

This is a draft paper for discussion. It should not be quoted, cited or reproduced.

CC/06/12

**COMMITTEE ON THE CARCINOGENICITY OF CHEMICALS IN FOOD,
CONSUMER PRODUCTS AND THE ENVIRONMENT**

Creative Accounting": Report by People for the Ethical Treatment of Animals (PETA)

Introduction

1. The Home Office has received a report from the organisation "People for the Ethical Treatment of Animals (PETA)" entitled "Creative Accounting: (Mis)judging the costs and benefits of rodent cancer studies by the UK Home Office" (1). The Home Office has asked for advice from the COC on this report and on its recommendations.

PETA report

2. The report is attached at Annex 1. A number of publications are cited and the secretariat has annotated the report to indicate where recent publications are also attached as annexes. A summary of each paper is given on the covering page to the relevant annex.

3. The PETA report discusses the issue of carcinogenicity testing by making an assessment of the benefits of the studies (considering subjects such as repeatability, relevance to humans and interpretability of results) and an assessment of the costs (in terms of animals and money). The committee may wish to consider in particular PETA's comments and conclusions on the benefits of the studies i.e. their usefulness and scientific validity.

4. Note that on page 11 of the report, PETA states that a 1983 paper by Salzburg (2) reported that rodent studies showed no carcinogenic effects for 12 of 19 chemicals listed by IARC as known human carcinogens, which suggests that the false negative rate may be as high as 63.2%. The paper by Salzburg refers to a list of 26 *putative* human carcinogens listed by Tomatis *et al* in 1978 (3). Salzburg excluded the seven chemicals to which humans are exposed by inhalation only, and considered the results of long term feeding studies in rat or mice on the other 19 chemicals in making his comparison. Unfortunately, no details of the chemicals involved are included in the paper.

5. A more recent paper by Ennever and Lave (4; see Annex 4) cites a 2001 letter by Johnson (5) which lists 10 known chemicals which have been classified by IARC as Group 1 human carcinogens or by the US DHHS/NTP Report on Carcinogens as "known" human carcinogens, and which have been tested in the NTP rodent carcinogenicity studies. Using data on these 10 chemicals, Ennever and Lave calculate the following:

This is a draft paper for discussion. It should not be quoted, cited or reproduced.

- on the assumption that a chemical is positive for carcinogenicity in rodents if it gives a positive result in any rodent study, then only 1/10 of the human carcinogens was not also carcinogenic in rodents (10% false negative rate),

- on the assumption that a chemical is positive in rodents only if it gives a positive result in both species, then only 3/6 of the human carcinogens tested in both species were also carcinogenic in rodents (50% false negative rate),

6. Relevant to the discussion of the PETA report, Members may wish to note a recent review by Anisimov *et al* (6) which discusses the similarities and differences in cancer characteristics between humans and rodents. It concluded that although “there are still significant differences between rodents and humans in the way cancer develops...(these)... do not diminish the importance of animal modelling. Rather, they warn against simplified extrapolation of the results of rodent experiments to humans and call for further investigation of this problem to reliably predict cancer risks, as well as foster success in treating human cancers based on data from laboratory animal studies.”

7. The PETA report does not address more subtle regulatory questions such as the value of studies in 2 species or whether an *in vivo* mutagen needs to be tested for carcinogenicity. In 1997, the committee addressed the utility of carcinogenicity studies in the mouse (7). In this work, the Gold Carcinogenicity Potency Database was used to identify potential mouse-specific carcinogens. The COC identified only 15 convincing mouse-specific carcinogens out of 559 chemicals with published carcinogenicity data in 2 species. Of these, the committee considered that two chemicals, both of which were genotoxic *in vivo*, posed a potential public human health hazard. The other 13 chemicals were considered likely to pose a minimal or, at most, equivocal health hazard. The committee cautioned that these conclusions must be viewed within the limitations of the data available and the regulatory context within which mouse carcinogenicity studies were required.

8. The committee is asked for a view on the PETA report. Is there any further work which the committee considers it could usefully undertake on the scientific value of rodent carcinogenicity studies?

Secretariat
2006

References

1. PETA. Creative Accounting: (Mis)judging the costs and benefits of rodent cancer studies by the UK Home Office. PETA Europe Ltd, May 2006.

2. Salzburg D (1983). The lifetime feeding study in mice and rats – An examination of its validity as a bioassay for human carcinogens. *Fundam Appl Toxicol* 3: 63-67.

This is a draft paper for discussion. It should not be quoted, cited or reproduced.

