National Institute for Public Health
and the Environment
Ministry of Health, Welfare and Sport

Prediction of carcinogenic potential of substances using repeated dose toxicity data

RIVM Report 340700006/2012

M. Luijten | J.J.A. Muller | L.G. Hernandez |
L.T.M. van der Ven | J. van Benthem

National Institute for Public Health
and the Environment
Ministry of Health, Welfare and Sport

Prediction of carcinogenic potential of substances using repeated dose toxicity data

RIVM Report 340700006/2012

Colophon

M. Luijten
J.J.A. Muller
L.G. Hernández
L.T.M. van der Ven
J. van Benthem

Contact:
M. Luijten
Laboratory for Health Protection Research
Mirjam.Luijten@rivm.nl

This investigation has been performed by order and for the account of Ministry of Health, Welfare and Sports, within the framework of project V/340700 'Advisering en kennisbasis CMR'.

Abstract

Prediction of carcinogenic potential of substances based on data from repeated dose toxicity studies

Industry is required to test substances for their carcinogenic potential before they can be marketed. The conventional test for carcinogenicity is the long-term rodent carcinogenicity bioassay. However, this assay has several drawbacks. An alternative approach to determine the carcinogenic features of substances would be to use data from sub-chronic repeated dose toxicity studies. This approach could lead to a substantial reduction in the number of carcinogenicity studies performed without compromising human safety. This is concluded from a literature review carried out by the National Institute for Public Health and the Environment (RIVM).

The evaluation of relevant studies reported in the literature shows that the performance of histopathological lesions in sub-chronic toxicity studies as early indicators for the identification of carcinogenic substances varies highly. In contrast, the absence of evidence of histopathological lesions, hormonal perturbation and genotoxicity can be very accurate in predicting the lack of carcinogenic potential. Data derived from the literature support the refinement of current regulatory safety criteria for conducting two-year rat studies.

Keywords:
prediction, carcinogenicity, repeated dose toxicity, histopathology

Rapport in het kort

Voorspelling van carcinogeniteit van chemische stoffen door gebruik van gegevens uit kortdurende toxiciteitstudies

Voordat een chemische stof op de markt kan worden gebracht, is de producent verplicht om aan te tonen dat de stof geen verhoogd risico geeft op kanker. De klassieke test hiervoor is een tweejarige proefdiertest, die een zeer groot aantal proefdieren vereist en bovendien relatief veel vals positieve resultaten geeft. In bepaalde gevallen is een kortdurende toxiciteitstudie, die negentig dagen duurt, een mogelijk alternatief. Hierdoor zouden aanzienlijk minder proefdierstudies nodig zijn, zonder de volksgezondheid in gevaar te brengen. Dit blijkt uit een literatuurstudie van het RIVM.

In de literatuurstudie van het RIVM is onderzocht of resultaten uit kortdurende toxiciteitsstudies kunnen voorspellen of een stof kankerverwekkend is. Als je in een kortdurende test toxische effecten vindt, bijvoorbeeld een toename in de celdeling of sterk vergrootte cellen, dan zegt dit nog steeds weinig of de stof wel of niet kankerverwekkend is. Toxische effecten bleken namelijk met wisselend succes te voorspellen of een stof kankerverwekkend is. Bij dergelijke resultaten is dan toch de langdurige dierstudie nodig om hierover duidelijkheid te krijgen. Daar staat tegenover dat als je geen effect vindt de stof ook hoogstwaarschijnlijk geen kanker veroorzaakt. In dit geval is een tweejarige test niet nodig. Deze studie laat zien dat de huidige criteria om tot een tweejarige dierproefstudie over te gaan, kunnen worden verfijnd.

Trefwoorden:
voorspelling, carcinogeniteit, repeated dose toxicity, histopathologie

Contents

Summary

An accurate assessment of the carcinogenic potential of substances such as pharmaceutical drugs and pesticides is essential to protect humans and the environment. Therefore, many substances are extensively tested before they are marketed. Currently, the rodent 2-year bioassay is used to assess the carcinogenic potential of substances. However, over the years it has become clear that this assay yields many false positive results and also has several economic and ethical drawbacks including the use of large numbers of animals, the long duration, and the high cost. The need for a suitable alternative assay is therefore high.

