Examples of Uses of Databases for Quantitative and Qualitative Correlation Studies between Genotoxicity and Carcinogenicity

by S. Parodi,* D. Malacarne,† and M. Taningher*

In this paper we give some examples of using databases of genotoxicity and carcinogenicity for quantitative and qualitative correlation studies between short-term tests and carcinogenicity. The quality of the databases is obviously important, but one of the major deficiencies of present databases is that they are too small. Using relatively small, different databases, different results can be obtained. With small databases it is difficult to disaggregate data for homogeneous chemical classes or other types of subsets. Using the databases of Gold (carcinogenicity) and Würgler (genotoxicity), we have investigated the carcinogenic potency of genotoxic and nongenotoxic carcinogens for different chemical classes.

Introduction

Databases of genotoxicity and carcinogenicity can be used for gathering information about a specific chemical. In this case the data are used for a first-level risk analysis. The conclusions are less detailed and accurate than reading all the papers relevant to our analysis, but, it will be possible to reach an opinion in a much shorter time because the information contained in the databases is predigested. For this type of purpose, there are several requirements of the database: a) all useful data are easily accessible; b) the available database is as large and comprehensive as possible; and c) compilation and classification in the database is sufficiently detailed and accurate.

Databases of genotoxicity and carcinogenicity can be used not only for a risk analysis concerning an individual compound, but also for correlation studies between different types of short-term genotoxicity tests and long-term carcinogenicity tests for large sets of chemicals and for subsets of specific chemical classes. This type of extended use of databases also takes place in structure–activity relationship studies.

In this paper we consider the use of databases for quantitative correlation studies between carcinogenic potency and genotoxic potency in a given short-term test. Because of their log-normal distribution (1), we compare log of potencies. We have also made an attempt to compare qualitative and quantitative studies.

Past Quantitative Correlation Studies

For a reconstruction of the story of quantitative correlation studies, which to our knowledge goes back to 1977 (2), we make reference to a recent review article published by our group (3). In this review article we discussed the problems related to the computation of genotoxic and carcinogenic potencies. From the point of view of the use of databases of genotoxicity and carcinogenicity in quantitative correlation studies, we can distinguish the papers published before 1985 and the papers published after 1985. Before 1985, for each published paper the authors had generated their own database both for a given type of genotoxicity and for carcinogenic potency.

We have published several quantitative correlation studies (4–8) before 1985, and for these studies we had prepared our own database of carcinogenic potencies (9). In our database of carcinogenic potencies, the treatment of the data coming from the original publications was substantially similar to the treatment of data in the database of Gold and co-workers (10–12). The main difference between the Gold database and our database was that our database was an “ad hoc” database of 118 chemicals. These 118 chemicals were only the chemicals that we had used in correlation studies with short-term tests. After 1985, in quantitative correlation studies of our group and other authors, the investigators started to use the database of Gold (10–12) for carcinogenic potencies.

In 1990 we started the using already existing databases also for genotoxicity data. More specifically, we have used the database of Würgler and co-workers (13) for the qualitative component of the information in genotoxicity tests and the database of Waters and co-workers (14) for the quantitative component of the information, always in genotoxicity tests. A summary of the results obtained in quantitative correlation studies by our group and other authors is shown in Tables 1, 2, and 3.
As shown in Table 1, for the test of mutagenicity in Salmonella (+ microsomes), the average correlation level is about 0.4 (substantially unchanged for more than 10 years). The first value reported by Meselson and Russell (2) is artificially high because four chemicals with a poor correlation were arbitrarily discarded. In Table 2, the correlation with carcinogenic potency of different types of short-term tests related to DNA damage and repair is shown. The average level of correlation is again around 0.4. The higher level reported by Lutz (18) is probably dependent on two reasons: a) only genotoxic compounds were considered and b) genotoxicity and carcinogenicity were examined always in the same species. In Table 3, a more heterogeneous short-term tests are considered.

