4.2.3. Toxicological properties/health effects

4.2.3.1. Absorption

The likely extent of absorption of a chemical via skin, lungs and gastro-intestinal tract is predicted by the EPA experts on the basis of the physico-chemical properties of the chemical (particularly log P_{ow}, which is usually a predicted value, and the physical form of the chemical). The initial opinion on this basis may be modified in the light of any available test data on the chemical itself or on a closely related structural analogue. Good, moderate, poor or no absorption will be predicted for each route of exposure (dermal, inhalation and oral).

The prediction of the likely extent of absorption following exposure by a particular route will be used when taking decisions on whether the chemical may present an unreasonable risk to human health and/or on testing requirements in the USA.

Absorption is not investigated in the base-set level testing in the EC, but whether any absorption has occurred can be inferred to an extent from evidence of systemic toxicity in the acute and repeated dose studies. It is less easy to decide that absorption has not occurred - the chemical may be well absorbed and show no systemic toxicity in the particular test(s) already conducted. However, it may cause adverse effects in other test systems not yet applied. Evidence of absorption (i.e. systemic effects) may have an influence on the timing of further testing. When there is no evidence from the currently available test data, the timing of further testing may be influenced by the likelihood of absorption based on the physico-chemical properties of the chemical and/or the extent of human exposure expected.

- Conclusions

There were too few studies conducted using the inhalation route for an accurate assessment of concurrence between SAR calls for absorption from the lungs and derived absorption estimates from toxicity test results.

Based on the 136 chemicals for which dermal toxicity studies were available, it is considered that acute dermal studies are inadequate to judge dermal absorption. There were too few 28-day studies to serve as a basis for definitive judgement on dermal absorption calls.

The SAR calls for gastro-intestinal absorption were essentially in agreement with estimates based on the oral toxicity test results: when they differed it was only in degree of absorption and not, with one exception, giving a completely different assessment of whether or not a chemical was absorbed at all. For some chemicals, which were classified in the EC on the basis of their oral toxicity, the relatively low extent of absorption predicted may be of some concern. However, none of these chemicals were predicted to have "no absorption" and thus would not have been dropped from EPA evaluations.

4.2.3.2. Acute toxicity

Acute toxicity data are used to predict the potential effects in humans of a single exposure to a chemical (e.g. during maintenance work or in an accident). They are also used to help in setting dose levels for other toxicity tests.

Prediction of acute toxicity is not emphasised in the EPA evaluation of a new chemical which focuses on long-term or sub-chronic effects. For the purposes of this project, however, predictions of acute toxicity following oral administration were made. (There were too few chemicals with data from inhalation or dermal acute toxicity tests which were suitable for conducting comparisons of the two approaches to evaluation.)

The following criteria were used to rank chemicals on the basis of their oral LD50 values, and so provide a means of comparing the predicted toxicity with that observed in the tests:

Oral LD50 (mg/kg)	Toxicity
>2000	Low (L)
>1000 <2000	Low-Medium (L-M)
> 500 <1000	Medium (M)
> 50 < 500	Medium-High (M-H)
< 50	High (H)

These criteria give more categories of acute toxicity than are conferred by the EC classification system (below), but the same criterion (LD50 >2000 mg/kg) is used to differentiate chemicals of low concern with regard to acute oral toxicity from those of some level of concern.

Oral LD50 (mg/kg)	EC classification
>2000	Not classified
> 200 <2000	Harmful
> 25 < 200	Toxic
< 25	Very toxic

Acute oral toxicity tests had been conducted on 142 chemicals (two chemicals had not been tested: chemicals 4 and 107 are corrosive and react violently with water). A prediction of acute oral toxicity had been made for all of the 142 chemicals which had been tested, plus the two which had not.

There were 21 chemicals for which the toxicity indicated by the test data differed from that predicted (15%). Twenty of these were found to have greater acute toxicity than had been predicted, but for fourteen of these there was overlap between the predicted and observed toxicity categories, (see Table 10). One chemical had lower toxicity than had been predicted (number 124).