3. Tomatis L, Agthe C, Bartsch H, Huff J, Montesano R, Saracci R, Walker E and Wilbourn J. (1978). Evaluation of the carcinogenicity of chemicals: A review of the monograph program of the IARC (1971 to 1977). *Cancer Research* 38: 877-885.
4. Ennerver FK and Lave LB (2003). Implications of the lack of accuracy of the lifetime rodent bioassay for predicting human carcinogenicity. *Regulatory Toxicol and Pharmacol.* 38: 52-57.
5. Johnson FM (2001). Response to Tennant al: Attempts to replace the NTP rodent bioassay with transgenic alternatives are unlikely to succeed. *Environ Mol Mut* 37: 89-92.
6. Anisimov VN, Ukraintseva SV and Yashin AI (2005). Cancer in rodents: does it tell us about cancer in humans? *Nature Reviews Cancer.* 5: 807-819.
7. 1997 Annual Report of the Committees on Toxicity, Mutagenicity, Carcinogenicity, pp 117-121. www.advisorybodies.doh.gov.uk/coc/1997ar.pdf

This is a draft paper for discussion. It should not be quoted, cited or reproduced.

Annex 2 to CC/06/12

Gottman E, Kramer S, Pfahringer B and Helma C (2001). Data quality in predictive toxicology: Reproducibility of rodent carcinogenicity experiments. *Env Health Persp* 109 (5), 509-514.

Summary

1. Gottman *et al* compared the results of NTP rodent carcinogenicity studies on 121 chemicals with the results of studies on the same chemicals taken from the general scientific literature¹. Both sets of data were taken from the Gold Carcinogenic Potency Database. It is noted that only the NTP studies were conducted with a standard protocol, thus it appears that the studies which were compared were likely to have had different protocols. The criteria used by Gold for selection of the literature studies are known to be far less stringent in terms of design and reporting than that for the NTP studies or the relevant OECD guidelines (Gold and Zeiger, 1997). Studies using different routes of administration also appear to have been compared. The classification as to carcinogenicity of the chemicals used in each study was based on the opinions of the study's authors and a compound was classified as a carcinogen if a positive result was obtained in at least one study.

2. Overall, 69 chemicals (57%) showed concordance in the authors' classifications from the NTP studies and literature studies. 57% of these were consistently classified as positive and 43% had negative results in studies from both sources. A number of sub-analyses were made:

Species: from 70 investigations with mice, 49% had concordant results; from 71 with rats, 62% were concordant, indicating a higher reproducibility in the rat.

Sex: The concordance of male mice was 46%, of female mice 36%, of male rats 55% and of female rats 69%.

Strain: In the NTP studies, 3 different rat strains and 1 mouse strain were used; in the literature studies, 29 different rat strains and 37 different rat strains were used. The most frequently used strains in both data sets were Fischer F344/N rats and B6C3F₁ mice. The concordance for these strains was 53% for male rats, 64% for female rats, 39% for male mice and 33% for female mice. The authors conclude that, as these results are close to the overall concordance rates, the poor reproducibility of carcinogenicity assays may not be due to different strains in the two literature data sets.

One- or two-species carcinogens: for carcinogenic compounds, 48% were one-species carcinogens and 52% were two-species carcinogens in the NTP data set; in the literature, 73% were one-species carcinogens and 27% were two-species carcinogens. A comparison of NTP data with literature data showed that 58% of

¹ A list of the chemicals is given in the paper.

This is a draft paper for discussion. It should not be quoted, cited or reproduced.

compounds were classified concordantly, but 'with low coefficients of association and without an indication for a better concordance as obtained for the overall data set'. From concordant one-species data, only 31% affected the same species.

Target-organ specificity: target sites for carcinogenicity were grouped into 11 basic target categories (eg digestive system, liver, cardiovascular system). In the NTP data set, 36% of chemicals were classified as one-category carcinogens and 64% as multicategory carcinogens. In the literature data set, results were similar with 42% classified as one-category carcinogens and most (56%) as multicategory carcinogens. A comparison of the two data sets showed that only 52% of the chemicals were concordantly classified. The majority were multicategory carcinogens but less than 50% of these caused tumours in the same categories.

Ames positive or negative chemicals: An insufficient number of the 121 chemicals also had Ames test evaluation data in the Gold Database to allow calculations to be made. Therefore, the authors used the Database as a whole and identified mutagenicity data on 178 chemicals from the NTP part and 272 from the literature part (overlapping data excluded). 57% of chemicals in the NTP studies and 64% in the literature studies were Ames positive. Using the NTP studies only, because they were conducted to standardised protocols, they found that 52% of one-tissue category carcinogens and 55% of one-species carcinogens were Ames negative, whereas the majority of multicategory carcinogens (66%) and two-species carcinogens (73%) were Ames positive.

