Animal Carcinogenicity Studies: Poor Human Predictivity

Andrew Knight BSc., BVMS, Cert AW, MRCVS, Jarrod Bailey PhD, Jonathan Balcombe PhD
Animal Consultants International www.AnimalConsultants.org

Abstract

The regulation of human exposures to potentially carcinogenic chemicals constitutes society's most important use of animal carcinogenicity data. Environmental contaminants of greatest U.S. concern are listed in the Environmental Protection Agency's (EPA) Integrated Risk Information System (IRIS) chemicals database. Most of these lack human exposure data. However, of the 160 IRIS chemicals possessing animal data but lacking human data as of January 1, 2004, we found that in most cases (58.1 %; 93/160) the EPA considered the animal data inadequate to support a classification of probable human carcinogen or non-carcinogen. For the 128 chemicals with human or animal data also assessed by the World Health Organization's International Agency for Research on Cancer (IARC), human carcinogenicity classifications were compatible with EPA classifications only for those 17 having at least limited human data ($p = 0.5896$). For those 111 primarily reliant on animal data, the EPA was much likelier than the IARC to assign carcinogenicity classifications indicative of greater human risk ($p < 0.0001$). The IARC is a leading international authority on carcinogenicity assessments, and the very significant differences in human carcinogenicity classifications of identical chemicals between the IARC and the EPA indicate that: (i) in the absence of significant human data the EPA is over-reliant on animal carcinogenicity data, (ii) as a result, the EPA tends to over-predict carcinogenic risk, and (iii) the true human specificity, and hence predictivity, of animal carcinogenicity data is even poorer than indicated by EPA figures alone.

Introduction

Since the first chemical bioassay in 1915, when Yamagiwa and Ichikawa showed that coal tar applied to rabbit ears caused skin carcinomas, several thousand have been conducted, with the objective of determining human carcinogenic risks for the great majority of chemicals lacking human exposure data. However, animal carcinogenicity testing remains a controversial area of research.

Proponents claim that all known human carcinogens that have been studied in sufficient animal species have produced positive results in one or more species. Critics respond that if enough animal testing is conducted, carcinogenesis will eventually occur in *some* species, regardless of human cancer risk. A *Mutagenesis* study found that of 20 human *non*-carcinogens, 19 produced carcinogenic effects in animals.

The most important use of animal carcinogenicity data lies in the regulation of human exposures to potential carcinogens. The U.S. Federal agency most responsible for regulating exposures to environmental contaminants is the Environmental Protection Agency (EPA), and the chemicals of greatest public health concern are listed within its Integrated Risk Information System (IRIS) chemicals database, along with their animal toxicity data and consequent human carcinogenicity assessments.

To assess the utility of animal carcinogenicity data in deriving human carcinogenicity assessments, we surveyed the IRIS chemicals database. To assess the reliability of the EPA carcinogenicity assessments obtained from animal test data, we compared them with those of a leading world authority, the World Health Organization's International Agency for Research on Cancer (IARC).

Methods

The 543 chemicals catalogued in the EPA's IRIS chemicals database (as of January 1, 2004) were examined to determine the proportion for which the EPA was able to derive classifications of "probable human carcinogen" or "probable human non-carcinogen" based primarily on animal carcinogenicity data. The relatively few classifications of "definite human carcinogen" relied primarily on available human exposure data. The remaining classifications of "possible human carcinogen" or "unclassifiable" were not considered substantially useful for risk assessment or regulatory purposes. They are excluded from the U.S. National Toxicology Program annual *Report on Carcinogens*.

Of the 177 chemicals considered by the EPA to possess at least limited human or animal data, 128 were assigned human carcinogenicity classifications by both the EPA and the IARC. Of these 128, 17 were considered by the EPA to possess at least limited human data, while 111 were primarily reliant on animal data.

The consistency of classifications between the EPA and IARC was examined for these two groups by comparing the carcinogenicity classification proportions within each group via chi-square tests, and also by comparing the individual classifications of the 111 chemicals primarily reliant on animal carcinogenicity data.

Chi-squared tests provide statistical calculations of the probability that two data sets, such as EPA and IARC human carcinogenicity classifications, are samples from the same underlying data population, and that any observed differences are simply due to random sampling variation. Large chi-squared values reflect increased probabilities that observed differences are due to real differences in underlying data populations.

