Arch Toxicol (1995) 69: 659–670

ORIGINAL INVESTIGATION

E. Schlede · U. Mischke · W. Diener · D. Kayser

The international validation study of the acute toxic class method (oral)

Received: 10 March 1995 / Accepted: 27 April 1995

Abstract An alternative to the oral LD_{50} test, the acute toxic class (ATC) method (oral), was validated with 20 substances in an international collaborative study with nine laboratories in five countries. The ATC method is a stepwise procedure with the use of three animals per step. It has been designed with three fixed doses (25, 200 and 2000 mg/kg). In general, this testing is sufficient for allocation to the toxicity classes of the majority of the international classification systems currently in use. The selection of testing at additional fixed doses (5, 50 and 500 mg/kg) may be considered if further refinement is necessary or for specific allocation to those international classification systems with a cut-off value of 5 mg/kg. On average, two to four steps are necessary to complete a test. With the ATC method substances can be ranked in a similar or even better manner than with an LD_{50} test but it uses up to 90% fewer animals, the average being 70% fewer. This also results in substantially fewer moribund/dead animals. The ATC method is based on biometric evaluations that, together with the experimental results, demonstrate that this method is a sensitive and reliable alternative to the LD_{50} test.

Key words Acute toxic class method (oral) · Alternative · LD_{50} test · Animal welfare · International classification systems

Introduction

Alternatives to conventional in vivo acute toxicity testing should be in vitro methods or animal tests with the use of substantially fewer animals than the LD_{50} test (Organisation for Economic Cooperation and Development, OECD 1981a, 1987; Commission of the European Communities 1992a). Until now, no validated in vitro test for the determination of acute oral toxicity exist. Therefore animal testing methods should be designed in such a way as to meet the demands of animal welfare without impairing the requirements for human health protection. Some in vivo alternatives to the classical LD_{50} test have been published (Deichmann and Le Blanc 1943; Lorke 1983; Bruce 1985; Kennedy et al. 1986; Yamanaka et al. 1990) and one alternative – the fixed dose procedure (Van den Heuvel et al. 1990) – has gained international acceptance (OECD 1992; Commission of the European Communities 1992b).

For animal welfare reasons, and also for other toxicological purposes, the LD_{50} test is not necessary. There should be no need to determine precisely a numerical value to describe the toxicity of a substance; it should suffice to define the toxicity class within justifiable limits.

The ATC method is an alternative, with the use of substantially fewer animals leading to significantly fewer moribund/dead animals. Moreover, this method increases consistency from laboratory to laboratory. The ATC method is based on biometric evaluations (Roll et al. 1989; Diener et al. 1994) and has been successfully tested in a national validation study for the classification system of chemicals of the European Union (EU) (Schlede et al. 1992).

In this paper the experimental and biometrical results of an international collaborative study of the ATC method are presented, taking into consideration all currently used international classification systems. This study provides a worldwide acceptance of this method in the same manner as the LD_{50} test.

The study was sponsored by the German Government and was carried out under the patronage of the OECD.

Study design

Nine laboratories from five countries participated in this study (Table 1). The Bundesgesundheitsamt was responsible for the coordination of the study.

E. Schlede (✉) · U. Mischke · W. Diener · D. Kayser
Fachbereich Chemikalienbewertung, Bundesinstitut für gesundheitlichen Verbraucherschutz und Veterinärmedizin, Thielallee 88–92, D-14195 Berlin, Germany

Table 1 List of participants in the international validation study

Country	Laboratory
Austria	Austrian Research Center Seibersdorf, Vienna (H. Hofer, W. Klein)
Belgium	Ministère de la Santé Publique et de l'Environment, Institut d'Hygiene et d'Epidémiologie, Brussels (M. P. Delcour-Firquet)
Japan	Biosafety Research Center, Foods, Drugs and Pesticides, Shizuoka Prefecture (H. Inoue)
	Hatano Research Institute, Food & Drug Safety Center, Kanagawa Prefecture (H. Ono)
	National Institute of Health Sciences, Tokyo (Y. Kurokawa)
Switzerland	Ciba Geigy Ltd, Basel (H. R. Hartmann)
	Research and Consulting Company Ltd (RCC), Itingen/Basel (L. Ullmann)
USA	Bushy Run Research Center, Export, Pa. (R. Myers)
	US Environmental Protection Agency, Health Effects Research Laboratory, Research Triangle Park, N.C. (E. Grose, S. McMaster)

Twenty of the 30 substances tested in the national validation study were selected for the tests. All LD_{50} data were derived from the literature (see Schlede et al. 1992). For most of the 20 test substances (Table 2), the geometrical means or in some cases the arithmetical means (nos 10, 12, 14 and 15) of the LD_{50} values were determined. Due to lack of information for no. 19, a fictitious value of >2000 mg/kg was assumed. With these LD_{50} data, toxicity classes were allocated to the currently used classification systems and to the particular classes of the international validation study.