An alternative approach to determine the carcinogenic potential of substances without carcinogenicity data is to use data from repeated dose toxicity studies. In the present report we evaluated whether histopathological parameters determined in repeated dose toxicity studies, particularly in sub-chronic 90-day studies, can be used to predict the carcinogenic potential of chemicals. For this, we searched the published literature and reviewed studies that examined possible correlations between substance toxicity and substance carcinogenicity. The studies found were analyzed in terms of their prediction accuracy. For the evaluation of the performance of a set of parameters to predict the carcinogenic potential of a substance we used as acceptance criteria a sensitivity of at least 80 percent, and a specificity of at least 70 percent.

In total, six different studies were identified, in which either a possible correlation between organ-specific parameters and carcinogenicity was examined or the predictive value of a so-called “whole animal response” was investigated. Results showed large variations in concordance ranging from 55-86 percent. Of the studies reviewed in this report, only two had a performance that met our criteria. Studies with sub-chronic 90-day exposures did not show an acceptable performance for carcinogenic potential. In the two reports that fulfilled the criteria, the repeated dose toxicity studies used for analyses consisted, at least in part, of studies with a minimum duration of 6 months. Early indicators were either histopathological lesions complemented with organ weight, or a whole animal response, together with hormonal perturbation and genotoxicity. A negative result for the latter set of parameters appeared to be highly predictive of a negative tumor outcome in 2-year rat studies. Using this approach, approximately forty percent of the rat carcinogenicity studies for pharmaceutical drugs could be eliminated without significantly altering human cancer risk assessment and with a substantial reduction in animal usage.

1 Introduction

Cancer is responsible for 7.6 million deaths worldwide per year, with 3 million new cancer cases per year in Europe alone. Whilst there is no single cause of cancer, evidence is emerging that exposure to substances in our every-day life may be a contributing factor to the increasing cancer burden. Identifying substances that cause cancer is of social and health concern. The carcinogenic potential of substances is determined using epidemiologic data and the 2-year rodent bioassay. Epidemiological data are only occasionally available, and only for existing substances. If available, they are often not conclusive because of many confounding factors. Further, epidemiological data cannot be used to determine the carcinogenic potential before a substance is marketed. Therefore, the 2-year cancer bioassay is in most instances the decisive test. Unfortunately, the 2-year cancer bioassay has also several economic and ethical drawbacks including the use of large numbers of animals, the long duration, and the high cost. In addition, there is considerable scientific doubt about the reliability of the assay, since too many false positive results have been observed. These are probably due to the prolonged use of rather high exposure levels and the high tumor background levels observed in rodents (Jacobson-Kram et al., 2004). For this reason, alternative methods to predict the carcinogenic potential of substances are in demand.

For carcinogenicity assessment of pharmaceuticals, in general, a 2-year bioassay in rats plus either a 2-year bioassay in mouse or an alternative 6-month transgenic mouse carcinogenicity study are required (Guidelines S1A, S1B, and S1C; International Conference on Harmonization, 1995, 1997, 2008). For substances in general, on the other hand, only clear positive evidence of genotoxic potential (*i.e.* positive in *in vivo* genotoxicity tests) or evidence from repeated dose studies that the substance is able to induce hyperplasia and/or preneoplastic lesions in combination with widespread dispersive use or evidence of frequent or long-term human exposure may in Europe trigger a 2-year cancer bioassay. Positive *in vivo* genotoxicity tests are generally seen as good determinants of the carcinogenic potential of substances. For instance, in Europe under REACH (EC, 2006) there are a number of conditions which can/may allow waiving of the carcinogenicity study, such as classification as a mutagen category 1A or category 1B. Here, the default presumption is that a genotoxic mechanism for carcinogenicity is likely. In this instance, cancer risk assessment has to be performed without carcinogenicity data.

