Arh. hig. rada toksikol., **30** *(1979) suppl. pp. 15–33.*

PROCEDURES FOR SCREENING CHEMICALS FOR CARCINOGENICITY

I.F.H. PURCHASE

Imperial Chemical Industries Limited, Central Toxicology Laboratory, Alderley Park, Macclesfield, Cheshire, United Kingdom

ABSTRACT

A major aspect of concern about the effects of chemicals on man centres on their carcinogenicity. By definition, human carcinogens can only be identified by epidemiological studies in man, a procedure that is manifestly too late to prevent suffering. Attention has thus been concentrated on developing laboratory methods based on parameters other than human cancer, which are nevertheless capable of identifying chemicals with biological activity classifying them as candidate human carcinogens.

Laboratory models of this type have certain characteristics that define their performance and thus their utility in identifying carcinogens and their potential effects in man. Predictivity is one such characteristic and is normally defined in terms of the correlation between carcinogenicity and activity as defined in the model system. Animal carcinogens can be identified with about 90% accuracy by using short-term tests such as the Ames *Salmonella* mutation assay and the Styles cell transformation assay. About 85% of chemicals that are carcinogenic in one rodent species are detected by a second species. Predicting the carcinogenic effects of chemicals in man on the basis of results from short-term tests and animal studies is thus likely to be subject to similar error rates.

A second characteristic is the reliability of the model. Much is now known about the reliability of animal models, with factors such as diet, food consumption, species, strain, environmental conditions, age, bedding, virus infection and route of exposure having been identified as having a profound influence on tumour incidence and thus the outcome of the test. As the short-term tests are used more widely, factors such as bacterial strain or cell type, incubation time and temperature, conditions of exposure of the organism, and parameters of the metabolising system (species, strain, organ, co-factor ratios, enzyme induction) are being found to influence the outcome of the test. The short-term tests have accelerated the speed with which these parameters can be investigated. The result is that an understanding of the effect of the various parameters on the test result and of their relevance to human carcinogenicity (in both animal and short-term models) is required before an assessment of the relevance of the test outcome can be made.

A third useful characteristic is the ability to assess potential potency in man. Animal models have not been generally useful in defining potency of most biological effects (e.g. antibacterial activity) and the same is generally true for carcinogenicity. The problem using short-term tests has only recently been tackled and, owing to the ease with which the magnitude of responses in these tests can be manipulated (i.e. a lack of ruggedness), general correlations seem unlikely. Other characteristics, such as organ specificity, are desirable but not yet feasible. The cost and time are characteristics which may influence the feasibility of carrying out a test, but should not influence interpretation of the results.

The confidence with which results from animal and short-term tests can be extrapolated to human carcinogenicity will depend on, amongst other things, the predictivity and reliability of the model. The probability that the results from these tests are relevant to man may also depend on the

"ruggedness" of the results: that is, the degree to which a test result can withstand manipulation of the conditions of testing. Similarly, the pattern of results from a range of test systems may also be important. Thus, the range of test systems in which positive results are found and the ruggedness of the result in each test system are measures of the relevance of the results to man.

Industrial societies have developed a way of life that is increasingly dependent on artefacts produced from chemicals. The benefits that man has derived from the use of these artefacts are abundant and include the control of disease, increase in agricultural production and improvement in the quality of life. However, it is only in recent years that the price that society has paid in terms of damage to human health has become the subject of concern, as is evidenced by the large number of chemicals of unknown biological activity that were introduced into manufacture during the early decades of this century.

A major aspect of concern about the effects of chemicals on man centres on their carcinogenicity. By definition, human carcinogens can only be identified by epidemiological studies in man, a procedure that is manifestly too late to prevent suffering. Attention has thus been concentrated on developing laboratory methods based on parameters other than human cancer, which are nevertheless capable of identifying chemicals with biological activity classifying them as candidate human carcinogens.

These laboratory models include both *in vitro* or "short-term" tests for carcinogenicity and the more conventional chronic animal studies.

Laboratory models of this type have certain characteristics that define their performance and thus their utility in identifying carcinogens and their potential effects in man. This paper discusses the predictivity and reliability of short-term tests and their ability to assess carcinogenic potency. Similar characteristics for animal carcinogenicity studies are also discussed.

SHORT-TERM TESTS

During the course of the last few years, a large variety of short-term tests has been described. The majority of these are based on mutation as an end-point and have contributed substantially to the theory that cancer is caused by somatic mutation. Others, however, are based on a variety of end-points (some of which may also be due to mutation) including biochemical changes, behaviour of tissues exposed to carcinogens, chemical reactivity and alterations in behaviour of cells in culture. In most cases there are few data on which to judge the merits of the tests, but the *Salmonella* plate test[3] and the cell transformation assay[40] are amongst the most extensively studied.

The *Salmonella* plate test, which is very extensively used for screening for carcinogens, is based on the induction of mutations in strains of *Salmonella typhi murium* using a system which incorporates liver enzymes to metabolise the chemical under test (Fig. 1). The cell transformation assay uses BHK21 cells in culture which are 'transformed', allowing growth in agar, by carcinogens. A metabolic system is also incorporated (Fig. 2).