It is difficult to say if the correlation with carcinogenic potency of some of these short-term tests is significantly better than the correlation level found in the previous tests shown in Table 1 and in Table 2. The high correlation level obtained by Clive (20) for 25 chemicals is probably dependent on the concomitance of several factors: most of the compounds tested were potent genotoxic carcinogens; diethylstilbestrol and saccharin, which might well be nongenotoxic carcinogens, showed a good correlation between mutagenicity and carcinogenicity; and 4-acetylaminofluorene, benzo[alpyrene, and diphenylnitrosamine were negative or questionable as carcinogens: a positive potency, well correlated with mutagenicity, was partially arbitrarily given.

### New Quantitative Correlation Studies

The data reported in Table 4 were recently obtained from an intersection of the database of Waters (14) and the database of Gold (10-12). Chromosomal aberrations and chromatid exchanges in vitro were investigated. We have not separated tests with and without metabolic activation to avoid the generation of sets of chemicals that are too small.

We have compared the two cytogenetic tests with the Salmonella test. The overall quantitative predictivity of the four in vitro cytogenetic tests seems significantly better than the predictivity of the Salmonella test.

This impression is confirmed by Table 5, in which we have reported the performance of three different cytogenetic tests *in vivo*. The data have been obtained from the same intersection database (Waters + Gold) as in Table 4.

From Table 1 (obtained from the data of several authors) we concluded an average predictivity for the Salmonella test around 0.4. In Tables 4 and 5 (and also for the sister chromatid exchange data of Table 3), in eight different investigations of the predictivity of cytogenetic tests, we found that they perform better than a correlation level around 0.4, typical of the Salmonella test. The probability of obtaining this kind of result by chance is 1/2^8 = 0.004. Therefore we conclude that the cytogenetic tests are more predictive of carcinogenic potency than the Salmonella test. The
Table 4. Quantitative correlation between log₁₀(LED) or [log₁₀(HID) + 1]^* for cytogenetic in vitro tests and logTD₅₀ for carcinogenicity in rodents.

| Test | No. of chemicals (positive only) | r | No. of chemicals (positive or negative) | r |
|------|---------------------------------|---|----------------------------------------|---|
| CIC  | 28                              | 0.59* | 38                      | 0.59* |
| SIC  | 30                              | 0.76* | 37                      | 0.73* |
| CHL  | 17                              | 0.90* | 24                      | 0.78* |
| SHL  | 19                              | 0.90* | 24                      | 0.80* |
| SAL  | 54                              | 0.33* | 80                      | 0.25  |

Abbreviations: r, parametric correlation coefficient; CIC, chromosomal aberrations, Chinese hamster cells; SIC, sister chromatid exchange, Chinese hamster cells; CHL, chromosomal aberrations, human lymphocytes; SHL, sister chromatid exchange, human lymphocytes; SAL, Salmonella typhimurium, all strains, reverse mutation.

*According to the conservative hypothesis that a compound negative in the test could still be positive with a potency 10 times lower than the threshold of detectability.

*p < 0.01, one-tailed.

*p < 0.001.

Table 5. Quantitative correlation between log₁₀(LED) or [log₁₀(HID) + 1]^* for cytogenetic in vivo tests and logTD₅₀ for carcinogenicity in rodents.

| Test | No. of chemicals (positive only) | r | No. of chemicals (positive or negative) | r |
|------|---------------------------------|---|----------------------------------------|---|
| CBA  | 13                              | 0.44  | 30                      | 0.61* |
| MVM  | 20                              | 0.55* | 39                      | 0.69* |
| SVA  | 21                              | 0.63* | 29                      | 0.64* |

Abbreviations: r, parametric correlation coefficient; CBA, chromosomal aberrations, animal bone marrow cells; MVM, micronucleus test, mice; SVA, sister chromatid exchange, animal cells.

*According to the conservative hypothesis that a compound negative in the test could still be positive with a potency 10 times lower than the threshold of detectability.

*p < 0.01, one-tailed.

*p < 0.001.

strength of this conclusion is tempered by the fact that we should verify that this better quantitative predictivity is also present when exactly the same chemicals are considered for the two types of tests. Because of this dishomogeneity even the spread of the X (short-term tests) and the Y (carcinogenicity) distribution is different for different couples of parameters. These differences can affect the correlation coefficient.

