Reproductive Toxicology 64 (2016) 57–63

Contents lists available at ScienceDirect

Reproductive Toxicology

journal homepage: www.elsevier.com/locate/reprotox

Revision of the ICH guideline on detection of toxicity to reproduction for medicinal products: SWOT analysis

Paul Barrow*

Roche Pharmaceutical Research and Early Development, F. Hoffmann-La-Roche, Ltd., Basel, Switzerland

ARTICLE INFO

Article history:
Received 2 March 2016
Received in revised form 29 March 2016
Accepted 31 March 2016
Available online 1 April 2016

Keywords:
Reproductive and developmental safety testing
Guidelines
Teratogenicity
ICHS5

ABSTRACT

SWOT analysis was used to gain insights and perspectives into the revision of the ICH S5(R2) guideline on detection of toxicity to reproduction for medicinal products. The current ICH guideline was rapidly adopted worldwide and has an excellent safety record for more than 20 years. The revised guideline should aim to further improve reproductive and developmental (DART) safety testing for new drugs. Alternative methods to animal experiments should be used whenever possible. Modern technology should be used to obtain high quality data from fewer animals. Additions to the guideline should include considerations on the following: limit dose setting, maternal toxicity, biopharmaceuticals, vaccines, testing strategies by indication, developmental immunotoxicity, and male-mediated developmental toxicity. Emerging issues, such as epigenetics and the microbiome, will most likely pose challenges to DART testing in the future. It is hoped that the new guideline will be adopted even outside the ICH regions.

1. Introduction

This article is intended to accompany an invited talk at the 44th annual meeting of the European Teratology Society, the views expressed are those of the author alone and do not represent the policies, positions or opinions of any organization, group or company.

The ICH guideline on "Detection of toxicity to reproduction for medicinal products" was the first ICH safety guideline. Following its issue in 1993 [1], it rapidly gained worldwide acceptance (see below for the notable exceptions of China and India). This guideline successfully harmonized the disparate requirements of the regulatory authorities in Europe, the USA and Japan. A significant reduction in experimental animal use was thus accomplished by abolishing the need to duplicate the various non-clinical reproductive toxicity studies to achieve marketing authorization for a new drug across the three regions. This remarkable achievement marked the birth of the International Conference on Harmonisation (recently renamed International Council on Harmonisation). The principles of the ICH guideline were based on a previous guideline issued by the FDA in 1966 [2] in response to the thalidomide tragedy and subsequent guidelines from the EC and Japan.

The initial ICH S5 guideline left open some questions on the minimum duration of treatment of males before mating and the relative value of semen analysis, mating performance and histopathology for the evaluation of testicular toxicity. Subsequent literature surveys [3] and validation studies [4,5] concluded that histopathology of reproductive organs is the most sensitive method for detecting effects on spermatogenesis. An addendum to the guideline to incorporate these findings was issued in 2000. The ICH S5(R2) nomenclature was added in 2005, when the title of the guideline was changed to "Detection of Toxicity to Reproduction for Medicinal Products & Toxicity to Male Fertility".

The studies described in the guideline are designed to detect each of the four known manifestations of developmental toxicity, i.e.: (1) death (embryo–fetal resorption, abortion, stillbirth or post-natal mortality), (2) growth retardation (resulting in low birth weight or depressed post-natal growth), (3) malformation, and (4) functional deficit [6]. The guidelines define six phases of reproduction that need to be assessed: (A) adult fertility, (B) early embryonic development before implantation on the uterus, (C) embryonic organogenesis, (D) fetal development, (E) birth and pre-weaning development and (F) post-weaning development up to sexual maturity. A three segment strategy is proposed to cover the evaluation of all of these phases, comprised of a fertility study (usually in the rat), embryo-fetal development (EFD) studies in two species (usually rat and rabbit) and pre- and post-natal development (PPND) studies (usually in the rat). Various options are also

* Correspondence address: Roche Pharmaceutical Research and Early Development, Pharmaceutical Sciences, Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd., Grenzacherstrasse 124, CH 4070 Basel, Switzerland.
E-mail address: paul.barrow.pb1@roche.com

http://dx.doi.org/10.1016/j.reprotox.2016.03.048
0890-6238/

proposed for combining two or more of the various rodent studies into a single experiment.

More than 20 years after taking effect, ICH S5 guideline is about to undergo its first major revision [7]. This review uses a SWOT (strengths, weaknesses, opportunities and threats) analysis to identify and discuss salient points to be considered in the ICH S5 guideline revision.

