Regulatory Toxicology and Pharmacology 68 (2014) 108–118

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

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

Reassessing the two-year rodent carcinogenicity bioassay: A review of the applicability to human risk and current perspectives

Palma Ann Marone [a,*], William C. Hall [b], A. Wallace Hayes [c,1]

[a] Product Safety Labs, 2394 Highway 130, Dayton, NJ 08810, USA
[b] Hall Consulting, Inc., 110 Shady Brook Circle #300, St. Simons Island, GA 31522, USA
[c] Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115, USA

ARTICLE INFO

Article history:
Received 19 October 2013
Available online 25 November 2013

Keywords:
2-Year carcinogenicity
Genotoxicity
Human risk assessment
Prediction

ABSTRACT

The 2-year rodent carcinogenicity test has been the regulatory standard for the prediction of human outcomes for exposure to industrial and agro-chemicals, food additives, pharmaceuticals and environmental pollutants for over 50 years. The extensive experience and data accumulated over that time has spurred a vigorous debate and assessment, particularly over the last 10 years, of the usefulness of this test in terms of cost and time for the information obtained. With renewed interest in the United States and globally, plus new regulations in the European Union, to reduce, refine and replace sentinel animals, this review offers the recommendation that reliance on information obtained from detailed shorter-term, 6 months rodent studies, combined with genotoxicity and chemical mode of action can realize effective prediction of human carcinogenicity instead of the classical two year rodent bioassay. The aim of carcinogenicity studies should not be on the length of time, and by obligation, number of animals expended but on the combined systemic pathophysiologic influence of a suspected chemical in determining disease. This perspective is in coordination with progressive regulatory standards and goals globally to utilize effectively resources of animal usage, time and cost for the goal of human disease predictability.

1. Introduction[2]

It has been over 40 years since the signing of the US National Cancer Act in December, 1971. Great strides have been made in realizing cures, but elimination of the disease and its predictability remain elusive goals. As mortality from curable infectious and cardiac disease decreases, the developing world is expected to see rises in cancer rates from 12.7 to 21 million new cases/year from 2008–2030 (Ferlay et al., 2010). Therefore as more people are exposed to chemicals in the form of agricultural, industrial and pharmaceutical products, food additives and natural/environmental pollutants, risk assessment remains as important today as ever. On the assumption that the same cancers that arise in humans can also be applicable to rodents on an accelerated timescale, human hazard identification screening assays for the more than 200 types of known human cancers oftentimes employ the use of rodents to study mechanisms of disease if and when the dose can be extrapolated to the human condition (Cohen, 2004). This has led to the universal acceptance of the 2-year rodent bioassay as a model for which governmental regulatory agencies have developed standardized guidelines for use over the years. With the benefit of decades of available data and comments for review, and with current guidelines carefully reconsidering animal use and replacement in hazard assessment, a need to revisit the usefulness and applicability of this study continues.

1.1. History

The association between human cancer and chemicals has been known from epidemiologic studies of occupational exposure since the time of the industrial revolution (Infante, 1993). Historical evidence for individual agents of chronic irritation causing pre-cancerous lesions is well known (IARC, 1987; NIOSH, 1983; Yamagiwa and Ichikawa, 1918). The cumulative review of the results from long term animal studies concludes a high concordance between response in humans and that of other mammalian species as predictive evidence of carcinogenicity (Chu et al., 1981; Rall, 2000; Tomatis and Huff, 2002; Tomatis et al., 1989). Over time,

* Corresponding author. Fax: +1 (732) 230 4209.
E-mail addresses: PamMarone@ProductSafetyLabs.com (P.A. Marone), hallconsulting@earthlink.net (W.C. Hall), awallacehayes@comcast.net (A.W. Hayes).
[1] Fax: +1 (978) 409 1154.