Twenty-one chemicals had been classified in the EC on the basis of their acute oral toxicity: twenty of them are included in Table 10 and were predicted to have lower toxicity than was observed, though for 14 there was an overlap between predicted and observed toxicity categories. However, 18 of the classified chemicals (12%) were predicted to be of "low" acute oral toxicity, and thus would apparently be considered of low concern with regard to this end-point (false negatives). The classified chemical which is not in Table 10 (number 281) was predicted, by analogy to data in the EPA confidential data base, to have "medium" acute toxicity and this was observed (LD50 = 850 mg/kg). Details of the oral toxicity predictions and test results are given for all chemicals in the project in Annex 12.

- Conclusions

Using arbitrary criteria to compare LD50 values with descriptions of predicted acute oral toxicity, there was a tendency to under-prediction of the level of toxicity for chemicals which, when tested, were shown to have significant acute oral toxicity. However, the majority of the chemicals were correctly predicted to be of low concern with regard to acute oral toxicity.

Predicted toxicity for 18 (12%) of the classified chemicals was "low", indicating that one-to-one substitution of predictive methods for testing would result in chemicals being missed which are, in fact, of some potential concern because of acute toxicity. It should be noted that two of these chemicals had been classified as "Toxic if swallowed" (numbers 307 and 330).

In most cases there were overlaps between the predicted and the observed toxicity for the classified chemicals, and between the toxicity predicted for the classified chemicals and those not classified. Hence, the predictive methods could not readily be used to classify chemicals within the context of a scheme using pre-defined criteria.

Thus, this comparative study shows that the predictive methods can be used to identify correctly the >80% of a batch of 142 heterogeneous new chemicals which are of low acute toxicity. However, it is of concern that some 12% of this set of chemicals did have an appreciable level of acute oral toxicity which was not predicted (false negatives). Because of this outcome, if assessment of acute toxicity is an important consideration in a given evaluation scheme, the submission of test data will be needed to assess this end-point adequately. This is especially so in instances where a quantitative assessment of acute toxicity is needed.

TABLE 10: Differences between SAR evaluations and acute oral toxicity test data

Chemical	LD50	Label[1]	MPD tox[2]	SAR tox[2]
47	1800	R22	L-M	L
49	>200 <2000	R22	L-M	L
54	1984	R22	L-M	L
124	2300	-	L	M
156	1800M 1960F	R22	L-M	L
197	612	R22	M	L
219	1670	R22	L-M	L
241	585	R22	M	L
242	520	R22	M	L
300	1011	R22	L-M	L
307	88	R25	M-H	L
312	1774	R22	L-M	L
330	104	R25	M-H	L
340	1750	R22	L-M	L
360	>1000 <2000	R22	L-M	L
370	1400	R22	L-M	L
413	1200	R22	L-M	L
425	1650	R22	L-M	L
436	899	R22	M	L
441	450	R22	M-H	M
443	320	R22	M-H	M

[1] See Appendix 3 for list of "R phrases".
[2] See abbreviations above.

4.2.3.3. Irritation

Knowledge of the potential for skin, eye and respiratory irritation is important when evaluating safe handling practices for chemicals. Skin and eye irritation test data are used to predict the likelihood that exposure of human skin or eyes to a chemical will result in adverse effects (corrosion or irritation). An indication of the duration/reversibility of effects is also usually obtained.

There is not a test method for respiratory irritation in either the EC or the OECD set of accepted test methods for the toxicity testing of chemicals.

Prediction of irritation is not usually part of the routine evaluation of new chemicals in the US, but predictions were made for the purposes of this project, although EPA did not attempt to characterise the degree of irritation.

4.2.3.3.i Skin irritation

The criteria used for conducting the comparisons were to compute "primary irritation scores" from the test data, by taking the average of the total erythema and oedema scores for both the 24 and 72 hour readings:

Primary irritation index	Irritant category
2 or less	Mild/nil (low)
>2 to 5	Moderate
>6	Severe

The category "corrosive" was also used.

In addition, chemicals were also considered according to whether they had been classified as "Corrosive" or "Irritating to skin" in the EC.

Of the total of 144 chemicals in the project, there were 140 on which skin irritation tests had been conducted. All 144 chemicals had been considered when predicting the potential for skin irritation as a consequence of dermal exposure to the chemicals.