Results

EPA human carcinogenicity classifications

Of the 543 chemicals catalogued in the EPA's IRIS chemicals database, 235 had been assigned human carcinogenicity classifications. Of these, 17 were classified as definite (A) or probable (B1) human carcinogens on the basis of their human carcinogenicity data. Of the remaining 218 chemicals lacking even limited human data, 160 were deemed to possess animal carcinogenicity data, primarily sourced from the biomedical literature (B2, C, subset of D, and E; Table 1).

Table 1: EPA Human Carcinogenicity Classifications of IRIS Chemicals

EPA human carcinogenicity classification (with basis for classification)	No. of chemicals	% of total
A: Human Carcinogen (convincing human data)	11	4.7
B1: Probable Human Carcinogen (limited human data)	6	2.6
B2: Probable Human Carcinogen (sufficient animal data)	64	27.2
C: Possible Human Carcinogen (animal data inadequate for stronger classification)	40	17
D: Unclassifiable (animal data inadequate for stronger classification)	53	22.6
D: Unclassifiable (no animal or human data)	58	24.7
E: Probable Human Non-Carcinogen (sufficient animal data)	3	1.3
TOTAL	235	

160 chemicals lacked even limited human data but possessed animal data.
Source: EPA Integrated Risk Information System database, Jan. 1, 2004.

Human specificity and utility of animal carcinogenicity data based on EPA figures

Of the 160 EPA chemicals lacking even limited human data (A or B1) but having animal data (B2, C, subset of D, and E), 64 were considered probable human carcinogens (B2), and three were considered probably not carcinogenic to humans (E). The remaining 93 chemicals were considered possible human carcinogens (C; 40) or unclassifiable as to their human carcinogenicity (D; 53) based on animal data considered inadequate to support a stronger classification (Table 1).

In sum, of those 160 chemicals lacking even limited human data but having animal data, the EPA considered the animal data inadequate to support the substantially useful classifications of probable human carcinogen or probable human non-carcinogen in the majority of cases (93/160; 58.1%, 95% CI: 50.4 - 65.5).

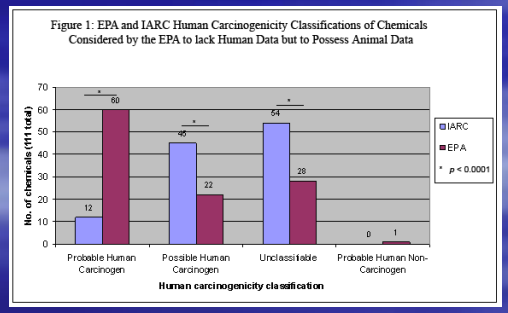
Comparison of EPA and IARC human carcinogenicity classifications

Of those 177 chemicals considered by the EPA to possess human or animal data (A, B1, B2, C, D with animal data, or E), 128 were also assessed by the IARC. Of these, 17 were considered by the EPA to possess at least limited human data (A or B1), and the remaining 111 EPA carcinogenicity classifications were primarily reliant on animal data.

For those 17 chemicals considered by the EPA to possess at least limited human data, overall EPA classifications were not found to differ significantly from those predicted by IARC classifications (Chi-squared = 0.291, 1 df, $p = 0.5896$).

However, for those 111 chemicals considered by the EPA to lack even limited human data, but to possess animal data, EPA and IARC classifications were very significantly different overall (chi-squared = 215.548, 2 df, $p < 0.0001$; Figure 1).

Figure 1: EPA and IARC Human Carcinogenicity Classifications of Chemicals Considered by the EPA to lack Human Data but to Possess Animal Data



Source: *IARC Monographs Programme on the Evaluation of Carcinogenic Risks to Humans*, and the EPA Integrated Risk Information System database, Jan. 1, 2004.

The EPA was much likelier than the IARC to assign carcinogenicity classifications indicative of greater human hazard. The EPA classified 60 chemicals as probable human carcinogens and 51 in all other categories, which was very significantly different from the IARC figures of 12 and 99 respectively (chi-squared = 215.273, 1 df, $p < 0.0001$). Similar disparities were found for possible human carcinogens (chi-squared = 19.771, 1 df, $p < 0.0001$) and unclassifiable chemicals (chi-squared = 24.378, 1 df, $p < 0.0001$).