[2] *Abbreviations:* EFSA, European Food Safety Authority; EMA, European Medicines Agency; EPA, Environmental Protection Agency; ESTP, European Society of Toxicologic Pathology; EU, European Union; FDA, Food and Drug Administration; IACUC, Institutional Animal Care and Use Committee; IARC, International Agency for Research on Cancer; ICH, International Conference on Harmonization; LH, Luteinizing Hormone; NCI, National Cancer Institute; NIEHS, National Institute of Environmental Health Sciences; NIOSH, National Institute of Occupational Safety and Health; NTP, National Toxicology Program; OECD, Office of Economic Cooperative Development; PPAR, Peroxisome Proliferator-Activated Receptor; QSAR, Qualitative and quantitative structure activity relationship; REACH, Registration, Evaluation, Authorization and Restriction of Chemicals; STP, Society of Toxicologic Pathology.

0273-2300/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.yrtph.2013.11.011

as more compounds were tested, sensitivity improved, and additions to the listing of potential carcinogens prompted the development of detailed classification schemes to confirm and validate the general usefulness of these studies (U.S. EPA, 1986). The advent of molecular biology and the confirmation of viral oncogenes in the 1960's, advanced chemical safety testing making the consideration of species and strain differences of animal and human data significant (Javier and Butel, 2008).

The current 2-year design was adapted from the original FDA systemic carcinogenicity protocols for food and drugs (Lehman et al., 1949; Lehman et al. 1955 as referenced in Jacobs and Hatfield, 2013; Weisburger, 1981; Weisburger and Weisberger, 1967; Weisburger and Williams, 1981a; Williams and Weisburger, 1981) and since published and revised in European and international guidelines (U.S. EPA, 1986, 1998a,b, 2005a,b,c; OECD, 1983, 2009a,b,c), the US FDA Redbook (FDA, 2000a), the US EPA guidelines (http://www.epa.gov/ocspp/pubs/frs/home/testmeth.html), International Conference on Harmonization (ICH) guidelines (http://www.ich.org/products/guidelines/safety/article/safety-guidelines.html) and the European Medicines Agency guidelines (ICH, 1996, 1998, 2008a, b, 2011, 2012a,b,c; Jena et al., 2005). Since 1971, the International Agency for Research on Cancer (IARC, 1980, 1999, 2006) has evaluated over 900 chemicals for carcinogenicity (Cogliano et al., 2004; IARC monographs 1–107, 2013). Other agencies in the United States (EPA, NIEHS) and independent groups within the international community also have evaluated a number of chemicals for their carcinogenicity potential (Gold et al., 1991, 1999; Huff, 2002; Ward, 2007). In addition, comprehensive pharmacologic/toxicologic records exist for over 5000 chemicals (Wexler et al., 2011).

Concurrently, *in vivo* and *in vitro* techniques designed to provide rapid, inexpensive screens for genotoxicity including mutagenicity (*Salmonella typhimurium* Reverse Mutation Test, Ames Test; Ames et al., 1973; McCann et al., 1975) and cytotoxicity (Micronucleus Test; Countryman and Heddle, 1976; Heddle, 1973; Schmid, 1975) were developed to study possible mutations arising from chemical exposure (Weisburger, 1999; Weisburger and Williams, 2011). Since that time, continued refinement of regulatory guidelines to strengthen the 2-year bioassay for environmental and industrial chemicals as the conclusive primary criteria for safety have evolved. Two-year studies for pharmaceuticals were adapted as an adjunct to the standardized chemical testing, but the pharmaceutical industry continues to rely most importantly on human data. Despite the vigorous debate for the relevance of risk evaluation in humans, long-term studies assessing the possibility of cancer for various compounds available to date indicate little fundamental change in the number of published reports and abstracts through the decades (187 in 1976–1979; 151 in 1980–1989; 126 in 1990–1999; 90 in 2010–2012, to date; NTP, 2006, 2013). Evaluation of the current practice of carcinogenicity testing (Combes et al., 2002; Locke and Goldberg, 2006) corresponds to the broader issue of animal use in scientific study at a time when the safety testing of advanced technological goods (genetically modified products, for example) does not lend themselves easily to standard study designs. More recently, guidance for the safety of biopharmaceuticals such as oligonucleotides, siRNA recombinant peptides, monoclonal antibodies and other like molecules according to ICH S6 guidance should be conducted on a "case by case" basis according to their functional risk in Japan and the US (Nakazawa et al., 2008; ICH, 2011, 2012a; Vargas et al., 2013).