Correct predictions of low concern for skin irritation were made for 104 of the 122 chemicals (including the untested polymer, chemical number 267) for which the test results indicated little or no irritancy (83% of the 122 chemicals; 73% of the total number of chemicals in the project). There were 18 chemicals which were predicted to be irritating to skin, but were found not to be irritant in the test conducted, i.e. false positives.

The test results (or physico-chemical characteristics of three chemicals: numbers 4, 107 and 194) showed that 22 chemicals either were, or could be expected to be, at least moderate skin irritants. Twelve of these had been classified as "Corrosive" in the EC, and six as "Irritating to skin". The outcome of the comparisons for the classified chemicals is shown in Table 11. Ten of these were identified by EPA as being skin irritants, while for the remaining 8, EPA did not identify a concern for skin irritation (false negatives). The group of false negatives included six corrosive chemicals.

TABLE 11: Comparison of predicted skin irritancy with that observed

Chemical	Label[1]	MPD result[2]	SAR result	Agreement[3]
4	R35	Corrosive[4]	Acute	Yes
49	R34	Corrosive	Irritant	Yes
53	R38	Mod - Sev	Irritant	Yes
107	R35	Corrosive[4]	Acute	Yes
118	R34	Corrosive	No comment	False -ve
182	R34	Corrosive	No comment	False -ve
192	R34	Corrosive	No comment	False -ve
194	R34	Corrosive[4]	No comment	False -ve
222	R38	Moderate	Irritant	Yes
235	R34	Corrosive	No comment	False -ve
237	R38	Low - Mod	Irritant	Yes
278	R38	Moderate	Irritant	Yes
370	R34	Corrosive	Irritant	Yes
373	R38	Moderate	No comment	False -ve
425	R34	Corrosive	Irritant	Yes
436	R34	Corrosive	No comment	False -ve
437	R38	Mod - Sev	Irritant	Yes
443	R34	Corrosive	No comment	False -ve

[1] See Appendix 3 for list of "R phrases".
[2] According to the criteria above, using primary irritation score.
[3] Predicted relative to test-derived level of skin irritancy.
[4] Chemicals not tested: EC assumed corrosivity based on physico-chemical properties.

The overall results for the comparison of SAR calls and MPD data for skin irritation are summarised in Table 12. In this Table, MPD positive includes the three chemicals considered corrosive in the EC on the basis of physico-chemical properties (chemicals 4, 107 and 194); and SAR negative includes the two chemicals for which the prediction was "uncertain". Details of the data on skin irritation for all chemicals are to be found in Annex 13.

TABLE 12: Overall results for skin irritation

	SAR Positive	SAR Negative
MPD Positive	14 (10%)	8 (5.5%)
MPD Negative	18 (12.5%)	104 (72%)

- Conclusions - skin irritation

Incorrect predictions were obtained for 18% of the chemicals: 12.5% were false positives and 5.5% were false negatives. The predictive methods used are not adequate for classification of chemicals using a system based on severity of response and thus the test cannot be replaced on a one-to-one basis by the predictive approach when knowledge of the potential for skin irritation/corrosion is needed.

4.2.3.3.ii Eye irritation

The criteria used to compare the test data with the SAR call for eye irritation could not be made on a severity index as the SAR evaluations did not usually include this index. From the test data summaries, a chemical was considered to produce significant eye irritation if redness, swelling or corneal opacity persisted beyond seven days or if effects were not reversible by 21 days or corrosion was reported. Eye testing was not conducted on chemicals with predictable corrosivity because of their physico-chemical characteristics or, for some chemicals (see Table 13), if corrosive effects had been recorded in a previously conducted skin test.

Classification according to the EC system (for which the criteria are a combination of scores and duration of effects), on the basis of the results of the eye irritation studies, was obviously also considered as indicating that the classified chemicals were eye irritants.

Of the total of 144 chemicals in the project, there were 140 on which eye irritation tests had been conducted, three were predicted to be corrosive and one (number 267) could not be tested for technical reasons. All 144 chemicals had been considered when predicting the potential for eye irritation as a consequence of ocular exposure to the chemicals.