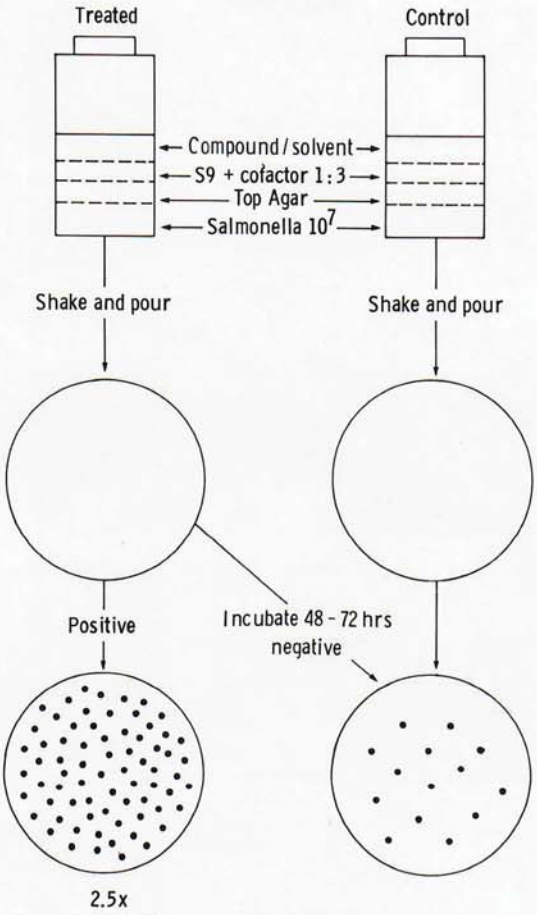

FIG. 1 – *Salmonella* plate test

The performance of these tests, in terms of their ability to discriminate between carcinogens and non-carcinogens, has been reported from several studies. McCann and colleagues[29], using data from a number of laboratories, found that 90% of carcinogens and 87% of non-carcinogens were correctly identified by the *Salmonella* plate test. Purchase and colleagues[34,35], in a controlled blind study, and Sugimura and colleagues[6], reported similar results (Table 1). There are some chemicals which have been tested twice in these studies and in 12 cases inconsistencies between the results from the McCann and colleagues[29] and Purchase and colleagues[34,35] studies occur. Nevertheless, about 88% of the carcinogens and 84% of the non-carcinogens were correctly identified.

Only one validation study of the Styles cell transformation assay has been carried out[34,35], where 91% of the carcinogens and 97% of the non-carcinogens were correctly identified.

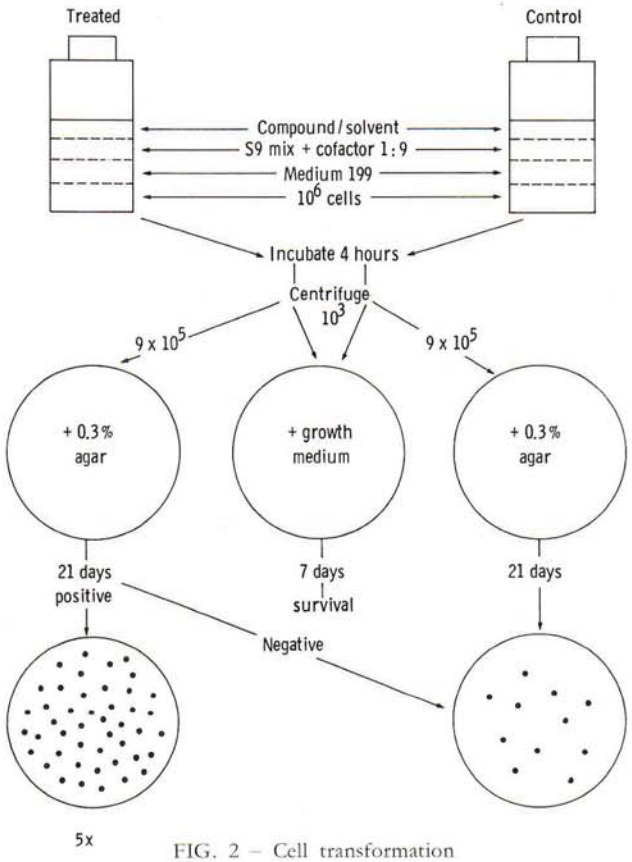

FIG. 2 – Cell transformation

TABLE 1

Results from validation studies of the *Salmonella* plate test.

Reference	Carcinogens			Non-carcinogens		
	Tested No.	Identified No.	%	Tested No.	Identified No.	%
McCann *et al.*[(a)29]	174	156	90	108	94	87
Purchase *et al.*[(b)35]	58	53	91	62	58	93
Sugimura *et al.*[(c,d)41]	39	25	64	54	47	87
Sugimura *et al.*[(e)41]	121	111	92	27	13	48

(a) using up to seven strains (TA98, 100, 1530, 1535, 1138 and his G46)[29]
(b) using four strains (TA98, 100, 1535 and 1538)[34]
(c) using two strains (TA98, 100)
(d) from Table 5[41]
(e) from Table 6[41]

Predictivity

The validation studies suggest that when groups of chemicals are tested there will be a definite error in extrapolation rate (expressed as false negative and false positive results from the short-term test). The size of this error rate is influenced strongly by the classes of chemicals being tested and many other factors, for example, test conditions. In the extreme example, if a group of non--carcinogens was tested, all positive results would be false. The relationship between the accuracy of the *Salmonella* plate and cell transformation tests and the composition of the sample of chemicals being tested can be calculated from the data of Purchase and colleagues[34,35], both for positive and negative results (Fig. 3). The more carcinogens there are in the sample, the better the prediction from a positive test result becomes.

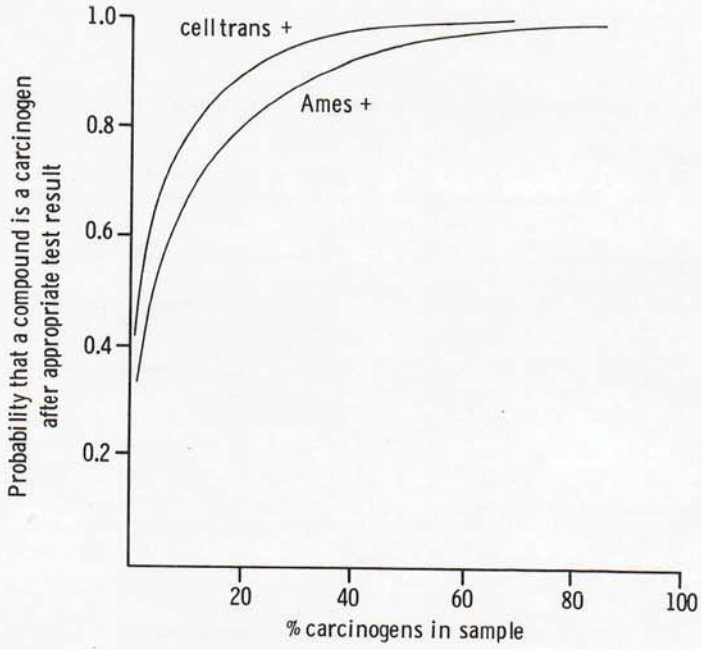

FIG. 3 – The probability that a compound with a positive test result will be carcinogenic. The graphs are drawn from data derived from Purchase *et al.*[35] for the cell transformation and *Salmonella* plate (Ames) assay.

