances that have taken place in the biologics field since mid-1990. Below, some major points are reviewed in each of the five Addendum topics and highlight the corresponding public comments that were submitted as part of the ICH review process. Principal sources for this article are personal experience and public comments submitted to the EWG by the Biotechnology Industry Organization³.

Species Selection

A key principle in the original ICH S6 guidance is the emphasis on the design of toxicology studies and the use of a relevant animal model, that is, a species that is pharmacologically responsive to the intended human drug. However, because biopharmaceuticals have unique attributes and mechanisms of action, choosing the most appropriate animal model for toxicology testing required a case-by-case approach. This was a departure from the then routine practice with small molecule drug candidates of testing primarily in albino rats and beagle dogs. Driving the case-by-case scenario was the desire to avoid or minimize the anticipated immunogenicity to

administered human test proteins, especially in chronic toxicology studies. Another consideration was the fact that most monoclonal antibodies, designed to be highly specific for human antigens, would cross-react with the target only in a nonhuman primate (NHP). When the NHP was the only relevant pharmacologically reactive species, the requirement for toxicity testing in two species was generally waived, which also then became accepted practice for other non-monoclonal antibody biopharmaceuticals (the exception is endocrine peptides, which are typically tested in rat and dog species). Moreover, avoidance of immunogenicity became the *raison d'être* for the NHP to be the preferred animal model, because there is greater sequence homology (less immunogenicity) of human biologics in these higher animal species. An unintended consequence of this practice was an increase in demand for monkeys, in parallel with the rising numbers of biopharmaceuticals requiring toxicity studies.

In time, experience in a number of pharmaceutical companies revealed that many of the more recently engineered biopharmaceuticals would cross-react in other species, such as the rat, without inducing immunogenicity. This was also true for some human monoclonal antibody therapeutics. Because the ICH S6 guidance states that two species are preferred and a second pharmacologically relevant species was available, the circumstances produced indecision among Sponsors who were not convinced that testing in a lower animal species would inform human safety risks. This also created a regulatory climate where Sponsors were challenged to demonstrate that a rodent model was not relevant or feasible. Usually this meant conducting a rodent 30-day toxicity study to address the question, effectively creating a two-species-testing paradigm by default. If it turned out that the rat was not pharmacologically appropriate, the results were still submitted to regulatory authorities as demonstration of off-target effects—a practice discouraged in the original ICH S6 guidance.

Public comments received by the EWG on this issue agree with the consensus opinion in the draft Addendum that a single species may be appropriate, given adequate scientific justification. The example is given of developmental and reproductive toxicology (DART) studies where a single species is acceptable (see topic below). Respondents have asked the EWG for criteria that would define when it is appropriate to go forward in development with only one species. Intertwined with this clarification, however, is the question of how much the determination of drug/target binding information and species sequence homology to target factor into the decision, and whether these data should be included as criteria for species selection. Current language in the Step 3 Addendum appears to eliminate tissue cross-reactivity analyses as a criterion, a conclusion that has been challenged by some scientists via the public comments. Characterization of target tissue binding, or tissue cross-reactivity, in the pharmacology animal model is frequently done during the early development or discovery phase to verify that a candidate molecule is hitting the appropriate target receptor. These experiments are not necessarily conducted in the eventual toxicology species, but often they are. Some Sponsors leveraged this information in their rationale for selecting a given species for toxicity testing, and it seems appropriate to continue the practice as a useful tool to support the pharmacology of the molecule. However, the results should not be the sole reason for justifying the selection of the toxicity species. It was also recommended in the original ICH S6 guidance that the reactivity of candidate monoclonal antibody drugs with human tissues and tissues from the preclinical test species be characterized. Public

comments favorable to this guidance point out that the tissue cross-reactivity data can provide important information to supplement knowledge of target distribution and reveal potential unexpected binding when evaluated in the context of the overall pharmacology and safety assessment data. Respondents agreed that tissue cross-reactivity methods should not be used for assessing molecule-to-molecule comparability, or for assessment of each binding site on bi-specific antibodies and protein drugs. Implicit in the positive support for retaining tissue cross-reactivity is that those laboratories conducting the assays have used the proper scientific rigor to develop the methods and are themselves proficient in the art.