On the basis of the test results, 105 chemicals were considered to be of low concern for eye irritation, as was chemical 267, which had not been tested. Correct predictions of low concern were made for 87 of these (83% of the "negative " chemicals, 60% of the total set of chemicals). The other 18 were predicted by the EPA to be irritant i.e. they were false positives.

The 38 remaining chemicals were either corrosive (12 chemicals), or irritant according to the criteria given above. The outcome of the comparisons between the predicted and test results for the classified chemicals is given in Table 13, the detailed analysis for all chemicals in the project is given in Annex 13.

TABLE 13: Comparison of predicted eye irritancy with that observed

Chemical	Label[1]	MPD result[2]	SAR result	Agreement[3]
4	R35	Corrosive[4]	Acute	Yes
47	R41	Severe	Uncertain	False -ve
49	R34	Corrosive	Irritant	Low
87	R41	Severe	No comment	False -ve
107	R35	Corrosive[4]	Acute	Yes
118	R34	Corrosive[4]	No comment	False -ve
124	R36	Irritant	Irritant	Yes
151	R36	Irritant	Irritant	Yes
170	R41	Severe	Irritant	Yes
182	R34	Corrosive	Irritant	Low
192	R34	Corrosive	No comment	False -ve
194	R34	Corrosive[4]	No comment	False -ve
197	R41	Severe	No comment	False -ve
222	R36	Irritant	Irritant	Yes
235	R34	Corrosive[4]	No comment	False -ve
237	R36	Irritant	Irritant	Yes
256	R36	Irritant	Irritant	Yes
263	R36	Irritant	Irritant	Yes
270	R36	Irritant	No comment	False -ve
281	R36	Irritant	Irritant	Yes
370	R34	Corrosive[4]	Irritant	Low
425	R34	Corrosive[4]	Irritant	Low

TABLE 13 - continued

Chemical	Label[1]	MPD result[2]	SAR result	Agreement[3]
436	R34	Corrosive[4]	No comment	False -ve
441	R41	Severe	Irritant	Yes
442	R41	Severe	Irritant	Yes
443	R34	Corrosive[4]	No comment	False -ve

[1] See Appendix 3 for list of "R phrases"

[2] According to the criteria given in the text

[3] Predicted relative to test-derived result

[4] Chemicals not tested: corrosivity assumed based on physico-chemical properties or results of skin irritation study

From the comparisons given in Table 13, it can be seen that, for the 26 classified chemicals, 16 were correctly predicted to be eye irritants and 10 were incorrectly assessed (false negatives).

The overall results for the comparison of the SAR calls and the MPD test results are summarised in Table 14.

TABLE 14: Overall results for eye irritation

	SAR Positive	SAR Negative
MPD Positive	26 (18%)	13 (9%)
MPD Negative	18 (13%)	87 (60%)

Conclusions - eye irritation

Incorrect predictions were made for 22% of the chemicals (9% were false negatives, 13% false positives). As with skin irritation, predictive methods are not adequate for classification of chemicals with regard to severity of the response and thus cannot replace test results on a one-to-one basis.

4.2.3.3.iii Respiratory irritation

New chemicals are not tested for respiratory irritation in the EC, but the potential for respiratory respiration had been considered by the EPA predictors.

Predictions of potential respiratory or mucous membrane irritation had been made for 9 (6%) of the chemicals in this study.

- General conclusions

The majority of this group of new chemicals was of low concern for skin (85%) and eye (74%) irritancy. Thus, the extent to which an assessment can be made of the power of the predictive methods to discriminate between chemicals on the basis of their skin or eye irritation potential is limited.

The majority (>80%) of the low concern chemicals were predicted correctly and 18% were over-predicted for either or both of skin and eye irritancy. The latter observation means that for these substances, the risk assessment would err on the side of caution but would lead to "over-labelling" if the predictive methods replaced the tests.

The incidence of false negatives and the limitations in assessing severity of response are of some concern and indicate that replacement of testing with prediction cannot yet be recommended with confidence.

Respiratory irritation is an important end-point which is not investigated in the MPD. It would be prudent to take note of chemicals predicted to be respiratory irritants.