If we look at a manual of statistics (21), adopting the symbol usage of these authors, then sₓₓ is the estimated variance of the random noise affecting the dependent variable (Y) in respect to the regression line; the mean of this random noise is assumed to be 0 with a normal distribution around the mean; this random noise is assumed to have the same distribution all along the regression line; sᵧ is the estimated global variance of the dependent variable Y; sₓ is the estimated global variance of the independent variable X; b is the estimate of the slope of the regression line; and r is the estimated correlation coefficient. Thus, we have (21):

\[
s_{yx}^2 = (1 - r^2)
\]

(1)

\[
b = r (s_y/s_x)
\]

(2)

and therefore, with a simple transformation:

\[
\frac{s_{yx}^2}{s_x^2} + b^2 = \frac{1}{r^2}
\]

(3)

As shown from Eq. (1), when sₓ is not much larger than sₓₓ, the correlation is low. This ratio will be close to 1 when the correlation with the independent variable is intrinsically poor.

On the other hand, as shown from Eq. (3), a sample with an unusually small spread of the independent variable will also show a poor correlation, all the other parameters remaining equal. Our correlation coefficients relative to Salmonella do not display an especially narrow spread of the independent variable (Fig. 1). We conclude that this parameter is not relevant to the lower values found for Salmonella in respect to the cytogenetic tests.

In our previous investigations on the predictivity of short-term tests, we found large variations in predictivity for different chemical classes. Usually, when data were disaggregated for chemical class, the obtained subsets were too small. Still, we have repeatedly obtained indications that dramatic variations in predictivity can take place for different chemical classes (1,4,6,17,22). This point stresses the importance of enlarging the size of the available databases. Only databases 5 to 10 times larger than the present databases could allow a reasonable analysis disaggregating different chemical classes or different significant molecular fragments. We would feel much more secure about the relevance of cytogenetic tests if our sets of data were made not of 30 to 60 chemicals, but of hundreds of chemicals. If data generated by the industry could be included in the databases, the
situation would improve significantly. Cytogenetic tests are often used in biomonitoring. It is important to have a better idea about the relevance of this type of biomonitoring.

Comparison between Quantitative and Qualitative Correlation Studies

In a recent analysis (23), Rosenkranz and collaborators have found a poor qualitative predictivity of carcinogenicity for sister chromatic exchanges (SCEs) in Chinese hamster ovary (CHO) cells. One hundred chemicals (65 carcinogens and 35 noncarcinogens) were examined in the framework of the National Toxicology Program (NTP) study. Sensitivity was 0.71 and specificity 0.31; 70.8% of the carcinogens were positive and 68.6% of the noncarcinogens were also positive. A small improvement was obtained if equivocal chemicals were considered carcinogens (sensitivity = 0.72, specificity = 0.38).

In Tables 6 and 7, we have examined our short-term tests of Tables 4 and 5 for their qualitative predictivity of carcinogenicity. Even in the intersection database of Waters and Gold, SCEs in vitro (Chinese hamster cells [SIC] and human lymphocytes [SAL]) have a low specificity, but the overall performance is a little bit better than the performance in the NTP database, which is probably due to differences of content between the two databases. What is perhaps more interesting is that, in the range of qualitative and quantitative predictivities examined, there is apparently no observable correlation between the two types of predictivities (Table 8). Not only are \( r \) and \([\text{sensitivity} + \text{specificity}]/2\) not correlated, but \( r \) and specificity do not reach a significant correlation (rank correlation analysis with the Spearman test: \( r_s = 0.43, p > 0.05, \text{one tailed} \). One reason for caution is that, for quantitative predictivity, the carcinogenicity data of the database of Gold have not been submitted to the same elaboration [usage of \((\log_{10} \text{LED}) + \log_{10} \text{HID} + 1\) values] as the genotoxicity data from the database of Waters. For the moment we have used only positive carcinogenicity data. For us this lack of correlation between qualitative and quantitative predictivity was largely unexpected, and we plan to investigate in greater detail the relationship between the two types of predictivity in the future.