Study Design

The Step 3 Addendum advises that a rationale be provided for high dose selection in toxicity studies and specifies that it should be either the dose that produces the maximum pharmacological effect in the preclinical species, or a dose that gives a 10-fold exposure margin over the maximum exposure to be achieved in the clinic. However, public comments request consideration for cases when the scientific data support a lower dose, as when solubility or overt pharmacology limits the top dose that feasibly can be administered. This could also apply to cases where the population has limited tolerability (e.g., nausea) to the drug or in high-risk oncology indications.

Further clarification is necessary around the issue of dose adjustment in preclinical studies if there is a large difference in binding affinity or potency between the human target and the target in animals. Regulators have in the past asked for Sponsors to adjust upwardly the high dose in preclinical studies proportional to the difference in reported affinities. Public comments point out that experiments have shown that *in vitro* potency or binding affinity with the target does not necessarily directly translate to observed *in vivo* biology or pharmacology. Because affinity relationships are a large unknown, it may be the best compromise to default to the 10-fold margin rule.

On the point of study duration for repeat dose toxicity studies, the Step 3 Addendum states that six-month chronic studies have proved adequate over time for preclinical safety testing. However, there is apparently still some skepticism among regulatory authorities who believe that longer studies may reveal potential hazards that would not otherwise be manifested. Two examples that are often cited are results obtained from longer duration studies conducted with insulin aspart (tumor promotion) and adalimumab (immune complex formation). The counterargument is that a weight of evidence approach would also identify insulins as potential promoters as a class, and the risk of immune complex formation has previously been identified for most monoclonal antibodies. Hence addressing these in the product label as potential hazards would seem to be appropriate and serve to inform patients of the safety risk, even if not identified in six-month studies. A retrospective review by Clarke and colleagues⁴ of preclinical studies conducted with 23 biologics supports this conclusion and offers the example that proportionally fewer biopharmaceuticals than small molecular weight drugs have been withdrawn from Phase II or Phase III clinical trials, an indication that no significant new hazards are surfacing in patient populations receiving the test drugs.

Immunogenicity

The original ICH S6 guidance recommends that Sponsors use a sensitive assay (able to detect <1 ug/ml of anti-drug antibody (ADA)) to characterize immunogenicity at pretreatment, end of study, and/or recovery, and to report the incidence and titers of individual animals

with a positive ADA response. These findings were to be evaluated in context with adverse toxicology findings, including evidence of hypersensitivity reactions or immune complex disease.

Over time, emphasis has shifted from concerns for the immunotoxic effects of ADA to the potential impact of ADA on the pharmacokinetics and pharmacodynamics of the test material. If ADA is detected preclinically, does it abrogate or diminish the pharmacologic activity of the biopharmaceutical and potentially invalidate the toxicity study? This point is highlighted in the Step 3 Addendum and further states that immunogenicity characterization is not required if there is sustained pharmacodynamic activity at study termination, as evidenced by the detection of a reliable marker of pharmacologic effect. If such a marker is unavailable, the Sponsor is obligated to fully characterize the ADA response per the ICH S6 and the EMEA draft immunogenicity guidance⁵ and published white papers^{6,7}. This topic received the fewest public comments, with most agreeing with the language and recommendations for immunogenicity testing.

Developmental and Reproductive Toxicology Testing (DART)

Although the conduct of reproductive toxicology studies is governed by the ICH S5 (R2) guidance, the ICH S6 Step 3 Addendum provides detailed instruction on how to conduct DART studies for those biopharmaceuticals where nonhuman primates are the only pharmacologically relevant species. The recommendations also reflect the intent to reduce the overall use of NHPs for reproductive safety testing. Illustrative of this point is the statement that a single, pharmacologically relevant species is sufficient to evaluate the teratogenic potential (embryo/fetal) of a test molecule. Furthermore, if there is more than one (NHP) relevant species, such as a rat or rabbit, the evaluation should be conducted in one of these lower animal species. If a Sponsor still wants to use NHP in a reproductive toxicity study, a scientific rationale must be provided for selection of monkeys. It is also regulatory preference that when the candidate biopharmaceutical is active only in NHP, the default is to conduct reproductive toxicology in this species rather than employing alternative approaches with homologous proteins or surrogate molecules.