4.2.3.4. Sensitisation

Knowledge of the sensitising potential of chemicals is important when evaluating safe handling practices.

Prediction of sensitisation is not usually part of the routine evaluation of a new chemical in the US, but it was considered for this project.

In the EC, chemicals are tested for their skin sensitising potential. There is not an internationally recognised test method for respiratory sensitisation. Classification of notified new chemicals as skin sensitisers in the EC is based on the proportion of animals showing a positive response in a particular test. In the EC, chemicals may be classified as respiratory sensitisers if they show close structural similarity to known chemical respiratory sensitisers.

Skin sensitisation tests, mostly maximisation tests, were conducted on 137 of the chemicals in the project. Twenty eight chemicals were classified as skin sensitisers (including one of those which had not been tested). A further 18 induced some positive responses but the number of animals responding was below the threshold for classification in the EC.

Seventeen chemicals were predicted to be sensitisers; four of these were predicted to be respiratory sensitisers and one was predicted to be a photosensitiser. Two were predicted not to be sensitisers. For most of the chemicals there was no comment on skin sensitisation - this is equivalent to considering the chemical of low concern/negative for this end-point.

For 108 chemicals (75% of the whole set in the project), both the test results and the predictions indicated low concern for skin sensitisation.

The results of the comparisons of the test data and the predictions are given in Table 15 for the 28 chemicals classified as skin sensitisers in the EC.

TABLE 15: Comparison of results for chemicals classified as skin sensitisers

Chemical	SAR	Result and comments
47	-	False negative
76	+	Agree
96	-	False negative
118	-	False negative
133	-	False negative
173	+	Agree
194	-	False negative NB: chemical not tested
196	+	Agree
197	-	False negative
200	-	False negative
222	+	Agree Chemical also classified and predicted as a respiratory sensitiser
235	-	False negative
256	+	Agree
271	-	False negative
275	-	False negative

TABLE 15 - continued

Chemical	SAR	Result and comments
330	+	Agree
341	+	Agree Chemical also classified and predicted as a respiratory sensitiser
344	-	False negative
348	-	False negative
376	+	Agree
393	-	False negative
401	-	False negative
413	-	False negative
416	-	False negative
437	-	False negative
442	-	False negative
444	-	False negative

Five other chemicals were predicted by the US to be skin sensitisers: one did not have adequate test data (240); two did induce some positive responses in the tests conducted (253, 312) and two were apparently false positive predictions (340, 364).

Two other chemicals were predicted to be potential respiratory sensitisers (69, 101).

For the set of comparable skin sensitisation data (140 chemicals) the comparisons in Table 16 can be made.

TABLE 16: Overall results for skin sensitisation

	SAR Positive	SAR Negative
MPD Positive	9 (6.5%)	19 (13.5%)
MPD Negative	4*(3%)	108 (77%)

* includes two substances for which positive responses, below the threshold for classification, where observed in the tests

Conclusions

The incidence of false negatives precludes use of the predictive methods to replace the tests on a one-to-one basis or to classify chemicals for their skin sensitisation potential. However, the concurrence of positive predictions with positive test results needs to be further assessed with a larger set of chemicals as confidence in the ability to predict positives could perhaps replace testing of chemicals predicted to be skin sensitisers.

For respiratory sensitisation, reliance is currently placed on predictive methods, based on structure, to classify new chemicals in the EC, and the unclassified substances predicted, in this project, to be potential respiratory sensitisers should be re-evaluated in the EC with regard to classification.

It is not possible to comment on the single prediction of potential photosensitisation.

4.2.3.5. Repeated dose toxicity

Repeated dose toxicity covers the adverse effects which may arise in humans exposed to a chemical at frequent, regular intervals over a prolonged period of time, for example at their daily work. To facilitate evaluation of safe handling practices for chemicals, it is important to have knowledge of the potential systemic effects which may occur on repeated exposure.

In the EC, general effects on the whole animal and effects on tissues, organs and/or systems are investigated. Special effects (e.g. neurotoxicity, reproductive toxicity, carcinogenicity) are investigated in specific tests, but indications of potential reproductive toxicity, neurotoxicity or immunotoxicity may be detected in repeated dose toxicity studies.

