METHODS

The data on each chemical were reviewed in detail before the meeting by selected members of the group: the animal studies and short-term test results were evaluated by experimentalists and the human studies by an epidemiologist. During the meeting of the Working Group these assessments were debated and adopted, and overall evaluations of carcinogenicity for humans were made on the basis of the combined evidence from humans and experimental systems (Table 1). Brief descriptions of the data on which the assessments and evaluations were based are given in the section on Results, together with references to the *Monographs* volumes in which they were evaluated previously and, when applicable, to papers published subsequently.

Assessment of evidence for carcinogenicity from studies in humans

Evidence of carcinogenicity from human studies comes from three main sources:

1. Case reports of individual cancer patients who were exposed to the chemical or process.
2. Descriptive epidemiological studies in which the incidence of cancer in human populations was found to vary in space or time with exposure to the agents.
3. Analytical epidemiological (case-control and cohort) studies in which individual exposure to the chemical or group of chemicals was found to be associated with an increased risk of cancer.

Three criteria must be met before a causal association can be inferred between exposure and cancer in humans:

1. There is no identified bias which could explain the association.
2. The possibility of confounding has been considered and ruled out as explaining the association.
3. The association is unlikely to be due to chance.

In general, although a single study may be indicative of a cause-effect relationship, confidence in inferring a causal association is increased when several independent studies are concordant in showing the association, when the association is strong, when there is a dose-response relationship, or when a reduction in exposure is followed by a reduction in the incidence of cancer.

The degrees of evidence for carcinogenicity from studies in humans were categorized as:

i. *Sufficient evidence* of carcinogenicity, which indicates that there is a causal relationship between the agent and human cancer.
ii. *Limited evidence* of carcinogenicity, which indicates that a causal interpretation is credible, but that alternative explanations, such as chance, bias or confounding, could not adequately be excluded.
iii. *Inadequate evidence*, which indicates that one of three conditions prevailed: (a) there were few pertinent data; (b) the available studies, while showing evidence of association, did not exclude chance, bias or confounding; (c) studies were available which do not show evidence of carcinogenicity.

Assessment of evidence for carcinogenicity from studies in experimental animals

These assessments were classified into four groups:

i. *Sufficient evidence* of carcinogenicity, which indicates that there is an increased incidence of malignant tumours: (a) in multiple species or strains; or (b) in multiple experiments (preferably with different routes of administration or using different dose levels); or (c) to an unusual degree with regard to incidence, site or type of tumour, or age at onset. Additional evidence may be provided by data on dose-response effects, as well as information from short-term tests or on chemical structure.

ii. *Limited evidence* of carcinogenicity, which means that the data suggest a carcinogenic effect but are limited because: (a) the studies involve a single species, strain, or experiment; or (b) the experiments are restricted by inadequate dosage levels, inadequate duration of exposure to the agent, inadequate period of follow-up, poor survival, too few animals, or inadequate reporting; or (c) the neoplasms produced often occur spontaneously and, in the past, have been difficult to classify as malignant by histological criteria alone (e.g., lung and liver tumours in mice).

iii. *Inadequate evidence*, which indicates that because of major qualitative or quantitative limitations, the studies cannot be interpreted as showing either the presence or absence of a carcinogenic effect; or that within the limits of the tests used, the chemical is not carcinogenic. The number of negative studies is small, since, in general, studies that show no effect are less likely to be published than those suggesting carcinogenicity.

iv. *No data* indicates that data were not available to the Working Group.

The categories *sufficient evidence* and *limited evidence* refer only to the strength of the experimental evidence that these chemicals are carcinogenic and not to the extent of their carcinogenic activity nor to the mechanism involved. The classification of any chemical may change as new information becomes available.