Alternative approaches are needed to determine the carcinogenic potential of substances without carcinogenicity data. For this, two approaches may be considered. Obviously, genotoxicity data can be used qualitatively to predict carcinogenicity (as done in REACH), because these tests have a rather good sensitivity (65-78 percent, Lambert et al., 2005). However, this does not solve the problem for non-genotoxic carcinogens since they are negative in genotoxicity tests. Another approach, relevant for both genotoxic and non-genotoxic substances, may be to use toxicity and histopathological data from repeated dose studies, particularly from sub-chronic 90-day toxicity studies.

The key question addressed in this report is whether parameters determined in sub-chronic repeated dose toxicity studies can be used to predict the carcinogenic potential of substances. Several studies have been reported in which parameters that are routinely measured in sub-chronic studies have been evaluated in order to identify whether one or multiple parameters may be useful early indicators for the identification of carcinogens. In this report, we critically reviewed the published literature. Hereby, we focused on the applicability of histopathological parameters for carcinogen prediction. Other substance and

biochemical parameters (*i.e.* hematology, clinical chemistry) may be also relevant, but these were not addressed in this report.

2 Approach

2.1 Literature search and methods

A search was performed on internationally published literature for studies that examine possible correlations between substance toxicity and substance carcinogenicity. For this purpose PubMed was used to search for relevant articles with the keywords: "carcinogenicity", "prediction", and "bioassay"; and "carcinogenicity" and "predictors". We also searched for additional relevant studies in the reference lists of selected publications. The results of this search are summarized in Table 1. We explicitly did not include any immunotoxicity or reproductive toxicity studies because generally these studies do not include parameters that might be useful as early indicators for early carcinogenesis (reproductive studies) or are only available for a very small number of substances (immunotoxicity studies).

2.2 Acceptance criteria for parameters with predictive value

The studies found were analyzed in terms of their performance for identifying carcinogenic substances. For the evaluation of the performance of a (set of) parameter(s) to predict the carcinogenic potential of a substance we used the following acceptance criteria: the proportion of correctly predicted carcinogens should be high (*i.e.* sensitivity $\geq$ 80 percent), whereas the proportion of correctly predicted non-carcinogens should be reasonable (*i.e.* specificity $\geq$ 70 percent).

3 Literature review

Several studies have been reported in which the question was addressed whether parameters measured in repeated dose toxicity studies can be used to predict the carcinogenic potential of substances. These studies are summarized in Table 1. Generally, these studies can be divided in studies that only examined a possible correlation between organ-specific parameters and carcinogenicity and studies that investigated the predictive value of tissue changes and/or biochemistry in general, irrespective of the tissue specificity of the carcinogenic response.

Hoel et al. (1988) explored the inter-relationship between toxicity, genotoxicity and carcinogenicity in laboratory rodents. The evaluation was based on information obtained from both 90-day exposure studies and 2-year bioassays performed by the National Toxicology Program (NTP) on 99 substances. Of these 99 substances, 53 (54 percent) had positive results in the 2-year cancer bioassay in either rat or mice without taking into account sex differences. Organ toxicity was evaluated in both 90-day and 2-year studies.

Results of the 90-day exposure studies were available for 51 of the 53 carcinogens. For 22 chemicals, histopathological lesions were observed at some (13 substances) or all (9 substances) tumor sites. For the remaining chemicals, histopathological lesions were either absent (13 carcinogens) or observed at non-tumor sites (16 carcinogens). The predictive value of these lesions had a sensitivity of 43 percent (22/51). Findings for the 46 non-carcinogens were not reported. These data suggest that the identification of histopathological effects observed in 90-day exposure studies is not sufficient in itself for discriminating carcinogens.