### Potency of Genotoxic and Nongenotoxic Carcinogens

Another example of using the databases of Gold (10–12) and Würgler (13) is presented in Figure 2. The database of Gold et al. was used for computing the \( \log_{10} \) of carcinogenic potencies. The database of Würgler was used as a source of qualitative responses for a large set of genotoxicity tests. Our intersection database was extended using 28 additional chemicals from the NTP (all the chemicals that could satisfy our conditions). In our study we defined as genotoxic those chemicals positive in at least three short-term tests and at least 75% of the considered tests. We defined chemicals negative in at least three short-term tests and at least 75% of the considered tests as nongenotoxic.

We found 141 chemicals positive for carcinogenicity in small rodents and at the same type genotoxic or nongenotoxic according to the definition given above. In general, genotoxic carcinogens were much more potent than nongenotoxic ones; however, the difference in potency varied significantly from chemical class to chemical class.

For nitroso compounds, azo compounds, alcohols and phenols, miscellaneous compounds, and polycyclic aromatic hydrocarbons, we found that genotoxic carcinogens were about 20–100 times more potent than nongenotoxic ones. For aromatic and heterocyclic amines and amides, nitro compounds, esters and carboxamides, and hydrazine derivatives, we found that genotoxic carcinogens were 3–13 times more potent than nongenotoxic ones.

### Table 6. Qualitative concordance between different *in vitro* short-term tests and carcinogenicity.

| Test | +  | −  | Total | Sensitivity | Specificity | Accuracy |
|------|----|----|-------|-------------|-------------|----------|
| CIC  | 28 | 14 | 42    | 0.74        | 0.36        | 0.60     |
| SIC  | 30 | 16 | 46    | 0.81        | 0.38        | 0.63     |
| CHL  | 17 | 14 | 31    | 0.71        | 0.38        | 0.53     |
| SHL  | 19 | 12 | 31    | 0.79        | 0.43        | 0.62     |
| SAL  | 54 | 28 | 82    | 0.53        | 0.48        | 0.51     |
| Total| 79 | 75 | 154   | 0.68        | 0.63        | 0.66     |

### Table 7. Qualitative concordance between different *in vivo* short-term tests and carcinogenicity.

| Test | +  | −  | Total | Sensitivity | Specificity | Accuracy |
|------|----|----|-------|-------------|-------------|----------|
| CBA  | 13 | 12 | 25    | 0.45        | 0.25        | 0.38     |
| −    | 16 | 4  | 20    |             |             |          |
| MVM  | 20 | 11 | 31    | 0.53        | 0.48        | 0.51     |
| SVA  | 21 | 4  | 25    | 0.72        | 0.64        | 0.70     |

### Table 8. Lack of correlation between qualitative and quantitative parameters for measuring predictivity.*

| Test | \( r \) | \((\text{Sensitivity} + \text{specificity})/2\) |
|------|---------|----------------------------------------|
| CIC  | 0.59    | 0.55                                   |
| SIC  | 0.73    | 0.60                                   |
| CHL  | 0.78    | 0.52                                   |
| SHL  | 0.80    | 0.61                                   |
| SAL  | 0.25    | 0.66                                   |
| CBA  | 0.61    | 0.35                                   |
| MVM  | 0.69    | 0.51                                   |
| SVA  | 0.64    | 0.68                                   |

Abbreviations: CIC, chromosomal aberrations, Chinese hamster cells; SIC, sister chromatid exchange, Chinese hamster cells; CHL, chromosomal aberrations, human lymphocytes; SHL, sister chromatid exchange, human lymphocytes; SAL, *Salmonella typhimurium*, all strains, reverse mutation; CBA, chromosomal aberrations, animal bone marrow cells; MVM, micronucleus test, mice; SVA, sister chromatid exchange, animal cells.
Only for halogenated aliphatics were nongenotoxic carcinogens slightly more potent than genotoxic carcinogens. In addition, in our database, we had nine nitrogen mustards all genotoxic and four halogenated aromatics all nongenotoxic. The detailed results observed for the different chemical classes are shown in Figure 2.