If NHP is the only relevant model, language in the Addendum allows Sponsors to conduct DART studies with a control group and only one dose group (preferably a dose that provides a 10-fold exposure margin). This agrees with public comments that an NHP DART study is essentially a hazard identification study, given the low numbers of females conventionally assigned to treatment groups. The Addendum further acknowledges that placental transport of many large molecular weight biologics does not occur via passive diffusion but may enter through specific transport mechanisms that may be temporally or differentially active only at certain stages of gestation. This is especially pertinent to immunoglobulins and monoclonal antibody therapeutics that are transported preferentially via neonatal Fc receptors (FcRn) and primarily during the last trimester, similar to humans. Therefore the recommendation is that maternal drug exposure in NHP DART studies should be maintained from gestation day 20 to birth, rather than only during organogenesis (day 50), as done with small molecule DART studies. The reader is referred to the publication authored by Martin and Weinbauer⁸, who expertly review appropriate NHP DART study designs that are based on the known developmental biology of primates and are consistent with recommendations in the Step 3 Addendum.

Carcinogenicity

For novel small molecular drug candidates, there is firm regulatory expectation that two-year rodent carcinogenicity bioassays will be conducted. This thinking changed with the introduction of human biologics that were either immunogenic in rodents or not pharmacologically active, and led to the language in the original ICH S6 guidance stating that standard carcinogenicity bioassays are generally inappropriate for biotechnology-derived pharmaceuticals. However, the guidance left the door open to consideration of rodent bioassays under certain situations. This door has been jarred further open with the demonstration, discussed earlier, that for some biopharmaceuticals, immunogenicity is not an issue, and the agents are also active in rodents. In perhaps a combination of a sense of fair play and concern for safety, regulators have asked Sponsors to conduct bioassay studies, thereby overriding a decade-old practice in some companies of not conducting carcinogenicity studies for biologics. The situation also has created anxiety during the planning and development phase for new molecules, as Sponsors attempt to anticipate whether regulatory authorities will require a bioassay for their product.

At this writing, the EWG has neither published its response to public comments nor given its approval to this topic in the Step 3 Addendum. Public comments agree with concepts in the Addendum that a weight of evidence approach should be considered when deciding if a two-year bioassay is needed. This would include a formal evaluation of the scientific literature pertinent to the drug or drug class, with the purpose of understanding any real or theoretical risks. Consideration would also be given to the utility of alternative methods for evaluating drug-induced mitogenicity or cell proliferation or the use of genetically altered rodent models. For example, if the literature indicates that the mechanism of action of a candidate biopharmaceutical renders it a human tissue mitogen, or a promoter, then the liability has been identified, and it is unlikely that a two-year assay will further inform the human risk. In other words, if animal data clearly demonstrate a cancer risk, further animal studies are not warranted. Many biopharmaceuticals do not target DNA (non-genotoxic), are non-immunosuppressive, and do not induce cell proliferation (non-promoter) in chronic six-month studies in one or two species. Additionally, the majority of biologic drugs are highly target-specific, do not generate reactive metabolites, and produce minimal to no off-target effects. Post-marketing surveillance is believed to be the best means of characterizing or assessing carcinogenic risk with this class of drugs. The Addendum gives the example of using this approach for immunomodulators where it is known that significant immunosuppression in humans increases the risk of developing lymphoid neoplasms. This risk, as advised in the Addendum, should then be communicated in the product label.

There may be cases where the science and a weight of evidence evaluation are insufficient for making an assessment about carcinogenic potential and will require additional information. But rather than default automatically to conducting a rodent bioassay, Sponsors should ask whether there is a specific concern or attribute of the pharmacology or toxicology that could be studied with further experimentation based on a specific hypothesis. The results from the additional experiments, combined with information from a weight of evidence evaluation, may provide sufficient information to make an informed risk assessment of the molecule's carcinogenic potential. The public comments are in favor of this approach for carcinogenicity,

but implementation by the industry remains in limbo until further clarification is forthcoming, or approval is granted by the EWG to accept the current Step 3 Addendum language on this topic.