Tennant et al. (1991) evaluated the relationship between chemical toxicity and carcinogenicity for a group of 31 substances that had been subjected to sub-chronic 90-day and chronic (*i.e.* 2-year) exposure experiments by the NTP. Based on the chronic study results, 22 of the 31 chemicals were considered to be carcinogenic. When evaluating the relationship between toxicity and carcinogenicity, Tennant and co-workers combined the data for rats and mice. Their results show that a diversity of toxic effects occurred independently of the mutagenic or carcinogenic potential of the chemicals. For both carcinogens and non-carcinogens, there were sites of sub-chronic (or even chronic) toxicity, or both, not associated with any neoplasia. For 10 of the 22 carcinogens, tumor induction with no associated toxicity at the target site was observed. However, among these were nine carcinogens that also induced sub-chronic toxicity at other sites not associated with neoplasia. The other 12 carcinogens induced tumors at a site for which sub-chronic toxicity was seen. Eight of the nine non-carcinogens did induce histopathological lesions after 90 days of exposure. However, only four of these induced lesions that were suggestive of a proliferative response. Based on these data, the sensitivity of histopathological lesions to predict the carcinogenic potential of substances is 55 percent (see Table 2). The authors concluded that "there are some types of sub-chronic toxicity that foreshadow carcinogenic potential, although causal associations are not yet established". This was particularly evident in their study for a group of carcinogens that induced sub-chronic nephrotoxicity. Although morphological characteristics of the renal neoplasms induced by these substances are similar, because of the diversity of the substances it is unlikely that there is a single mechanism common to all of these substances that can account for tumorigenesis.

Table 1: Inventory of studies

Reference	Duration of study	Carcinogens	NC	Total	Species	Strain	Sex
Hoel et al. (1988)	90-day	51	0	51	Rat	Fischer 344	M+F
	2-year	53	46	99	Mouse	$B6C3F_1$	
Tennant et al. (1991)	90-day	22	9	31	Rat	Fischer 344	M+F
					Mouse	$B6C3F_1$	
Elcombe et al. (2002)	7-, 28- and 90-days	9	0	9	Rat	Fischer 344	M+F
					Mouse	$B6C3F_1$	
Allen et al. (2004)†	≤12-month	11	76	87	Rat	Fischer 344	M+F
		27	56	83	Mouse	$B6C3F_1$	
Reddy et al. (2010)	6-and 12-month	30	50	80	Rat	NS	NS
Sistare et al. (2011)	6-and 12-month	66	116	182	Rat	NS	NS

NC = non-carcinogen; NS = not specified; †liver only.

Elcombe et al. (2002) evaluated the predictive value of several organ-specific parameters following exposure to nine non-genotoxic carcinogens. The substances selected included six liver, three thyroid gland, and four kidney carcinogens. The substances were administered to the same strains of mice and rats used in the original NTP bioassays. Liver, thyroid gland, and kidney were collected after 7, 28, and 90 days of exposure. Parameters measured as potential predictors included relative organ weight, histopathological observations, and biochemical measurements. No single parameter alerted specifically to carcinogenic outcome in rodent liver, thyroid gland, or kidney. For liver, only an increase in the relative weight of mouse liver following chemical treatment had a substantial predictive value, *i.e.* 4 out of 5 carcinogens were correctly predicted. Overall, none of the markers correlated well with rat liver carcinogenicity. For the thyroid, elevated levels of thyroid-stimulating hormone (TSH) were highly predictive in the mouse (moderate in rat). For two of the four rat renal carcinogens, increases in cell labeling indices and hyaline droplet formation acted as good predictors of carcinogenicity. However, the carcinogenicity of the remaining two rat renal carcinogens was poorly predicted by these parameters. The majority of the useful markers was evident at the early times studied (7 days and 28 days), but no overall best time for the measurement of all markers was identified. Despite the extensive results generated in this study, the number of substances evaluated was too small to draw general conclusions.

Allen et al. (2004) assessed the effectiveness of correlating specific hepatocellular pathology in pre-chronic studies (*i.e.* studies ≤ 12 months duration) with carcinogenic endpoints in the liver in the 2-year bioassay. Only studies using B6C3F1 mice and/or Fisher 344 rats were used for the analysis. In total, the number of substances reviewed was 83 for mice, and 87 for rat. The hepatic histopathological lesions that were evaluated were: hepatocellular necrosis, hepatocellular hypertrophy, hepatocellular cytomegaly, bile duct hyperplasia and hepatocellular degeneration. Increased liver weight (relative and/or absolute) was also used as an endpoint for analysis.