We conclude that genotoxicity seems to contribute significantly to carcinogenic potency. For some chemical classes this contribution is especially strong; it is weaker for others. A genotoxic carcinogen can have both genotoxic and epigenetic activities relevant for carcinogenicity, whereas a nongenotoxic carcinogen can have only epigenetic activities.

A qualitative observation in agreement with our quantitative analysis has been reported by Gold et al. (24). These authors have observed that "more toxic carcinogens are significantly more likely to be mutagenic than less toxic carcinogen." In addition, going from a highest administered dose of less than 1 mg/kg/day to a highest administered dose of more than 1000 mg/kg/day, the fraction of carcinogens mutagenic in Salmonella decreases regularly from 71-76 %, to 28-13 % in mice and rats, respectively (L.S. Gold, personal communication).

Conclusions

In this short review we have given some examples of using databases for correlation studies and for the analysis of the carcinogenic potency of genotoxic and nongenotoxic carcinogens. In our opinion, two improvements in the existing databases would be most useful for the type of studies illustrated in this report: a) larger databases would make the conclusions that have been reached safer and more solid; b) better organization of the data would make it easier to extract subsets of data needed for a given type of correlation study. Finally, we suggest that a network of investigators interested in improving the existing databases could be very useful, for accelerated progress in this important field.

For the future, we hope that frequencies of induced mutations and rearrangements in dominant and recessive proto-oncogenes in different target organs after treatment with chemical agents, will become available (a new type of database). The frequency of these irreversible alterations in the genome are the true end point to which short-term genotoxicity tests should be compared.

Tumor frequency is a complex function of stimulations of proliferation, clonal expansions of preneoplastic cells, modulations of differentiation, and, finally, multiple irreversible alterations in the genome. Short-term tests of genotoxicity should be compared only with this last variable. New, entirely different tests are needed for assessing the other variables.

This research was supported by MURST (Progetto INTERLAB) and STEP (EEC) project 199. We are grateful to Gabriella Frigerio for her careful typing of the manuscript.

REFERENCES

1. Parodi, S., Taningher, M., and Santi, L. Utilization of the quantitative component of positive and negative results of short-term tests. Mutat. Res. 205: 283-294 (1988).

2. Meselson, M., and Russel, K. Comparison of carcinogenic and mutagenic potency. In: Origins of Human Cancer, Book C (H. H. Hiatt, J. D. Watson, and J. A. Wisten, Eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1977, pp. 1473-1481.

3. Parodi, S., Taningher, M., Romano, P., Grilli, S., and Santi, L. Mutagenic and carcinogenic potency indices and their correlation. Teratog. Carcinog. Mutagen. 10: 177-197 (1990).

4. Parodi, S., Taningher, M., Russo, P., Pala, M., Vecchio, D., Fassina, G. F., and Santi, L. Quantitative predictivity of the transformation in vitro assay compared with the Ames test. J. Toxicol. Environ. Health 12: 493-500 (1983).

5. Parodi, S., Taningher, M., Baldi, C., and Santi, L. Predictive ability of the autoradiographic repair assay in rat liver cells compared with the Ames test. J. Toxicol. Environ. Health 10: 513-539 (1982).

6. Parodi, S., Taningher, M., and Santi, L. Alkaline elution in vivo: fluorometric analysis in rats. Quantitative predictivity of carcinogenicity as compared with other short-term tests. In: Indicators of Genotoxic Exposure, Banbury Report No. 13 (B. A. Bridges, B. E. Butterworth, and I. B. Weinstein, Eds.), Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1982, pp. 137-155.

7. Parodi, S., Taningher, M., and Santi, L. Induction of preneoplastic nodules: quantitative predictivity of carcinogenicity. Anticancer Res. 3: 393-400 (1983).

8. Parodi, S., Taningher, M., Zulinio, A., Ottaggio, L., De Ferrari, M., and Santi, L. Quantitative predictivity of carcinogenicity for sister chromatid exchanges in vivo. In: Sister Chromatid Exchange: 25 Years of Experimental Research, Part A (R. R. Tice and A. Hollaender, Eds.), Plenum Press, New York, 1989, pp. 409-429.