One implication of this observation is that it would be possible to obtain virtually any result from a validation study by carefully selecting the compounds to be tested. Accuracy of 100% could be obtained using a group of carcinogens drawn from the highly active polycyclic, nitroso and alkylating carcinogens and a group of non-carcinogens from insoluble or inactive chemicals which are part of the test organism's metabolic pathway. Very low levels of accuracy would be

obtained by selecting carcinogens, such as diethylstilbestrol, or using chemicals such as dieldrin, DDT and phenobarbitone which are negative in the *in vitro* test. Classification of the chemicals into broad chemical classes (such as alkylating agents, polycyclics and arylamines[35]) suggested that these short *in vitro* tests were equally accurate for these three chemical classes. However, this classification was so broad that it obscured the fact that the tests are highly dependent on the structural class of the chemical being tested. Thus, although in the validation studies[29,35] pairs of chemicals from eleven classes were correctly identified by the *Salmonella* plate test (Table 2), there are several other examples where chemicals were not correctly identified (Table 3). Subsequently, other chemicals have been identified, such as saccharin[10], hexamethylphosphoramide[11] and dimethylnitrosamine (which is negative in the standard *Salmonella* test, although positive when incubated before plating[13]) which are negative in the *Salmonella* plate test but of varying potency in animal carcinogenicity studies.

TABLE 2
Non-carcinogens and their carcinogenic analogues which were correctly identified by the *Salmonella* plate test.

Non-carcinogen	Carcinogenic analogue
Ref. 35	
3-methyl-4-nitroquinoline-N-oxide	4-nitroquinoline-N-oxide
3,3',5,5'-tetramethylbenzidine	benzidine
4-acetylaminofluorene[35], fluorene[41]	2-acetylaminofluorene
anthracene	9,10-dimethylanthracene
dimethylformamide	dimethylcarbamoyl chloride
1,3-dinitrobenzene	1-fluoro-2,4-dinitrobenzene
α-naphthylamine	β-naphthylamine
diphenylnitrosamine	nitrosofolic acid
Ref. 3	
15,16-dihydro-3-methyl-cyclopenta (a)phenanthrene-17-one	15,16-dihydro-11-methyl-cyclo-penta(a)phenanthrene-17-one
Ref. 41	
pyrene	benzo(a)pyrene
stilbene	4-dimethylaminostilbene

At first sight it may seem surprising that these short-term tests which have an "accuracy" of 90%, do not identify so many carcinogens (Table 3). However, the 90% figure can be considered to be composed of 90% of compounds in which the accuracy is 100% and 10% of compounds in which the accuracy is 0%. This is reflected in the graphs describing the relationship between "accuracy" and proportion of carcinogens in the sample tested (Fig. 3). Both considerations suggest that it is important to decide whether the short-term test selected is sensitive to the chemical class under test. This can be approached by structural analogy and the use of chemical class control pairs, as is discussed later.

TABLE 3
Some carcinogens reported negative in the *Salmonella* plate test during the validation studies.

Diethylstilbestrol	Urethane
Ethionine	1'-Hydroxysafrole
Carbon tetrachloride	Dieldrin
3-Amino-1,2,4-triazole	DDE
Thioacetamide	2,7-Diaminofluorene
Thiourea	PCB
Natulan	Testosterone propionate
Dimethylhydrazine	3-Hydroxyanthranilic acid

Reliability

There have been few systematic studies of the reliability (defined as the qualitative and quantitative reproducibility) of short-term tests, although there are data from many laboratories on the results of testing chemicals under different conditions. It is already clear that the results of short-term tests are sensitive to alterations in minor details of the test procedures. This is important because it can result in differences in test results on the same compound between different laboratories. An understanding of which factors are most sensitive to change will enable a test protocol to be designed which is much more 'rugged'. When that is achieved, reproducibility of the results both within and between laboratories will be improved.

There are, however, two important consequences of considerations of the reliability or ruggedness of short-term tests. The first, that of extrapolation of results to other species, is shared by chronic animal studies, and is discussed later. The second refers to the choice which the experimentalist must make in the selection of the protocol. Should a rigidly defined protocol always be used even if minor modifications would, for some chemicals, result in a different outcome? In the case of the *Salmonella* plate test, the validation is based on certain fixed protocols. It could be claimed therefore that variations in the protocol invalidates the extrapolation to other species. This would impose an unwelcomed rigidity on the protocol and, as the examples given below illustrate, is not always a correct view. It is important that future work should be directed to defining just how important these changes in protocol are and how large the changes are that can be tolerated. The following discussion highlights features critical to the reliability of short-term tests.

Target cell

Assay system. Hexamethylphosphoramide is a potent rat carcinogen producing tumours of the nasal sinuses in rats exposed to 20 ppb by inhalation[25]. As this was the first indication that this chemical was carcinogenic, it immediately raised the question of which of its analogues would also be carcinogenic. Repeated testing using the *Salmonella* plate test resulted in a variable response, positive results only being obtained in about 30% of the tests[11,12]. The response in the cell transformation assay was much more consistent and, as a consequence, it was possible to establish that the structural

feature responsible for the carcinogenic activity was probably related to the N--dimethyl-amino function. The reason for the variable results from the *Salmonella* plate test has not been established but could relate to the differences in the metabolism or the target cells.

Bacterial strain. The strains used in the *Salmonella* plate test are sensitive to chemicals which induce either base pair substitutions (TA100, TA1535) or frame shift mutations (TA98, TA1538). Certain chemicals specifically produce only one type of lesion and these are used in routine testing to confirm the identity of the strains. Other chemicals produce both types of lesion, but to a different degree. Thus, two of the metabolites of benz(a)anthracene produce positive results of variable intensity in TA100 and TA98[28]. The importance of this is that the relative effects of these metabolites are different, being 1 : 3 in TA100 and 1 : 32 in TA98, thus affecting the quantitative extrapolation of the results.