As a group, hepatocellular hypertrophy, necrosis and cytomegaly appeared to be parameters with predictive potential for carcinogenicity. In mice, 17 of the 27 liver carcinogens (10 false negatives) were successfully predicted, with only 2 false positives (see Table 2). However, these lesions did not identify 10/27 or 37 percent of the liver carcinogens. In rats, the same three liver lesions would again be reasonably effective at predicting liver carcinogenesis (64 percent) (Table 2). The inclusion of increased liver weight increased the sensitivity of predicting carcinogens to ~93-100 percent (Table 2). The enhanced sensitivity had a cost in increasing the number of false positives. Using these four parameters as predictors resulted in a prediction concordance of 63-76 percent. Although these findings seem promising, they are based only on 11 and 27 carcinogens in mice and rats, respectively. Clinical chemistry endpoints such as changes in liver enzymes were also explored as potential predictors of liver carcinogenicity. Several inconsistencies in results observed in the technical reports precluded any possible inclusion of clinical chemistry parameters in the final analysis. Altogether, histopathological parameters from prechronic studies alone were not considered to be specific for carcinogen prediction.

Given that parameters from 90-day toxicity studies do not seem to have sufficient carcinogen discriminatory power, Reddy et al. (2010) investigated the predictivity of preneoplastic lesions observed in 6- and 12-month toxicity studies for pharmaceuticals. They not only evaluated the available data on a tissue-by-tissue basis, but also employed a whole animal approach. The working hypothesis for the latter approach was that the lack of evidence for preneoplastic lesions at any site accurately predicts a substance's lack of carcinogenic potential. A relational database was built consisting of 6- and

12-month rat chronic toxicity studies. The data consisted of 62 (6-month) and 54 (12-month) studies.

The performance of chronic toxicity studies to correctly predict tissue-specific tumor outcome was evaluated for over 20 different tissues. The majority of the data available for these specific analyses were derived for non-carcinogens. Analysis of the 6-month data showed that the accuracy of histopathological findings to predict the carcinogenic potential of a pharmaceutical varied from 0 percent to 100 percent, depending on the tissue. Only preneoplastic changes in liver, adrenal, and stomach in a 6-month study had a sensitivity of predicting carcinogens that was 80 percent or higher. The results for the 12-month studies were highly comparable, with only liver and stomach meeting the acceptance criteria as defined in paragraph 2.2. The authors concluded that this approach is not useful, since the preneoplastic changes were not reliable predictors of tumor outcome in the corresponding tissues.

For the whole animal approach, the dataset included 30 rat carcinogens and 50 non-carcinogens in 2-year studies. Twenty-five out of the 30 carcinogens (83 percent) showed preneoplasia in any tissue, with a concordance of 75 percent (60/80, Table 2). To evaluate the effect of the additional 6 months of treatment on the accuracy of prediction, data for substances containing both 6- and 12-month studies were examined. This comparison, made for 36 substances, revealed that the predictivity was comparable (see Table 2) and no clear advantage was gained by the additional 6 months of treatment. According to the authors, the absence of preneoplasia in any tissue was highly predictive (88 percent) of a negative tumor outcome in 2-year rat studies. Based on this rationale, approximately fifty percent of the rat carcinogenicity studies would have been eliminated for the 80 pharmaceuticals examined, with no risk to humans and with a substantial reduction in animal usage and drug development time.

Encouraged by the promising findings reported by Reddy et al. (2010), Sistare and co-workers combined available data from 6- and 12-month toxicity studies on 182 marketed and non-marketed pharmaceuticals (Sistare et al., 2011). These 182 pharmaceuticals included the 80 pharmaceuticals analyzed by Reddy et al. and consisted of 66 carcinogens and 116 non-carcinogens. Besides histopathological lesions, evidence of hormonal perturbation and genotoxicity were included as potential predictors for carcinogenicity.

Using a single positive result among these three criteria as a test for outcome in the 2-year study, 52 of the 66 carcinogens were correctly identified, yielding 79 percent test sensitivity. Specificity and concordance were 53 percent and 63 percent, respectively (Table 2). Similar to the study of Reddy et al. (2010) the added value of 12 months of exposure, instead of 6 months, was evaluated. Here, 6-month and 12-month toxicity data were available for 53 substances. Analysis of these data provided evidence that, using the three criteria, there is little value gained by running a study for 12 months (Table 2).