2. Methods

The strengths and weaknesses of the current ICH S5(R2) guideline were assessed, along with opportunities and threats pertaining to the revision process. Each item is listed by category and discussed with regard to the future revised guideline (R3). Identifying the strengths of the guideline highlights the current existing elements that should not be removed or disrupted in the revision process. The identified weaknesses highlight issues that have arisen since the adoption of the guideline or where scientific thinking or technology has evolved. Opportunities represent desirable additions to the guideline, such as alignments with other ICH guidance documents. Finally, threats are issues that should be dealt with proactively where possible in case they become obstacles to the successful completion of the revision process.

3. Results

3.1. Strengths of the current guideline

3.1.1. Safety record

The existing procedures described in the guideline have been remarkably successful in identifying the reproductive hazards of new drugs. Since the adoption of the guideline, there have been no developmental toxicity-related tragedies with marketed drugs. Effective drug labeling and more cautious drug prescription practices have of course also contributed to this success. The study designs described in the guideline have proven to be effective in detecting reproductive and developmental hazards associated with mechanisms of action that were not yet envisaged when the guideline was devised. This, however, is not a reason to be complacent (see Weaknesses).

3.1.2. Wide acceptance

ICH S5(R2) rapidly gained acceptance by most health authorities worldwide, including those outside of the ICH regions (with the exception of China and India, see Threats) and has been the *de facto* standard for more than 20 years.

3.1.3. 3 Rs

From its adoption, ICH S5 resulted in a significant reduction in the number of experimental animals used for the regulatory developmental and reproductive toxicity (DART) testing of new drugs. By harmonizing study designs, the guideline removed the need for specific DART studies to meet the regulatory requirements of each region for a worldwide marketing submission. We must continue to strive to reduce or eliminate animal use for drug safety testing, without compromising safety.

3.1.4. Established robust study designs

Despite their complexity, the DART study designs have proven to be practical and effective, having incorporated the best elements from the previous guidelines in the three ICH regions. The necessary equipment is readily available and the methods have become routine. Safety testing laboratories have built up large databases of reference values over the last two decades.

3.1.5. Flexibility

The guideline avoids mandatory rules, favoring flexibility (for instance, with respect to the combination of various rodent studies). This flexibility has allowed the guideline to remain relevant even for classes of drug that had not been invented when the guideline was published. This flexibility should be retained in the revised guideline.

3.1.6. Testicular toxicity

The research culminating in the amendment on testicular toxicity remains pertinent today and even anticipated issues that would arise years later with respect to the testing of biopharmaceuticals (see Weaknesses, below).

3.2. Weaknesses of the current guideline

3.2.1. The title

"Detection of Toxicity to Reproduction for Medicinal Products & Toxicity to Male Fertility" is long, clumsy and misleading. Male fertility is part of toxicity to reproduction and does not need to be repeated. Furthermore, female fertility is not mentioned, even though it is covered in the guideline. It is ironic that male fertility was added to the title in acknowledgment of work that concluded that the most sensitive marker of testicular toxicity can be determined in the general toxicity studies rather than in the DART studies. Hopefully, the revised guideline will have a more balanced title.

3.2.2. No provision for alternative tests

Some alternative tests—i.e. in-vitro, ex-vivo, in-silico and non-mammalian systems are mentioned in the current guideline "for encouragement", but no options are provided to replace studies in live mammals. While very few new alternative tests for DART have been developed over the last 20 years, many of the previously existing tests have been extensively validated and qualified [8]. Unfortunately, a predictability of more than 80% with respect to animal studies has not been reliably reported for any of the available tests [9]. Also, the physical characteristics of drug candidates (solubility, pH, osmolality, etc.) often render alternative test systems impractical. Because of these limitations, alternative test systems cannot be expected to completely replace live animal studies for regulatory DART testing in the near future. Nonetheless, alternative systems, such as the embryonic stem cell test [10] and the zebra fish teratogenicity test [11], are used by many companies for drug candidate selection and are thus contributing to reduced animal use in pharmaceutical development.

The revised guideline should give guidance on how to qualify an alternative test system and the conditions that should be met before an alternative method can be used in the place of a mammalian study. The (fortunate) paucity of human data for the majority of known teratogens, makes it very difficult to demonstrate that a new alternative test is more or as effective than an established animal study for the detection of human teratogenicity. One current school of thought considers that a prospective alternative test should be demonstrated to be predictive of the animal test that it is intended to replace. So, for instance, before a zebra fish test can be used in the place of the rabbit EFD, it would have to be shown to reliably detect the same list of known teratogens as the rabbit EFD. This principle is illogical at best, however, when applied to a human stem cell test, which logically could be expected to be more predictive for the human than an animal-based test.

Alternative tests for teratogenicity in general show poor specificity (i.e. true negative rate) due to a very poor capacity to predict developmental effects arising from maternal influences (e.g. reduced perfusion of the placenta) or as the result of active metabolites produced by maternal metabolism (e.g. allyl alcohol