Variation in details of testing protocol

Most of the details in this discussion refer to the *Salmonella* plate test and some have been discussed previously by Ashby and Styles[7,9].

Duration of exposure. The standard *Salmonella* plate test procedure is to add the chemical to the test medium and incubate for 72 hours. The chemical will be in contact with the *Salmonella* for up to 72 hours, the exact period depending on stability, rate of metabolism and other factors. With volatile gases this period appears to be particularly important. Vinyl chloride, added in solution to the *Salmonella* plate and cell transformation test, gave negative results although it was positive when tested as a gas[35]. The reason was probably that in the former case exposure was too brief and too low to produce an effect. Recently, fluorocarbon F22 was found to be negative when exposure was six hours, but positive when exposure was 24–72 hours[27].

Test protocol. The standard *Salmonella* plate test is the plate incorporation assay, where the chemical, S9 mix and bacteria are mixed in top agar and then poured on to a plate. Variations in this procedure, for example by preincubation of the chemical, S9 mix and bacteria before pouring the plate, can alter the response of the test[13]. Nitrosamines are more easily detected by this method and the response to benzo(a)pyrene metabolites is halved[28]. The use of the fluctuation technique will greatly increase the sensitivity of bacterial mutation tests in some cases[20].

Number of cells. When the number of cells plated is altered, the mutant yield also changes, but not in proportion to the change in cell number. Thus, a reduction in the number of *Salmonella* added to the plate to 10% reduces the mutant yield produced by N-nitrosomorpholine, N-nitroso-N'-methylpiperazine, N-nitrosopiperidine and N-nitrosopyrrolidine to 65, 55, 32 and 25% respectively of the values obtained with 10^8 bacteria[13].

Oxygen content. Assays are normally conducted under aerobic conditions but different results can be obtained by using anaerobic incubation. Azathioprine,

negative in the standard assay, is positive if incubated for 14 hours anaerobically[39]. Para-phenylenediamine is positive in the *Salmonella* plate test when it is added to the incubation mixture dissolved in DMSO and incubated aerobically. A negative result is obtained if the incubation is anaerobic[18].

The medium. Variations in the buffer and nucleophil (cysteine[33] or serum[28]) content of the medium can alter the response of the test.

Solvent. Various solvents have been shown to alter the response in the *Salmonella* plate test[4]. The use of DMSO as a solvent greatly enhances the positive response of some chemicals. Para-phenylenediamine is negative when tested dissolved in water, but is positive when dissolved in DMSO[18].

Background mutation frequency. A positive result from a bacterial mutation test is judged either in terms of the number of revertants per μmole of compound or as the increase in number of revertants expressed as the multiple of the background frequency. In the latter case a compound is considered positive if there is a greater than 2.5-fold increase in colonies. In both cases, the background frequency is crucial. Considerable variations in background frequency, which are currently unexplained, provide one of the largest variables in the interpretation of plate incorporation assays.

Metabolic activation system

The S9 mix currently employed is somewhat arbitrary and it is not surprising that it is a principal, if not the major, source of variability in the response of those chemicals which require metabolic activation. The following are among the factors so far identified.

Inducers. 'Aroclor'1254, 3-methylcholanthrene and phenobarbitone are the commonly used inducing agents. Their effects on the enzymes induced in the liver, or other organs, differ and hence they alter the response observed in these tests[43].

Diet. Rats fed a lipotrope-deficient diet were used to provide liver for preparation of S9 mix[42] which resulted in a reduction in the colonies produced by aflatoxin B_1 in the *Salmonella* plate test. Even when mixed with normal S9, a reduced response was observed, in spite of the fact that rats on lipotrope--deficient diets are more susceptible to the effects of aflatoxin B_1 carcinogenesis.

Organ. Different organs have different activities and specificities of metabolising enzymes and this is reflected in the variable response observed when S9 is prepared from various organs[17,38].

Species. Acetylaminofluorene is claimed to be non-carcinogenic in cotton rats, but is a potent carcinogen in rats. S9 mix derived from cotton rats is, however, much more efficient in producing a positive response than that from rats[30]. This effect has also been observed in S9 derived rom rats and mice[14,23] with vinylidene dichloride.

Proportions of S9 mix. There is a variation in the response of the *Salmonella* plate test when the ratio of S9 mix to co-factor is altered[1,13]. This is well exemplified by varying the amount of S9 in the mix from 25 μl to 150 μl which resulted in a 50-fold increase in response to acetylaminofluorene[30]. Mixing S9 derived from different species, which activates butter yellow to different extents, results in a modulation of the response[8].

Glutathione, which is present in liver homogenates and is normally associated with detoxification, was found to enhance greatly the response of the *Salmonella* plate test to 1,2-dichlorethane. The mechanism for this is probably the formation of the half mustard with subsequent alkylation of DNA[36].

Adequacy of enzymes in S9 mix. Addition of epoxide hydratase or inhibition of it either decreases or increases the response of the *Salmonella* plate test to benzo(a)pyrene as the amount of epoxide formed changes[33]. This observation highlights the sensitivity of the S9 mix to variations in critical enzyme activity. An extension of this is the absence of certain gut enzymes from the S9 which are responsible for metabolising chemicals (such as cycasin) into proximate carcinogens (in that case methylazoxymethanol), thus producing a negative result.

Mixtures of chemicals

The interplay of toxicity and mutagenicity when two or more chemicals are added to the *Salmonella* plate test can be difficult to predict and provide inaccurate results.

Enhanced response. The addition of the non-mutagenic norharman to the medium resulted in a marked increase in the response[32]. Similarly, azobenzene will greatly stimulate the response of butter yellow[8], presumably by inhibiting or saturating detoxication pathways.

Another way in which enhanced responses are seen is when chemicals give positive results because of potent mutagenic impurities[19]. An example of this is the positive result obtainable with trichloroethylene, caused by the presence of stabilisers, in the *Salmonella* plate assay[21].