The combination of the three test criteria of histopathologic risk factors, evidence of hormonal disruption, and genotoxicity were negative for 76 substances. Of these 76 pharmaceuticals, 62 substances (82 percent) were correctly predicted to be rat non-carcinogens. The authors claim that these analyses support a proposal to refine regulatory criteria for conducting a 2-year rat studies to be based on histopathological findings from rat 6-month studies, evidence for hormonal perturbation, and findings of a 6-month transgenic mouse carcinogenicity study. This proposed decision paradigm aims to reduce over 40 percent of 2-year rat studies without compromising human safety. Further investigation is needed to address whether the carcinogens that were negative for histopathological risk factors, hormonal perturbations and genotoxicity, have positive results in 6-month transgenic mouse studies, as proposed in this new paradigm.

Table 2: Summary of results

Reference	Parameter	Sensitivity (% correctly predicted carcinogens)	Specificity (% correctly predicted non-carcinogens)	Concordance (% correctly predicted substances)
Hoel et al. (1988)	Mouse and rat, histopathological lesions	43% (22/51)	-	-
Tennant et al. (1991)	Mouse and rat, histopathological lesions	55% (12/22)	56% (5/9)	55% (17/31)
Allen et al. (2004)	Mouse liver hypertrophy, necrosis, and cytomegaly	63% (17/27)	96% (54/56)	86% (71/83)
Allen et al. (2004)	Mouse liver hypertrophy, necrosis, cytomegaly and increase in liver weight	93% (25/27)	68% (38/56)	76% (63/83)
Allen et al. (2004)	Rat liver hypertrophy, necrosis, and cytomegaly	64% (7/11)	79% (60/76)	77% (67/87)
Allen et al. (2004)	Rat liver hypertrophy, necrosis, cytomegaly and increase in liver weight	100% (11/11)	58% (44/76)	63% (55/87)
Reddy et al. (2010)	Rat, whole animal response, 6- and 12-month toxicity studies, combined	83% (25/30)	70% (35/50)	75% (60/80)
Reddy et al. (2010)	Rat, whole animal response, 6-month toxicity studies	80% (12/15)	67% (14/21)	72% (26/36)
Reddy et al. (2010)	Rat, whole animal response, 12-month toxicity studies	87% (13/15)	62% (13/21)	72% (26/36)
Sistare et al. (2011)	Rat, histopathology, genotoxicity, hormonal disruption, 6- and 12-month toxicity studies combined	79% (52/66)	53% (62/116)	63% (114/182)
Sistare et al. (2011)	Rat, histopathology, genotoxicity, hormonal disruption, 6-month toxicity studies	74% (14/19)	62% (21/34)	66% (35/53)
Sistare et al. (2011)	Rat, histopathology, genotoxicity, hormonal disruption, 12-month toxicity studies	84% (16/19)	65% (22/34)	71% (38/53)

4 Discussion and Conclusions

In the present report we evaluated whether parameters determined in repeated dose toxicity studies may be useful as early indicators for the identification of carcinogens. To assess the performance of a (set of) parameter(s) we applied the following acceptance criteria: the sensitivity should be at least 80 percent, whereas the specificity should reach 70 percent at minimum. The rationale for these settings was that parameters measured in a sub-chronic toxicity study should have such a predictive value that their use in cancer risk assessment is feasible.

Of the parameters sets reviewed in this report, only two seem to have a sufficient performance. The first set includes mouse liver hypertrophy, necrosis, cytomegaly and increase in liver weight, determined in toxicity studies with a maximum duration of 12 months (Allen et al., 2004). However, the added value of this parameter set is limited, because its predictivity has only been demonstrated for one organ and the human relevance of liver tumors in B6C3F1 mice is questionable given the high background levels in this strain of mice (Scheepmaker et al., 2005).

The other set consists of whole animal response, hormonal perturbation and genotoxicity (Reddy et al., 2010; Sistare et al., 2011). This approach was tested using 6-month and 12-month toxicity data. Comparison of these data indicated that the sensitivity of a 12-month study is not substantially higher than a 6-month study. Moreover, negative results for each of these criteria predicted negative findings in rat carcinogenicity studies with an accuracy of 82 percent. These analyses suggested that regulatory criteria for conducting a two-year rat study may be refined and be based on assessment of histopathological findings from a rat six-month study, evidence of hormonal perturbation, and genetic toxicology results, together with the findings of a six-month transgenic mouse carcinogenicity study.