Comparison of the individual classifications of these 111 chemicals revealed that 67 (60.4%) were assigned an EPA carcinogenicity classification indicative of greater human hazard, 38 (34.2%) were assigned an equivalent classification, and 6 (5.4%) were assigned a classification indicative of lesser human hazard, than the corresponding IARC classification of the same chemical.

Discussion

The sensitivity of a human carcinogenicity assay (test) refers to its ability to yield a positive result in the presence of a human carcinogen. Its specificity refers to its ability to yield a negative result in the presence of a human *non*-carcinogen. A desirable test has both high sensitivity (correctly detecting most carcinogens) and specificity (minimizing false positive results).

Based on EPA figures alone, the specificity of animal carcinogenicity data for deriving substantially useful human carcinogenicity classifications is clearly poor. Of those 160 IRIS chemicals lacking even limited human data but possessing animal data, the EPA considered the animal data inadequate to support substantially useful human carcinogenicity classifications in the majority (93) of cases.

However, IARC assessments of the same chemicals reveal that the human utility of animal carcinogenicity data is probably even lower than indicated by EPA figures. EPA and IARC carcinogenicity classifications were similar only for those chemicals possessing human data. For those possessing only animal data, the EPA was much likelier than the IARC to assign carcinogenicity classifications indicative of greater human hazard.

Given that the IARC is recognized as a leading international authority on human carcinogenicity classifications, the very significant differences in classifications of identical chemicals between the IARC and the EPA indicate that:

(i) in the absence of significant human data the EPA is over-reliant on animal carcinogenicity data,
(ii) as a result, the EPA tends to over-predict carcinogenic risk, and
(iii) the true human specificity, and hence predictivity, of animal carcinogenicity data is even poorer than indicated by EPA figures alone.

Our findings corroborate those of previous investigators. In response to a 2000 Congressional directive, the EPA undertook an evaluation of the data variability and uncertainty within its IRIS assessments. A representative sample of 16 IRIS assessments were subjected to in-depth evaluation by a panel of six independent experts, who concluded that despite being advertised as quantitative science-based classifications, some were, in fact, more grounded in EPA policy favoring classifications indicative of greater human risk.

However, EPA human carcinogenicity assessments are by no means more suspect than those of other U.S. regulatory agencies. In their survey of 350 representative chemicals, Viscusi and Hakes (1998) found that the carcinogenicity assessments of other U.S. regulatory authorities, particularly the Food and Drug Administration and the Occupational Safety and Health Administration, are even less reflective of carcinogenicity data than those of the EPA.

The poor human specificity of animal carcinogenicity studies was also demonstrated by Tomatis & Wilbourn (1993) and Haseman (2000), and further described by Rall (2000), Ashby and Purchase (1993), Fung *et al.* (1995) and Ennever and Lave (2003).

Conclusions

By 1998, only about 2,000 (2.6%) of the 75,000 industrial chemicals in use and listed in the EPA's Toxic Substances Control Act inventory, had been tested for carcinogenicity. The cost of testing these 2.6% of industrial chemicals was millions of animal lives, millions of skilled personnel hours, and hundreds of millions of dollars.

The most important use of the animal data thus derived is in the regulation of human exposures to potential carcinogens by governmental agencies such as the EPA. However, our results demonstrate that the human specificity of animal carcinogenicity data is inadequate for the EPA to derive substantially useful human carcinogenicity classifications for the majority (58.1%) of chemicals of greatest public health concern.

The very significantly different human carcinogenicity classifications of identical chemicals between the EPA and the IARC—a leading international authority on carcinogenicity assessments—clearly illustrate the over-reliance of the EPA on animal carcinogenicity data. The result is that the EPA over-predicts carcinogenic risk. Hence the true human specificity, and predictivity, of animal carcinogenicity data is even poorer than indicated by EPA figures alone.

The sensitivity of the traditional rodent bioassay in detecting human carcinogens for *some* sex-species combination is not in question. However, its poor human specificity severely undermines its utility for predicting human carcinogenicity, and consequently, its use in regulating exposures to potential human carcinogens. The implementation by regulatory authorities of alternative assays with superior human predictivity results is clearly necessary.

Acknowledgements

We gratefully acknowledge the assistance of the Physicians Committee for Responsible Medicine, Washington DC, in funding this research, and of the Japan Anti-Vivisection Association, Tokyo, in funding this poster.

References

Available on request.