2. Methods: study design

With some adaptations for the requirements in pharmaceuticals to include tumorigenic action, fundamentally there is little distinction between the rules and regulations developed for the safety studies required for chemicals, whether they be environmental, agricultural, industrial, food or medications (ICH, 1996). Generally, testing requirements for chemicals are standardized according to registration category whereas pharmaceuticals, particularly small biologic molecules, are evaluated on a case-by-case basis as it is not always appropriate to conduct traditional 2-year carcinogenic studies (Jacobs and Hatfield, 2013;Van Oosterhout et al, 1997). As a reflection of this, with a few modifications, guidelines for preclinical studies, short and long-term, are well standardized across world-wide regulatory agencies (Hayes et al., 2011; Robens et al., 1994) and, for pharmaceuticals, may be used to support clinical trials. Depending on the information available, and apart from pharmaceutical consideration, prior to performing the 2-year study, at least one and possibly two, prechronic studies in the form of a 14- and 90- days study (EFSA, 2011b; U.S. EPA, 1998c, 2000; FDA, 2000b, 2000c; ICH, 2008a; OECD, 1998, 2008) is often warranted for setting dose levels at the maximum tolerated dose (Gaylor et al., 1985; Haseman and Lockhart, 1994; Huff et al., 1986; NRC, 1993) The objective of the carcinogenicity study is to provide information on the health hazards likely to arise over the course of the lifetime of the given species, with the rodent serving as the current preferred model (Bucher, 2002; Portier and Hoel, 1983). With the goal of determining the potential for and characterization of tumor development and progression, presently the assessment of risk for long-term preclinical studies encompasses detailed directives on the housing conditions, record keeping, species, strain, the number and sex per dose group, dose level, duration, regimen and route of exposure, applicability to and availability of genotoxicity, pharmacokinetics, and clinical and histopathology process and assessment (ICH, 1996, 1998, 2008a; OECD, 2009a,b,c) under Good Laboratory Practices (U.S. EPA, 1983; FDA, 1978). Early on in the testing process, when it was noted that the carcinogenicity potential for one species was not necessarily adaptable to another, chronic toxicology studies were required in two species (Jacobs and Hatfield, 2013). The 18-month rat and 12-month dog were later changed to include a 12-month rat and dog when a 2-year mouse alone was available (D'Aguanno, 1973; Goldenthal, 1968). The standard NCI/NTP carcinogenicity protocol adopted in 1976, includes 50 animal per sex per group and 2 test groups plus control as part of the 2-year study design, is still used today (NCI, 1976) most often with the addition of one extra test group and in some cases, initiating exposure *in utero*. In the performance of a battery of preclinical testing (acute, subchronic and chronic studies), it is recommended that all *in vivo* testing for a given product be performed in the same animal strain and, where possible with the same batch of test material. Use of concordant strains is particularly critical when performing corresponding immunotoxicology testing for a given chemical (U.S. EPA, 1998d; ICH, 2005, 2006) or when shorter term and/or preclinical supporting studies are conducted prior to the 2-year study. Targeted supporting testing for neurotoxic and developmental and reproductive/teratogenicity potential may be performed for applicable chemicals (FDA, 2000d,e; ICH, 1993; OECD, 1983, 1995, 1996, 1997a, 2001a). The most recent, extended one-generation study (replacing the two-generational study) has the advantages of optional testing schemes for animal usage and duration of study (OECD, 2001b, 2011). Prior to data assessment, the culmination of the 2-year study for safety concludes with the histologic interpretation ascribed to toxicologic pathologists, the global societies of which have defined strict procedures and standards for nomenclature and diagnostic assessment central to the identification of translational hazard identification (Crissman et al., 2004; Devor et al., 1994; Dua and Jackson, 1988; ESTP, 2005; Ettlin, 2012; Ettlin et al., 2010a, 2010b; Faccini et al., 1992; Keenan et al., 2009; Long et al., 1992; Ward, 2010; Young et al., 2011). The overprovision of tissues and the time taken for their evaluation has also led to calls for targeted analysis (Leblanc, 2000).