Diminished response. If a mixture containing a very toxic chemical is tested in an *in vitro* test, the toxicity can obscure the positive result (Table 4). Cysteine, when added to the culture medium, can reduce or abolish the positive response of a number of otherwise mutagenic chemicals[37].

Ashby and Styles[7]describe a number of changes in the response of the *Salmonella* plate test (e.g. changes in response to safrole, dimethylnitrosamine and hexamethylphosphoramide) which are as yet unexplained.

The phenomenon of 'phenotypic reversion' is another example of an unexplained result[26]. In the standard plate-incorporation assay[3], the colonies growing on the plate 72 hours after adding the chemical are counted as 'revertants'. It is assumed that their growth represents a genetic or genotypic alteration in the *Salmonella* and therefore that the bacteria would continue to

grow if plated on to histidine-deficient medium. If the colonies observed at 72 hours will not grow when plated out, alteration to the genotype is not demonstrated and it is described as phenotypic reversion. As the plating out of revertant bacteria is not a standard practice, false positive results can easily be recorded.

TABLE 4
Results from testing a mixture and its fractions in the *Salmonella* plate test. The figures are the number of revertants per plate relative to the control.

Sample		TA1538	TA98
Mixture		2.5	1.2
Fraction	1	10.9	4.5
	2	10.1	3.0
	3	10.8	2.9
	4	0.9	1.1
	5	1.2	0.8
	6	1.8	1.4
	7	0.9	1.4
Control		1.0	1.0
DMSO		1.4	2.5

Potency

The *in vitro* short-term tests obtain much of their advantage over conventional carcinogenicity studies by their relative brevity and simplicity. Although many qualitative aspects of short-term test metabolic systems are similar to those of intact mammals, there is ample evidence (reviewed above) of sources of variability in the short-term tests that derive directly from their simplicity and many substantial differences from mammalian systems are known. The suggestion by Messelson and Russel[31] and Ames and Hooper[2] that a quantitative relationship exists between the potency of chemicals to the *Salmonella* plate test and their potency in animals has been discussed by Ashby and Styles[7]. The fact that so many factors can profoundly affect the response of a chemical in the *Salmonella* plate test or other *in vitro* tests suggests that the response in the short-term tests will form a poor basis on which to develop potency correlations. These major differences in metabolic competence would be expected to produce a different range and proportion of metabolites in the *in vitro* system when compared with *in vivo* metabolism[15,28].

If the results from the short-term tests are to be used for potency correlations, the conditions under which the correlations are made will have to be extremely carefully defined and tested and there is a strong likelihood that any such correlation would be restricted to within certain limited chemical classes.

False results

The validation studies indicate that there is a certain percentage of chemicals which are incorrectly identified by the short-term tests. Careful consideration of these chemicals suggests that they can be grouped into several types of false negative and positive results. (In the context of this discussion, it is assumed that a reproducible result is obtained from the short-term assay and the categorisation of this as a false result occurs which it is compared with *in vivo* carcinogenicity data.)

False negative results (or negative results on mammalian carcinogens)

"Technical" false negative. In this case some hitherto unsuspected technical artefact of the protocol prevents a positive result being observed. Two examples of this have already been given, viz. vinyl chloride when administered as a solution[35] and fluorocarbon 22 when administered for only six hours[27]. Other, as yet undefined, causes for technical false negatives may explain the results from such chemicals as hexamethylphosphoramide.

"Mechanism" false negatives. As more data on the mode of action of carcinogens are produced, it is becoming clear that many of them are carcinogenic as a result of one of a number of mechanisms other than mutation. These 'epigenetic' carcinogens will not always be detected by a mutation-based assay[10]. Examples include chemicals which act by interfering with the hormone balance (diethylstilbestrol, oestrogens, goitrogens and antithyroid agents), those which produce tumours as a secondary response to tissue damage (alcohol--induced liver tumours in man, diethylene glycol-induced bladder tumours in rats[44]) and a variety of others where the mechanism of action is unknown (chloroform, carbon tetrachloride, dieldrin, phenobarbitone and saccharin).

"Activation" false negative. As has been pointed out in the discussion on potency, the metabolic competence of the S9 mix may differ considerably from the situation in the whole animal and the spatial relationship between the site of production of active metabolites and the molecular target may be very different. A simple example of this is cycasin which is carcinogenic in rats but is negative in the *Salmonella* plate test. The reason for this is that it requires β-glucosidase from gut bacteria to produce the active metabolite, methylazoxymethanol, which is the proximate carcinogen. Although cycasin is converted to the active metabolite in rats this does not occur with the standard S9 mix, but requires addition of β--glucosidase. Another example of a carcinogenic compound which is often negative, but with a putative metabolite which is positive in the *Salmonella* plate test, is safrole.

False positive results (or positive results on mammalian non-carcinogens)

"Technical" false positive. Any characteristic of the test protocol which results in a positive may be described as an artefact or technical false positive. A rather facetious example would be histidine or any histidine analogue which supported

growth of the indicator organism. The more serious example is that provided by phenotypic reversion mentioned earlier.

"Mechanism" false positives. The simplicity of the *in vitro* models in comparison with the *in vivo* situation suggests that there may be chemicals positive in the *in vitro* tests which do not produce results *in vivo,* because, for example, immunological surveillance or DNA repair reverse the lesion. I know of no example of this type.

"Activation" false positives. There are two distinct mechanisms by which the metabolic system may produce false positive results. In the first, the chemical, or its metabolite, is so reactive that it irreversibly binds to other macromolecules *in vivo* before the ultimate metabolite can react with the DNA. Examples include dichlorvos, which is so rapidly hydrolysed that no systemic genetic effects are produced. A parallel example (which is not entirely relevant) is β-propiolactone which is positive in microbial mutagenicity assays, but does not procude systemic genetic effects[16], although it is carcinogenic at the site of application. The second type occurs when a chemical is metabolised to an electrophilic product not seen *in vivo* and not producing carcinogenic effects. The 5,6-epoxide of 7,12--dimethylbenz(a)anthracene is positive in the *Salmonella* plate test but probably non-carcinogenic. Similarly, the 4,5-epoxide of benzo(a)pyrene is strongly positive in TA1537 but does not produce skin tumours in mice[22]. (±)Trans-7β,8--dihydrodiol-9α,10α-epoxy-7,8,9,10-tetrahydro benzo(a)pyrene is another positive material which is inactive as a complete carcinogen or initiator[24] in mouse skin (although it produces benign pulmonary tumours in newborn mice). An excess of nitroreductases in *Salmonella typhi murium* which are less active *in vivo* may also provide false positive results for some nitrofurans[3].