9. Parodi, S., Taningher, M., Ugolini, F., and Santi, L. A Database of Carcinogenic Potencies (Problems Related to the Use of the Quantitative Component of the Information). Istituto Nazionale per la Ricerca sul Cancro, Genova, Italy, 1988.

10. Gold, J. S., Sawyer, C. B., Magaw, R., Backman, G. M., de Veciana, M., Levinson, R., Hooper, N. K., Havender, W. R., Bernstein, L., Peto, R., Pike, M. C., and Ames, B. N. A Carcinogenic Potency Database of the standardized results of animal bioassays. Environ. Health Perspect. 58: 9-319 (1989).

11. Gold, J. S., de Veciana, M., Backman, G. M., Magaw, R., Lopipero, P., Smith, M., Blumenthal, M., Levinson, R., Bernstein, L., and Ames, B. N. Chronological supplement to the Carcinogenic Potency Database: standardized results of animal bioassays published through December 1982. Environ. Health Perspect. 67: 161-200, 1986.

12. Gold, J. S., Slone, T. H., Backman, G. M., Magaw, R., Da Costa, M., Lopipero, P., Blumenthal, M., and Ames, B. N. Second chronological supplement to the Carcinogenic Potency Database: standardized results of animal bioassays published through December 1984 and from the National Toxicology Program through May 1986. Environ. Health Perspect. 74: 237-329 (1987).

13. Würgler, F. E. Predicting mammalian mutagenesis by submammalian assays: an application of database GEN. Environ. Health Perspect. 6: 37-39 (1991).

14. Water, M. D., Stack, H. F., Brady, A. L., Lohman, P. H. M., Haroun, L., and Vainio, H. Appendix I. Activity profiles for genetic and related tests. In: IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Genetic and Related Effects: An Updating of Selected IARC Monographs from Volumes 1 to 42, Supplement 6. International Agency for Research on Cancer, Lyon, 1987, pp. 687-696.
15. McCann, J., Gold, L. S., Horn, L., McGill, R., Graedel, T. E., and Kaldor, J. Statistical analysis of Salmonella test data and comparison to results of animal cancer test. Mutat. Res. 205: 183–195 (1988).
16. Piegorsch, W. W., and Hoel, D. G. Exploring relationships between mutagenic and carcinogenic potencies. Mutat. Res. 186: 161–175 (1988).
17. Bolognesi, C., Taningher, M., Parodi, S., and Santi, L. Quantitative predictivity of carcinogenicity of autoradiographic repair test (primary hepatocyte cultures) for a group of 80 chemicals belonging to different chemical classes. Environ. Health Perspect. 70: 247–253 (1989).
18. Lutz, W. K. Quantitative evaluation of DNA binding data for risk estimation and for classification of direct and indirect carcinogens. J. Cancer Res. Clin. Oncol. 112: 85–91 (1990).
19. Taningher, M., Saccomanno, G., Santi, L., Grilli, S., and Parodi, S. Quantitative predictivity of carcinogenicity of the covalent binding index of chemicals to DNA: comparison of the in vivo and in vitro assays. Environ. Health Perspect. 84: 183–192 (1990).
20. Clive, D., Johnson, K. O., Spector, J. F. S., Batson, A. G., and Brown, M. M. Validation and characterization of the L5178Y/TK mouse lymphoma mutagen assay system. Mutat. Res. 59: 91–108 (1979).
21. Snedecor, G. W., and Cochran, W. G. Statistical Methods, 6th ed. The Iowa State University Press, Ames, IA, 1967.
22. Parodi, S., Taningher, M., and Santi, L. DNA fragmentation: its predictivity as a short term test. In: Chemical Carcinogenesis (C. Nicolini, Ed.), Plenum Publishing Corp. 1982, pp. 121–154.
23. Rosenkranz, H. S., Ennever, F. K., and Klopman, G. Relationship between carcinogenicity in rodents and the induction of sister chromatid exchanges and chromosomal aberrations in Chinese hamster ovary cells. Mutagenesis 5: 559–571 (1990).
24. Gold, L. S., Bernstein, L., Magaw, R., and Stone, T. H. Interspecies extrapolation in carcinogenesis: prediction between rats and mice. Environ. Health Perspect. 81: 211–219 (1989).