Most of the individual histopathological parameters investigated were found to be insufficiently predictive. This may be due to the fact that several important aspects were not or only poorly considered. First, the dose levels used were in most studies not taken into account. In general, exposure levels tested in a sub-chronic repeated dose toxicity study are higher than in a chronic carcinogenicity study. The relevance of histopathological lesions at very high and thus probably non-physiological doses is questionable and may, if wrongly interpreted, influence the conclusion on the predictive value of these parameters. Possible future studies should include criteria how to deal with differences in dosing levels between sub-chronic and chronic studies. An example how to define and how to use these criteria has been reported by Reddy et al. (2010) and Sistare et al. (2011).

A second aspect is species and sex. Generally a substance is called a full carcinogen if it induces tumors in both sexes of at least two animal (rodent) species. Commonly, rats and mice are used, despite the known high background tumor sensitivity in the mouse (leading to false positive results). In the present report, species and sex specificity of the carcinogenicity of a substance was not considered. Substances that were carcinogenic exclusively in rats or mice or only in one sex were put on the same level. This may have considerable effects on the outcome of a comparison like the present one. It is particularly important to consider sex differences. For instance, histopathological findings in males but tumors in females may, if sex differences are not considered, lead to a "predictive" call whereas if sexes are distinguished it will be a negative call for each of the sexes.

A third aspect is the composition of the databases under study. Reliable databases are essential for the evaluation of histopathological parameters as early indicators for carcinogenicity. The composition of a database has a major

influence on the outcome of a study based on this database. A good database comprises a reliable number of relevant substances. In most studies reviewed in this report relatively small databases were used. A disadvantage of small databases is that they may contain substances that are strongly biased for the effect(s) under investigation. It is obvious that the reliability of generally valid conclusions decreases with size and that firm conclusions can never be drawn from too small databases. A reliable database, relevant for the domain for which a prediction is made, should contain both negative and positive substances in about the same numbers for the effect under investigation, which means for the present report carcinogens and non-carcinogens. The databases used for the present report either do not meet or only approximate these criteria. For instance, in the studies by Hoel et al. (1988) and Elcombe et al. (2002) the databases exclusively contained carcinogens. On the other hand, in the studies of Reddy et al. (2010) and Sistare et al. (2011) the non-carcinogenic pharmaceuticals outnumbered the carcinogenic ones.

In most studies, only histopathological findings were considered as predictors of carcinogenicity. However, exposure to a substance may result in other (adverse) effects, such as clinical, biochemical and hematological effects or endocrine disruption, which in turn may directly or indirectly influence carcinogenicity. These effects may influence both the sensitivity and the specificity of histopathological findings to predict the carcinogenic potential of a substance. For instance, Allen et al. (2004) demonstrated that where mouse liver hypertrophy, necrosis and cytomegaly were rather good predictors of carcinogenicity, the sensitivity increased when liver weight was added to these three parameters. Unfortunately, the addition of liver weight was detrimental to the specificity of these predictors. This may be related to the fact that some level of toxicity is required in repeated dose studies. Therefore, almost every substance will have some effect at the highest dose level resulting in a high sensitivity and a low specificity.

Apparently, the authors of the studies used comparing liver and multi-site parameters from sub-chronic repeated dose toxicity studies to carcinogenicity in the 2-year cancer bioassay suggest that repeated dose toxicity studies are not specific enough for discriminating carcinogens. However, if all the above mentioned remarks and/or shortcomings are considered, most of these studies do not justify such a strong conclusion. The conclusion whether histopathological findings with or without other toxicological parameters of sub-chronic repeated dose toxicity studies can be predictive for carcinogenicity should not be taken on the basis of the available review papers from the open literature. Such an investigation needs a specific approach. In the present report taking into account all remarks and shortcomings, still indications were observed that prediction of carcinogenicity by histopathological findings from sub-chronic repeated dose toxicity studies may be promising (*e.g.* Elcombe et al., 2002). To investigate this promising prediction a better comparison with better tests is needed. Therefore, we started an investigation with a reliable number of relevant substances for which robust sub-chronic repeated dose toxicity and carcinogenicity tests are available performed according generally accepted protocols (*e.g.* OECD guidelines). In this investigation criteria will be introduced under which conditions substances can be accepted or need to be removed from the database, taken into account the level of robustness of the carcinogenicity (positive in one of 2 species or in one or two sexes), the dosing regimes and the strain of rodents used. Only after this investigation it may be justified to conclude whether or not sub-chronic histopathological data allow the prediction of carcinogenicity.