"Incorrect" false positives. In many cases animal carcinogenicity data are of such a poor standard that the classification of the chemical as a non-carcinogen is incorrect. Thus, trimethylphosphate, positive in the *Salmonella* plate test, was considered non-carcinogenic in our validation study[35], but subsequently there is some evidence that it may be a carcinogen.

Strategies for minimising false results

The intelligent application of knowledge of the structural characteristics of chemical carcinogens and the likely metabolic processes which they might undergo is an integral component of short-term testing. A comprehensive review of the subject[5] has been made. It is noteworthy that the boundaries of the structural characteristics with which the carcinogenic activity can be associated can afford to be fairly wide, since short-term tests can be used to test suspect structures. Any serious discrepancy between the structural characteristics and activity in the test should be a signal for more detailed testing to be carried out. It is interesting to note that this use of structural analogy can be seen as a subjective means of altering the proportion of carcinogens and non-carcinogens in the sample being tested.

In situations when short-term tests[35] are being applied to chemicals from, or related to, a class of known carcinogens or non-carcinogens, the accuracy of the result can be improved by the use of chemical class control pairs[6]. The recognition that there are a large number of variables which can profoundly alter the quantitative and qualitative results from short-term tests, and that the precise importance of any one variable is difficult to assess, is further support for the idea that chemically related compounds of known carcinogenic activity should be used for controlling results. There are eleven examples of such pairs of compounds from the validation studies (Sec Table 2) and others are available in certain chemical classes.

An extension of these ideas is that the test to be used in evaluating potential carcinogenicity should be carefully chosen for its ability to discriminate between carcinogenicity and non-carcinogenicity in the class under study. In many cases, structurally related compounds with known activity may not be available and in that case the use of two tests, for example the *Salmonella* plate and cell transformation, will improve the chances of correctly identifying carcinogenic activity[35].

ANIMAL STUDIES

Carcinogenicity studies using life-time administration to animals are cumbersome and expensive, but they are, nevertheless, the commonest method for screening chemicals prior to widespread human exposure. In spite of the use of animal studies for over 40 years, there does not appear to be a systematic study of the comparability of results from different species.

Interspecies correlation

In one sense the term "validation" when applied to short-term tests refers to an interspecies comparison of biological activity. With the short-term tests the comparison is between activity with *Salmonella* or BHK21 cells on the one hand, and carcinogenicity in any mammalian species on the other. A similar, although phylogenetically much closer, comparison can be made between, say, carcinogenicity in the mouse and that in the rat.

The sources of data for making such a comparison are necessarily dependent on the scientific literature. Three convenient sources have been used, each in a slightly different way. Firstly, the NCI bioassay programme provides data on the carcinogenicity of chemicals administered to rats and mice by the same route and following similar protocols. These studies give possibly the best comparative data on carcinogenicity and the conclusions are published, which means that for comparative purposes it removes the need for the reviewer to judge the adequacy of the studies. The results from the first 30 compounds are given in Table 5. The second source of both data and a considered evaluation of the carcinogenicity of compounds is the IARC Monograph Series, "The Evaluation of Carcinogenic Risk to Man" and a summary of opinions given is in Table 5. (Where the Monograph did not provide an opinion on the carcinogenicity of a compound in at least two species, that compound was omitted). The problem with that series is

TABLE 5

Positive or negative results obtained in animal carcinogenicity studies. The numbers refer to the number of compounds which have been tested in at least two species with the result specified at the head of the column.

	Positive in all	Mixed result	Negative in all
NCI	8	11	11
IARC	60	15	8
PHS	23	7	64
PHS + IARC	83	22	72
Totals	91	33	83

NCI = National Cancer Institute, Bioassay Programme
IARC = International Agency for Research into Cancer Monograph Series, "Assessment of Carcinogenic Risk to Man" Vol 1–17
PHS = National Cancer Institute publication 156, "Survey of Compounds which have been tested for Carcinogenicity"

that the compounds have been selected because of suspicion of carcinogenicity; hence a requirement for a broader selection led to the use of the third source, "Survey of Compounds which have been tested for Carcinogenicity". Compounds were selected, using the index, which had data reported in at least two species. The original publications were then consulted and only those which were considered adequate accepted.

Data on carcinogenicity were available on 207 compounds which had been tested in two species (usually rat and mouse). Of these 92 compounds were positive in both species, 83 negative in all species and 33 had differences between species (Table 5). Thus, a compound which was carcinogenic in one species had about an 80% chance of being carcinogenic in the second. It should be stressed that these percentage figures are based on historical data and their use for future test results may present some problems. Nevertheless, the historical figures for extrapolation are remarkably similar to the figures produced from validation studies on the short-term tests. The latter may be numerically greater because the variation in animal carcinogenicity data is excluded from that comparison. The problems associated with extrapolation from *in vitro* tests to mammalian species, such as how to define the appropriate test protocol, difficulties with reproducibility and potency correlations, and the probability that the results are relevant to man, will also apply for inter-mammalian species extrapolation.

Reliability

The sources of variability in animal studies include route of administration, dose and distribution of dose, sex, strain, age and species of mammal, food consumption, dietary composition, environmental conditions, access to bedding,

hormonal status, intercurrent disease, duration of study and dietary contaminants. These factors can influence the qualitative and quantitative response of the animal to the compound under test.

CONCLUSION

Whatever model is used as a screening procedure for carcinogenicity, it is apparent that there are shortcomings and problems associated with its use. The purpose of using screening tests is to predict potential human carcinogenicity before exposure of man results in the development of cancer. The results of validation of short-term tests and of correlations between mammalian carcinogenicity data from different species lead to the conclusion that a positive result in any single test system, in the presence of negative results from several other tests, gives a lower predictivity of activity in man. In the case of short-term tests, this predictivity is considerably enhanced when two test systems are used and the results are the same. Thus, positive data in more than one test system provide an increased confidence in the predictivity of the result.