For most of the chemicals in this project only 28-day, and/or occasionally 90-day, study results were available. In the EC study summaries used for this project, dose levels used, a description of toxic signs, including clinical chemistry and haematology, gross and microscopic changes in a selected set of tissues/organs, and NOEL, NOAEL, LOEL and LOAEL (no/low observed effect/adverse effect level) values are usually included or can be deduced. In general, only effects of biological significance are included and species specific effects (e.g. peroxisome proliferation and, in the more recent summaries, male rat specific light hydrocarbon nephropathy) are not. Chemicals are classified for repeated dose toxicity in the EC on the basis of adverse effects (of biological/human significance) occurring at or below dose levels specified according to the route of exposure and the duration of the study.

Predictions of repeated dose toxicity are particularly important in the US EPA evaluation process, with identification of potentially toxic chemicals as the goal. Efforts are also made to assess potential target tissue/organ/system.

Test data were not available for seven chemicals (3 corrosive chemicals, 2 polymers, 1 organoclay and one chemical not tested in the light of test data available for another notified chemical, of very similar structure). Two chemicals had been tested in 28-day inhalation studies and eight in dermal 28-day studies. For one of the latter group, a 90-day study had also been conducted. The remaining 127 chemicals had been tested using 28-day oral toxicity studies and three also had results available from 90-day studies.

Eight chemicals had been classified in the EC on the basis of their repeated dose toxicity.

The comparison of repeated dose toxicity test results with predicted toxicity was the most difficult to do as interpretation of observed effects in terms of severity and significance is a matter of professional judgement. The factors considered in the evaluation were the perceived seriousness of the toxic effect, the number of organ-specific parameters affected, with microscopic pathology given the heaviest weight, multiplicity of target organs, the toxic effect(s) at the LOAEL, the numerical value of the NOAEL, dose-related effects and the spacing of the dose levels used.

The systemic toxicity data from the test results were scored as high, moderate or low using the following general criteria (sometimes modified according to professional judgement):

Concern level	Criteria
Low (L)	No systemic toxicity (NOAEL 1 g/kg/day or more); only minor clinical signs of toxicity; liver and/or kidney weight increase or clinical chemistry changes; LOAEL >500 mg/kg/day.
Moderate (M)	Organ pathology (gross and/or microscopic) with LOAEL 500 mg/kg/day or less; clinical chemistry changes and organ weight changes at <500 mg/kg/day; NOAEL <100 mg/kg/day.
High (H)	Death, organ pathology (microscopic) at LOAEL 100 mg/kg/day or less; multiple organ toxicity; NOAEL <10 mg/kg/day.

"Split-levels" (L-M; M-H) were adjustments for specific multiple organ toxicity, borderline effect levels and professional judgement.

The outcome of the comparisons of repeated dose toxicity on the basis of concern level is summarised in Table 17.

TABLE 17: Matrix analysis of systemic toxicity concern levels

SAR	L	L-M	M	M-H	H
MPD					
L	62	10	5	0	0
L-M	23	11	2	0	0
M	11	1	5	1	1
M-H	3	1	2	3*	0
H	1	0	0	0	1*

* 1 chemical in each of these groups was corrosive and predicted to have acute effects

One chemical (337) is not included in the matrix. It was M-H according to the test results, but there was no prediction of repeated toxicity.

Sixty-two chemicals (43%) were considered of low concern both following testing and by the predictive methods.

Twenty chemicals (14%) with greater than "low" concern were predicted to have the same level of concern as was deduced from the test data using the criteria given above. This group included the two corrosive chemicals which were predicted to have "acute" effects (numbers 4 and 107) and chemical 292 for which data were available from the product literature.

The concern level was under-predicted for 42 chemicals (29%) though for 27 chemicals there were overlapping concern levels from the test and predicted results; and 23 of these predicted to be of low concern were only low-moderate from test results. For the other 15 the concern level predicted was at least one whole level lower than that deduced from the test data. Six of this sub-set of 15 were chemicals classified in the EC on the basis of repeated dose toxicity.