3. The authors conclude that the unexpectedly large discordance between the results from the NTP data and the literature data in the Gold Databases were due to differences in the experimental protocols used e.g. the number and range of doses, administration route and group size. They consider the results provide an indication that rodent carcinogenicity studies are, in general, poorly reproducible. However, a more valid comparison might be between the results of studies which have been conducted to a standard protocol, e.g. the relevant OECD guideline.

This is a draft paper for discussion. It should not be quoted, cited or reproduced.

Annex 3 to CC/06/12

Gaylor DW (2005). Are tumor incidence rates from chronic bioassays telling us what we need to know about carcinogens?. Reg Toxicol Pharmacol 41, 128-133.

Summary

1. The primary purpose of this paper was 'to estimate the statistical false negative rate for carcinogenicity based on a retrospective analysis of tumour incidence rates observed in a series of standard NTP rodent carcinogenicity studies and an evaluation of the effect of using a sample size of 50 animals per group at the maximum tolerated dose'. 50 animals per dose group is the standard group size in NTP studies and the highest dose is usually the maximum tolerated dose (MTD).

2. The authors used existing statistical dose-response trend tests from published NTP carcinogenicity studies to estimate the probability that a statistically significant trend would be detected with a larger sample size of 200 animals per dose group. A statistical significance rate of $p \leq 0.01$ was used for all tissue sites. They considered studies on 156 chemicals, tested in both sexes of rats and mice. In the NTP studies, 97/156 (62%) of these were positive (some or clear evidence of carcinogenicity), 23 had equivocal evidence of carcinogenicity, and 36 were considered to show no evidence of carcinogenicity. With 200 animals per group, it was estimated that at least 95.1/156 would be positive, at least 22.5 show equivocal evidence, and 10 or fewer would be negative (see table 1 of paper). (The authors explain that because of inherent experimental variation, not all the chemicals deemed carcinogenic or equivocally carcinogenic in the standard study would necessarily produce a statistically significant dose-response trend at the $p \leq 0.01$ level if 200 animals per dose group had been used). Overall, it was estimated that at least 92% of all chemicals tested at the MTD would produce a statistically significant dose-response trend at one more more tissue site.

3. The authors conclude that the results demonstrate the lack of power of the standard NTP carcinogenicity study to detect carcinogens at the MTD, and that it appears that almost all chemicals would demonstrate carcinogenicity at high doses if tested in an adequate number of animals. They conclude that there appears to be little need to screen chemicals for carcinogenicity in 2-year chronic studies at the MTD. However they state that bioassays are needed to provide a dose-response relationship between dose and carcinogenic effects for use in cancer risk assessments and, therefore, do not recommend that the number of doses or the number of animals per dose should be altered.

This is a draft paper for discussion. It should not be quoted, cited or reproduced.

Annex 4 to CC/06/12

Ennerver FK and Lave LB (2003). Implications of the lack of accuracy of the lifetime rodent bioassay for predicting human carcinogenicity. Regulatory Toxicol and Pharmacol. 38: 52-57.

No summary is provided as the key aspect of this paper is highlighted in the covering paper.

This is a draft paper for discussion. It should not be quoted, cited or reproduced.

Annex 5 to CC/06/12

Knight A, Bailey J and Balcombe J (2006). ATLA 34, 19-27. Animal carcinogenicity studies: 1. Poor human predictivity.

Summary

1. The authors compared the carcinogenicity classifications for the 128 chemicals which had undergone assessments both by the EPA for the IRIS database and by IARC. (They excluded chemicals in the IRIS database regarded as 'unclassifiable due to no animal or human data' (a subset of D, see Table 1) but it is not clear that they excluded any from the IARC category of 'unclassifiable').

2. Of these 128 chemicals, 17 were considered by the EPA to possess at least limited human data and the remaining classifications were reliant mainly on animal data. For these 17 chemicals, overall EPA classifications were not found to differ significantly from those predicted by IARC classifications ($p > 0.5$).

3. However, for the 111 chemicals considered by the EPA to lack even limited human data, but to possess animal data, EPA and IARC classifications were highly significantly different overall ($p < 0.0001$). The EPA was much more likely than IARC to assign carcinogenicity classifications indicative of greater human hazard. Overall, 67 (60.4%) of chemicals 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 lower human hazard than the corresponding IARC classification of the same chemical.

4. On the assumption that the IARC assessments are definitive, the authors comment that, based on EPA figures alone, the predictivity of animal carcinogenicity data for human hazard, and hence its utility in deriving substantially useful human carcinogenicity classifications, is questionable. However, they also note that there are deficiencies in the EPA assessments e.g. inconsistency in the scope or depth with which different assessments are conducted, a lack of criticism about the standard of data used in the assessments compared to IARC, and a tendency to err on the side of caution.