As we have seen, the qualitative and quantitative results from *in vitro* and *in vivo* tests are subject to considerable change when certain parameters of the test are altered. In some cases the conditions of the test may have to be very precisely determined before a positive result is obtained. In others, such as in the case of powerful alkylating agents in short-term tests or nitrosamines in mammals, the conditions of testing are not critical, a positive being obtained very easily. This variability in the "ruggedness" of the data should be used in extrapolation to man. Aflatoxin B_1, which is positive in many *in vitro* tests and carcinogenic to rats, pigs, trout, ducks and other species, would thus have a high probability of being carcinogenic to man. Trichloroethylene, which is negative in a variety of *in vitro* tests, is non-carcinogenic in rats. The liver tumours produced in mice (if they are not due to stabilisers) must be considered with the negative data. It seems much less likely that it will be carcinogenic in an untested species (i.e. man).

Ultimately the only way of demonstrating that a chemical is a human carcinogen is by observing occupationally-associated increase in cancer incidence in working populations. Epidemiological evidence of this type would obviously take precedence over negative results from laboratory studies. The converse, however, does not give such a clear-cut conclusion. Negative epidemiology is often not sufficiently convincing to define non-carcinogenicity of a chemical, because of the numerous difficulties inherent in epidemiological studies of working populations (such as the length of the observation period and the difficulties in obtaining a sufficiently large group of workers with complete follow-up and records of exposure levels).

Negative studies are particularly unconvincing in the face of positive data from *in vitro* or conventional laboratory studies, even though, as explained above, direct extrapolation from laboratory models to man is bound to contain inaccuracies. The use of several laboratory systems allows a qualitative

assessment of the accuracy of the extrapolation. When all systems agree, greater confidence can be placed in the extrapolation than when there are substantial disagreements between the results from different laboratory models; in the latter case accurate extrapolation will have to await an understanding of why the differences occur and which result is most relevant to man.

REFERENCES

1. *Ames, B. N , Durston, W. E., Yamasaki, E. and Lee, F. D.* Carcinogens and Mutagens: a single test system combining liver homogenates for activation and bacteria for detection. Proc. Natl. Acad. Sci., USA, **70** (1973) 2281–2285.
2. *Ames, B. N. and Hooper, K.* Does carcinogenic potency correlate with mutagenic potency in the Ames assay? Nature, **274** (1978) 19–20.
3. *Ames, B. N., McCann, J. and Yamaski, E.* Methods for detecting carcinogens and mutagens with the *Salmonella* mammalian-microsome mutagenicity test. Mutat. Res., **31** (1975) 347–364.
4. *Anderson, D.* Personal communication.
5. *Ashby, J.* Appendix 1 to reference (35).
6. *Ashby, J and Purchase, I. F. H.* The selection of appropriate chemical class controls for use with short-term tests for potential carcinogenicity. Ann. Occup. Hyg., **20** (1977) 297–301.
7. *Ashby, J. and Styles, J. A.* Factors influencing mutagenic potency *in vitro*. Nature, **274** (1978) 20–22.
8. *Ashby, J. and Styles, J. A. In vitro* evaluation of some derivatives of the carcinogen Butter Yellow: implications for environmental screening. Br. J. Cancer, **38** (1978) 34–50.
9. *Ashby, J. and Styles, J. A.* Does carcinogenic potency correlate with mutagenic potency in Ames assay? Nature, **271** (1978) 452–455.
10. *Ashby, J., Styles, J. A., Anderson, D. and Paton, D.* Saccharin. Possible example of an epigenetic carcinogen mutagen. Fd. Cosmet. Toxicol., **16** (1978) 95–103.
11. *Ashby, J., Styles, J. A. and Anderson, D. A.* Selection of an *in vitro* carcinogenicity test for derivatives of the carcinogen hexamethylphosphoramide. Br. J. Cancer., **36** (1977) 564–571.
12. *Ashby, J., Styles, J. A. and Paton, D.* Potentially carcinogenic analogues of the carcinogen hexamethylphosphoramide (HMPA). Br. J. Cancer, **38** (1978) 418–427.
13. *Bartsch, H., Camus, A. and Malaveille, C.* Comparative mutagenicity of N-nitrosamines in a semi-solid and in a liquid incubation system in the presence of rat or human tissue fractions. Mutat. Res., **37** (1976) 149–162.
14. *Bartsch, H., Malaveille, C., Montesano, R. and Tomatis, L.* Tissue-mediated mutagenicity of vinylidene chloride and 2-chlorobutadiene in *Salmonella typhi murium*. Nature, **255** (1975) 641–643.
15. *Bigger, C. A. H., Tomaszewski, J. E. and Dipple, A.* Differences between products of binding of 7,12-dimethylbenz(a)-anthracene to DNA in mouse skin and in a rat liver microsomal system. Biochem. Biophys. Res. Commun., **80** (1978) 229–235.
16. *Brusick, D. J.* The genetic properties of beta-propiolactone. Mutat. Res., **39** (1977) 241–255.
17. *Brusick, D., Jagannath, D. and Weekes, U.* The utilization of *in vitro* mutagenesis techniques to explain strain, age and sex-related differences in dimethylnitrosamine tumour susceptibilities in mice. Mutat. Res., **41** (1976) 51–60.
18. *Burnett, C.* Personal communication.
19. *Donahue, E. V., McCann, J and Ames, B. N.* Detection of mutagenic impurities in carcinogens and non-carcinogens by high pressure liquid chromatography and the *Salmonella* microsome test. Cancer Res., **38** (1978) 431–438.

20. *Green, M. H. L., Bridges, B. A., Rogers, A. M., Horspool, G., Muriel. J. W., Bridges, J. W. and Fry, J. R.* Mutagen screening by a simplified bacterial fluctuation test: use of microsomal preparations and whole liver cells for metabolic activation. Mutat. Res., **48** (1977) 287–293.