In summary, the clarifications and recommendations contained in the Step 3 Addendum address many of the major regulatory and scientific questions that the biotechnology industry wrestles with as it attempts to balance animal welfare, economy, and patient safety in the conduct of preclinical toxicity testing. It is expected that the final Step 4 guidance will not differ substantially from the current version, but the public keenly awaits a ruling and finalization of the ICH S6 Addendum.

References

- 1 Food and Drug Administration. (CDER/CBER/ICH. July 1997). Guidance for Industry S6 Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals. Retrieved February 9, 2011 from <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm074957.pdf>
- 2 European Medicines Agency. (CHMP/ICH/302/95. November 2009). ICH Topic S6 (R1). Step 3. Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals. Retrieved February 9, 2011 from http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/01/WC500043470.pdf
- 3 BIO comments on FDA-2009-D-0573. Addendum to ICH S6: Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals S6 (R1). Retrieved on February 9, 2011 from <http://www.bio.org/reg/20100201.pdf>
- 4 Clarke, J., Hurst, C., Martin, P., Vahle, J., Ponce, R., Mounho, B., Heidel, S., Andrews, L., Reynold, T. and Cavagnaro, J. (2007). Duration of chronic toxicity studies for biotechnology-derived pharmaceuticals: is 6 months still appropriate? *Regul. Toxicol. Pharmacol.* 50: 2-22.
- 5 European Medicines Agency. (EMA/CHMP/BMWP/14327/2006 (DRAFT)). Guideline on Immunogenicity Assessment of Biotechnology-Derived Therapeutic Proteins. Retrieved February 9, 2011 from http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003947.pdf
- 6 Mire-Sluis, A.R., Barrett, Y.C., Devanarayan, V., Koren, E., Liu, H., Maia, M., Parish, T., Scott, G., Shankar, G., Shores, E., Swanson, S.J., Taniguchi, G., Wierda, D. and Zuckerman, L. A. (2004). Recommendations for the design and optimization of immunoassays used in the detection of host antibodies against biotechnology products. *J. Immunol. Meth.* 289: 1-16.
- 7 Koren E, Smith HW, Shores E, Shankar G, Finco-Kent D, Rup B, Barrett YC, Devanarayan V, Gorovits B, Gupta S, Parish T, Quarmby V, Moxness M, Swanson SJ, Taniguchi G, Zuckerman LA, Stebbins CC, Mire-Sluis A. (2008). Recommendations on risk-based strategies for detection and characterization of antibodies against biotechnology products. *J. Immunol. Meth.* 333:1-9.
- 8 Martin, P.L. and Weinbauer, G.F. (2010). Developmental toxicity testing of biopharmaceuticals in nonhuman primates: previous experience and future directions. *Intl. J. Toxicol.* published online 6 October 2010. doi: 10.1177/1091581810378896.



Daniel Wierda, MS, PhD, is a senior toxicologist and immunology safety consultant for B2S Consulting, a consultancy firm that specializes in preclinical and clinical statistical and bioanalytical support for diabetes.

Dr. Wierda has over 30 years of experience in academia and industry spanning the areas of myelotoxicity, immunotoxicity, immunogenicity, and biopharmaceutical preclinical drug development. After receiving his MS and PhD degrees in pharmacology and toxicology from the University of Kansas Medical School in Kansas City, Kansas, he conducted postgraduate research under the tutelage of Dr. Richard Irons in the area of benzene hematotoxicity at the Chemical Industry Institute for Toxicology, Research Triangle Park, North Carolina. He continued his research in benzene hematotoxicity as a faculty member in the Department of Pharmacology and Toxicology at the University of West Virginia, Morgantown. He was elected president of the Society of Toxicology Immunotoxicology Specialty section in 1994. Dr. Wierda has authored several manuscripts related to preclinical immunogenicity and has been an invited speaker on the topic of the use of preclinical animal models for characterizing the relative immunogenicity of biopharmaceuticals.