Toxicity concern was apparently over-predicted for 19 chemicals. However, the extent of repeated dose toxicity testing of these chemicals was limited to 28-day studies (18 oral studies, 1 dermal). It will be of interest, if/when 90-day, or longer, study data become available, to re-compare the predicted toxicity with that found on testing.

Overall, the correct level of concern (according to the criteria given above) was predicted for 57% of the chemicals, but was under-predicted for 29%. Toxicity was apparently over-predicted for 13% of the chemicals.

Details of the organ toxicity predictions and test results are given for all the chemicals in the project in Annex 14.

Conclusions

Just over half (57%) of this group of 143 heterogeneous chemicals were correctly predicted to be either of low concern (43% of the total) or to have the same level of concern (14% of total) in relation to repeated dose systemic toxicity. The concern level was apparently over-predicted for a further 13%, but if/when longer-term studies are conducted the predicted effects may be induced.

Under-prediction of the level of concern on the basis of repeated dose toxicity was noted for 42 chemicals (29% of the total), although for 23 of these, the test data indicated only low-moderate concern and EPA predicted low concern. For 15 chemicals, there was at least one whole "level of concern" difference, and six of the eight classified chemicals were in this group.

On the basis of these comparisons, although for 74% of the chemicals in this study, correct or near-correct predictions of concern level were made, it is not considered possible to consider the predictive methods as an adequate substitute for conducting repeated dose toxicity testing of a random/heterogeneous group of chemicals because of under-prediction of toxicity. As classification of a chemical as dangerous following repeated exposure depends not only on the effects seen, but also on the doses at which they occur, the predictive methods for repeated dose toxicity would not provide a firm basis for classification.

4.2.3.6 Mutagenicity

Chemicals which increase the incidence of mutations in the cells of exposed humans may thereby increase the incidence of cancer (from mutations in somatic cells) or genetic defects in the offspring (from mutations in germ cells). It is generally thought prudent to assume that there is no threshold exposure level, below which exposure would give rise to only low concern, for chemical mutagens. Thus, chemicals identified as mutagens are subject to stringent controls so that human exposure is minimised.

Because of the serious and irreversible effects which may occur in humans exposed to chemical mutagens, testing for mutagenicity usually employs a number of tests, in vitro and in vivo, which are conducted either as a battery or (as in the EC) in series. In the EC, all notified chemicals must, if it is technically possible, be tested in a bacteriological test for gene mutation and in a test in mammalian cells for chromosomal effects at the "base-set" level of supply. The latter test may be either an in vitro test or a test conducted in vivo. Maximised conditions are used, though short of conditions likely to cause artefactual positive results; and in vitro tests are conducted both with and without exogenous metabolic activation. Further testing is conducted to investigate in more detail positive test results, as necessary, and/or as supply tonnages reach the trigger levels. Classification of chemicals on the basis of mutagenicity is done according to criteria defined in Annex VI to the dangerous substances Directive. Chemicals are not usually classified unless there is evidence of mutagenicity from tests conducted in vivo, so positive in vitro test data will trigger the need for testing in vivo.

The EPA predictions for mutagenicity, based on e.g. chemical class, analogue data, likely metabolites, alkylating potential, represent an overall for mutagenic potential. EPA also considers available data concerning mutagenicity test systems and their sensitivity towards different classes of chemicals. Thus, the criteria for comparing the predicted with the test results involved more than a simple comparison of EPA predictions with the test data. In addition, the test results for a few (6, 4%) chemicals with borderline responses were not always interpreted in the same way by the EPA and EC experts.

Tests had not been conducted on five of the 144 chemicals in the project - 3 for technical reasons (chemicals 4 and 107 were corrosive and chemicals 267 was an insoluble polymer) and for the other two (chemicals 194 and 445) data from analogues were considered acceptable. Predictions had been made for the first three (all were "low concern" for mutagenicity) but there were no test data to compare them with. Thus, there were 141 data pairs for comparison. All of the 139 chemicals tested had Ames test data and all had at least a result from one other study. The in vivo micronucleus test occurred most frequently as the second study, and the in vitro chromosome aberration test was next most common. Tests in E coli (always alongside the Ames test when the E coli test had been conducted); in vivo chromosome aberration, nuclear anomaly and sister chromatid exchange (SCE) tests and in vitro mammalian cell gene mutation assays, unscheduled DNA synthesis, and SCE tests also occurred in this set of tests. Interestingly, for no chemical was there more than one positive test.