21. *Henschler, D., Eder, E., Neudecker, T. and Metzler, M.* Carcinogenicity of trichloroethylene: fact or artifact? Arch. Toxicol., **37** (1977) 233–236.

22. *Jerina, D. M., Yagi, H., Hernandez, O., Dansette, P. M., Wood, A. W., Levin, W., Chang, R. L., Wislocki, P. G. and Conney, A. H.* In: Carcinogenesis, vol 1, Polynuclear aromatic hydrocarbons: chemistry, metabolism and carcinogenesis, eds Freudenthal, R. I. and Jones, P. W., Raven Press, New York, 1976, p. 91–113.

23. *Jones, B. and Hathway, D. E.* Tissue-mediated mutagenicity of vinylidene chloride in *Salmonella typhi murium* TA1535. Cancer Letters, **5** (1978) 1–6.

24. *Kapitulnik, J., Wislocki, P. G., Levin, W., Yagi, H., Jerina, D. M. and Conney, A. H.* Tumorigenicity studies with diol-epoxides of benzo(a)pyrene which indicate that (±)-trans-7β, 8α-dihydroxy-9α, 10α-epoxy-7,8,9,10-tetrahydrobenzo(a)pyrene is an ultimate carcinogen in newborn mice. Cancer Res., **38** (1978) 354–358.

25. *Lee, K. P., Trochimowicz, H. J. and Sarver, J. W.* Induction of nasal tumours in rats exposed to hexamethylphosphoramide. Lab. Invest., **36** (1977) 344–345.

26. *Longstaff, E.* Personal communication.

27. *Longstaff, E. and McGregor, D.* Mutagenicity of a halocarbon refrigerant monochlorodifluoromethane (R-22) in *Salmonella typhi murium*. Toxicol. Letters, **2** (1978) 1–4.

28. *Malaveille, C., Kuroki, T., Sims, P., Grover, P. L. and Bartsch, H.* Mutagenicity of isomeric diolepoxides of benzo(a)pyrene and benz(a)anthracene in *S. typhi murium* TA98 and TA100 and in V79 Chinese hamster cells. Mutat. Res., **44** (1977) 313–326.

29. *McCann, J., Choi, E., Yamaski, E. and Ames, B. N.* Detection of carcinogens as mutagens in the *Salmonella* microsome test. Part I, assay of 300 chemicals. Proc. Natl. Acad. Sci. USA, **72** (1975) 5135–5139.

30. *McGregor, D. B.* Cotton rat anomaly. Nature, **274** (1978) 21.

31. *Messelson, M. and Russell, K.* Comparisons of carcinogenic and mutagenic potency. In: Origins of human cancer, eds Hiatt, H. H., Watson, J. D. and Winsten, J. A., Cold Spring Harbor Laboratory (1977) p.1473–1481.

32. *Nagao, M., Yahagi, T., Honda, M., Seino, Y., Kawachi, T., Sugimura, T., Wakabayashi, K., Tsuji, K. and Kosuge, T.* Comutagenic actions of norharman derivatives with 4-4-dimethylaminoazobenzene and related compounds. Cancer Letters, **3** (1977) 339–346.

33. *Oesch, F., Bentley, P. and Glatt, H. R.* Prevention of benzo(a)pyrene-induced mutagenicity by homogenous epoxide hydratase. Int. J. Cancer, **18** (1976) 448–452.

34. *Purchase, I. F. H., Longstaff, E., Ashby, J., Styles, J. A., Anderson, D., Lefevre, P. A. and Westwood, F. R.* Evaluation of six short-term tests for detecting organic chemical carcinogens and recommendations for their use. Nature, **264** (1976) 624–627.

35. *Purchase, I. F. H., Longstaff, E., Ashby, J., Styles, J. A., Anderson, D., Lefevre, P. A. and Westwood, F. R.* An evaluation of six short-term tests for detecting organic chemical carcinogens. Br. J. Cancer, **37** (1978) 873–903.

36. *Rannug, U., Sundvall, A. and Ramel, C.* The mutagenic effect of 1,2-dichloroethane on *Salmonella typhi murium* I. activation through conjugation with glutathione *in vitro*. Chem. Biol. Interact., **20** (1978) 1–16.

37. *Rosin, M. P. and Stich, H. F.* The inhibitory effect of cysteine on the mutagenic activities of several carcinogens. Mutat. Res., **54** (1978) 73–81.

38. *Sehgal, C. B., Hutton, J. J.* Homogenates of pregnant rat and fetal tissues metabolize carcinogens to mutagens detected by *Salmonella typhi murium* TA98 and TA100. Mutat. Res., **46** (1977) 325–344.

39. *Speck, W. T. and Rosenkranz, H. S.* Mutagenicity of azathioprine. Cancer Res., **36** (1976) 108.

40. *Styles, J. A.* A method for detecting carcinogenic organic chemicals using mammalian cells in culture. Br. J. Cancer, **36** (1977) 558–563.

41. *Sugimura, T., Kawachi, T., Matsuschima, R., Nagao, M., Sato, S. and Yahagi, T.* A critical review of submammalian systems for mutagen detection. In: Progress in Genetic Toxicology, eds. Scott, D., Bridges, B. A. and Sobels, F. H., Elsevier North Holland, Amsterdam, 1977, p. 126–154.

42. *Suit, J. L., Rogers, A. E., Jetten, M. E. R. and Luria, S. E.* Effects of diet on conversion of aflatoxin B_1 to bacterial mutagen(s) by rats *in vivo* and by rat hepatic microsomes *in vitro*. Mutat. Res., **46** (1977) 313–323.

43. *Suzuki, H., Kishida, F. and Miyamoto, J.* Induction of hepatic drug metabolizing enzymes and mutagen-activation capacity in several animal species. J. Pesticide Sci., **2** (1977) 421–425.

44. *Weil, C. S., Carpenter, C. P. and Smyth Jr, H. F.* Urinary bladder calculus and tumor response following either repeated feeding of diethylene glycol or calcium oxalate stone implantation Ind. Med. Surg., **36** (1967) 55–57.