References

Allen, D.G., G. Pearse, J.K. Haseman, and R.R. Maronpot (2004) Prediction of rodent carcinogenesis: an evaluation of prechronic liver lesions as forecasters of liver tumors in NTP carcinogenicity studies. Toxicol Pathol 32: 393-401.

EC (2006) Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC.

Elcombe, C.R., J. Odum, J.R. Foster, S. Stone, S. Hasmall, A.R. Soames, I. Kimber, and J. Ashby (2002) Prediction of rodent nongenotoxic carcinogenesis: evaluation of biochemical and tissue changes in rodents following exposure to nine nongenotoxic NTP carcinogens. Environ Health Perspect 110: 363-375.

Hoel, D.G., J.K. Haseman, M.D. Hogan, J. Huff, and E.E. McConnell (1988) The impact of toxicity on carcinogenicity studies: implications for risk assessment. Carcinogenesis 9: 2045-2052.

International Conference on Harmonization (1995) ICH Guideline S1A: Need for Carcinogenicity Studies of Pharmaceuticals. November 1995.

International Conference on Harmonization (1997) ICH Guideline S1B: Testing for Carcinogenicity of Pharmaceuticals. July 1997.

International Conference on Harmonization (2008) ICH Guideline S1C: Dose Selection for Carcinogenicity Studies. March 2008.

Jacobson-Kram, D., F.D. Sistare, and A.C. Jacobs (2004) Use of transgenic mice in carcinogenicity hazard assessment Toxicol Pathol 32 Suppl 1: 49-52.

Lambert, I.B., T.M. Singer, S.E. Boucher, and G.R. Douglas (2005) Detailed review of transgenic rodent mutation assays. Mutat Res 590: 1-280.

Reddy, M.V., F.D. Sistare, J.S. Christensen, J.G. Deluca, G.K. Wollenberg, and J.J. Degeorge (2010) An evaluation of chronic 6- and 12-month rat toxicology studies as predictors of 2-year tumor outcome. Vet Pathol 47: 614-629.

Scheepmaker, J.W.A., C.E. Smit, and M.T.M. van Raaij (eds.) (2005) Factsheets for the (eco)toxicological risk assessment strategy of the National Institute for Public Health and the Environment. Part V. RIVM report 601516013.

Sistare, F.D., D. Morton, C. Alden, J. Christensen, D. Keller, S.D. Jonghe, R.D. Storer, M.V. Reddy, A. Kraynak, B. Trela, J.G. Bienvenu, S. Bjurstrom,

V. Bosmans, D. Brewster, K. Colman, M. Dominick, J. Evans, J.R. Hailey, L. Kinter, M. Liu, C. Mahrt, D. Marien, J. Myer, R. Perry, D. Potenta, A. Roth, P. Sherratt, T. Singer, R. Slim, K. Soper, R. Fransson-Steen, J. Stoltz, O. Turner, S. Turnquist, M. van Heerden, J. Woicke, and J.J. Degeorge (2011) An analysis of pharmaceutical experience with decades of rat carcinogenicity testing: support for a proposal to modify current regulatory guidelines. Toxicol Pathol 39: 716-744.

Tennant, R.W., M.R. Elwell, J.W. Spalding, and R.A. Griesemer (1991) Evidence that toxic injury is not always associated with induction of chemical carcinogenesis. Mol Carcinog 4: 420-440.

Published by:

National Institute for Public Health and the Environment
P.O. Box 1 | 3720 BA Bilthoven
The Netherlands
www.rivm.com