Assessment of data from short-term tests

Because of the large number and wide variety of short-term tests that may be relevant for the prediction of potential carcinogens, the data relative to each compound have been summarized in the form of tables. These indicate both the type of test used and the biological complexity of the test system. *'DNA damage'* includes evidence for covalent binding to DNA, induction of DNA breakage or repair, induction of prophage in bacteria, and a positive response in tests of comparative survival in DNA repair-proficient and DNA repair-deficient bacteria. *'Mutagenicity'* refers to induction of mutations in cultured cells or in organisms (e.g., heritable alterations in phenotype, including forward or reverse point mutations, recombination, gene conversion, and specific-locus mutation). *'Chromosomal anomalies'* refers to the induction of chromosomal aberrations, including breaks, gaps, rearrangements and micronuclei, sister chromatid exchange and aneuploidy. *'Other'* refers to various additional endpoints, including cell transformation (T), i.e., morphological transformation and colony formation in agar; dominant lethal (DL) tests; morphological abnormalities in sperm (SA); and mitochondrial mutation (Mt). The biological systems include: *'Prokaryotes'*, i.e., bacteria, in the presence or absence of an exogenous metabolic activation system, and cellular systems; *'Fungi and green plants'*; *'Insects'*, usually *Drosophila melanogaster*; *'Mammalian cells* (in vitro)';,either rodent or human somatic cells or cell lines in culture; *'Mammals* (in vivo)', studies in which the test compound was administered to intact experimental animals; and *'Humans* (in vivo)', studies of cells from groups of individuals drawn from a population exposed to the substance in question.

In these tables, a '+' indicates that the result was judged by the Working Group to be significantly positive in one or more assays; '-' indicates that it was judged to be negative from an evaluation of one or more assays; and '?' indicates that contradictory results were obtained in assays from different laboratories or in different biological systems, or that the result was judged to be equivocal. The individual tables for each compound are summarized, for purposes of comparison, in Appendix 3.

The overall evidence summarized in the table was adjudged to fall into one of three categories, *sufficient*, *limited* and *inadequate*:

i. *Sufficient evidence*, when there were at least three positive results in at least two of three test systems measuring DNA damage, mutagenicity or chromosomal effects. When two of the positive results were for the same genetic effect, they had to be derived from systems of different biological complexity.
ii. *Limited evidence*, when there were at least two positive results, either for different endpoints or in systems representing two levels of biological complexity.
iii. *Inadequate evidence*, when there were generally negative or only one positive test results. Up to two positive test results were considered inadequate if they were accompanied by two or more negative test results.

The Working Group was unable to define criteria for 'negative' evidence.

In establishing these categories the Working Group gave greater weight to the three primary endpoints - DNA damage, mutagenicity and chromosomal effects - and judgements were made on the quality as well as on the quantity of the evidence. In a minority of cases, strict interpretation of these criteria was tempered by consideration of a variety of other factors (such as the purity of the test compound, problems of metabolic activation, appropriateness of the test system) which, in the judgement of the Working Group, would place a compound in a category above or below that indicated by the summary table.

Evaluation of carcinogenic risk to humans

At present, no objective criteria exist to interpret data from studies in experimental animals or from short-term tests directly in terms of human risk. Thus, in the absence of *sufficient evidence* from human studies, evaluation of the carcinogenic risk to humans was based on consideration of both the epidemiological and experimental evidence. The breadth of the categories of evidence defined above allows substantial variation within each. The decisions reached by the Group regarding overall risk incorporated these differences, even though they could not always be reflected adequately in the placement of an exposure into a particular category, as listed in Table 1.

The chemicals, groups of chemicals, industrial processes or occupational exposures were thus put into one of three groups:

Group 1

The chemical, group of chemicals, industrial process or occupational exposure is carcinogenic to humans. This category was used only when there was *sufficient evidence* from epidemiological studies to support a causal association between the exposure and cancer.

Group 2

The chemical, group of chemicals, industrial process or occupational exposure is probably carcinogenic to humans. This category includes exposures for which, at one extreme, the evidence of human carcinogenicity is almost *'sufficient'*, as well as exposures for which, at the other extreme, it is inadequate. To reflect this range, the category was divided into higher (*Group A*) and lower (*Group B*) degrees of evidence. Usually, category 2A was reserved for exposures for which there was at least *limited evidence* of carcinogenicity to humans. The data from studies in experimental animals played an important role in assigning studies to category 2, and particularly those in Group B; thus, the combination of *sufficient evidence* in animals and inadequate data in humans usually resulted in a classification of 2B.

In some cases, the Working Group considered that the known chemical properties of a compound and the results from short-term tests allowed its transfer from Group 3 to 2B or from Group 2B to 2A.

Group 3

The chemical, group of chemicals, industrial process or occupational exposure cannot be classified as to its carcinogenicity to humans.