One hundred and twenty chemicals gave negative results in both a bacteriological (Ames) test and a non-bacteriological test. Some of these chemicals also had negative results from gene mutation tests in E coli and/or from other non-bacteriological tests. Two chemicals were assumed by analogy to structurally similar chemicals to be negative and were not tested. Thus, following testing, 122 chemicals (85% of the chemicals in the project) were considered negative. SAR predictions of low concern for mutagenicity were made for 107 chemicals in this group (88% of the MPD "negatives").

Depending on how the analysis is done "false positive" predictions were made for 14 (10% of total) or 2 (1.4%) chemicals. A direct reading of the MPD results would lead one to conclude that there were 14 false positive predictions. However, EPA considers that positive results would be produced if tests were performed using assay systems other than those used already to test the affected

chemicals. The EPA conclusions are based on the existence of data on analogues (chemical or mechanistic) indicating positive results in certain test systems. It will be of interest (and potential importance) to see whether the predictions of positive mutagenicity are fulfilled if further test data become available.

Six chemicals (4% of total) with positive test data were predicted "low" (false negatives) because of absence of known positive data in analogues.

The test results (including, where appropriate, an indication of weak positive results), EPA predictions and results of comparison are given in Annex 15 for all of the chemicals in the project.

- Conclusions

A high proportion of the chemicals in this project were negative for mutagenicity and a high proportion of these were correctly identified by the EPA.

Although the number of test-positive chemicals was small, it is also of concern that six of them were called low. The observation that 123 of 142 data pairs (87%) were apparently correctly predicted thus has to be seen in the light of the above comment. For this reason it would not be prudent at this time to replace mutagenicity testing of new chemicals in the EC with the predictive methods used in the US for PMN chemicals.

As the EC classification system for mutagenicity, as applied to notified new chemicals, depends essentially on testing in vivo to investigate whether effects observed in vitro are expressed in vivo, the predictive methods used here, which do not make this distinction, could not be used for classification in the EC.

4.2.3.7. Other effects

A number of effects were considered using the predictive methods which had not yet been investigated in the EC testing programme for the chemicals in this project i.e. reproductive and developmental toxicity, neurotoxicity and oncogenicity. For some chemicals, indications of some of these effects (e.g. clinical signs of neurotoxicity; changes affecting the reproductive organs) may be reported for the acute or repeated dose tests. Such reports were made for some chemicals in this project: five chemicals had significant indications of potential reproductive toxicity (76, 151, 186, 200 and 292) and reproductive toxicity was predicted for chemicals 200 and 292 but not for the others (developmental toxicity was predicted for chemical 76). Signs of neurotoxicity were seen with six chemicals (54, 268, 340, 342, 431 and 434) and neurotoxicity was predicted for two of these (54 and 340).

Adverse effects on reproduction and/or development were predicted by EPA for 51 chemicals (35%); 27 chemicals were predicted to be neurotoxic (19%) and 33 (23%) to be oncogenic. This is of particular concern as these potential effects are not specifically investigated in the initial testing of new chemicals in the EC.

The health concerns for which the MPD data set does not provide data were analysed for number of chemicals for which such concerns were expressed and the frequency of occurrence. Of the 144 chemicals, 66 (44%) had concerns that addressed health effects outside the scope of the MPD data set. The breakdown by effect and frequency of occurrence is presented in Table 18.

TABLE 18: Health concerns not addressed by the MPD data set

Concern	Number of chemicals	% of Total chemicals
Oncogenicity	33	23
Developmental toxicity	46	32
Reproductive toxicity	13	9
Neurotoxicity	21	15
Immunotoxicity	2	-
Photosensitisation	1	-
Lung	1	-
Respiratory sensitisation	1	-

This table indicates that potential adverse effects beyond those in the MPD were identified for a substantial number of the chemicals, which implies that hazards and possibly risks may be underestimated if these effects are not considered. There may be a need for early focused testing in at least some of these cases.