Eight laboratories conducted the tests according to Good Laboratory Practice (OECD 1981b).

The choice of the rat strain was optional. Six laboratories used Sprague-Dawley rats, three used Wistar rats. The choice of sex for the first treatment was also optional. For the first treatment four laboratories used only male, three laboratories used only female rats, and the other two used either male or female rats.

The animals were carefully monitored after the administration of the substance. Observations were made at least six times during the day of treatment, followed by daily observations until the end of the test. Necropsy findings were also documented.

The participants were free to kill the animals for animal welfare reasons when they considered it necessary.

All participants knew the vehicle to be used.

The substances were supplied by the Bundesgesundheitsamt to the participants. They were at least of analytical grade and each substance was part of one lot. Altogether, 180 tests were carried out.

The experiments were started in October 1991 and all tests were completed in October 1993.

Method

The ATC method uses a sequential testing procedure with three animals of each sex per step. The dose to be used as the starting dose is selected from one of three fixed dose levels, 25, 200 and

Table 2 LD_{50} values, experimental and biometrical classification results of the ring study (*exp.* experimental, *bio.* biometrical)

No.	Substance	LD_{50}	Number of classifications for toxicity classes A–H															
			A		B		C		D		E		F		G		H	
			exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.
1	Aldicarb	1	–	–	–	–	–	–	–	–	–	–	–	–	–	0.4	9	8.6
2	Parathion	4	–	–	–	–	–	–	–	–	–	–	–	–	3	1.0	6	8.0
3	Thiosemicarbazide	12	–	–	–	–	–	–	–	–	–	–	–	–	9	9.0	–	–
4	Indomethacin	13	–	–	–	–	–	–	–	–	1	–	4	0.2	4	7.8	–	1.0
5	*N*-Phenylthiourea	9	–	–	–	–	–	–	3	–	1	0.1	1	0.9	2	2.7	2	5.2
6	Mercury (II) oxide	29	–	–	–	–	–	–	–	–	3	0.1	5	3.2	1	5.8	–	–
7	Sodium arsenite	38	–	–	–	–	–	–	–	–	3	0.5	5	6.0	1	2.5	–	–
8	Bis(tributyltin)oxide	147	–	–	–	–	2	–	5	1.3	2	5.4	–	1.8	–	0.5	–	–
9	Acrylamide	163	–	–	–	–	–	–	1	0.8	8	8.2	–	–	–	–	–	–
10	Cadmium chloride	* 237	–	–	–	–	–	–	2	3.6	7	5.4	–	–	–	–	–	–
11	Caffeine	270	–	–	–	–	–	–	8	5.4	1	3.6	–	–	–	–	–	–
12	Aniline	* 822	–	–	–	–	8	8.4	1	0.6	–	–	–	–	–	–	–	–
13	Ferrocene	1280	–	–	–	–	5	8.9	4	0.1	–	–	–	–	–	–	–	–
14	Sodium salicylate	*1601	–	–	–	–	9	9.0	–	–	–	–	–	–	–	–	–	–
15	Acetanilide	*1689	1	–	2	0.1	6	7.6	–	1.2	–	0.1	–	–	–	–	–	–
16	Acetonitrile	2515	–	–	4	0.4	4	7.5	–	0.1	–	–	–	–	–	–	–	–
17	Butylated hydroxyanisole	2853	–	–	2	0.4	6	7.3	–	0.2	–	–	–	–	–	–	–	–
18	*N,N*-Dimethylformamide	4604	3	–	3	1.0	2	6.8	–	0.2	–	–	–	–	–	–	–	–
19	Quercetin dihydrate	>2000	8	8.0	–	–	–	–	–	–	–	–	–	–	–	–	–	–
20	Ethylene glycol	6336	8	4.0	–	4.0	–	–	–	–	–	–	–	–	–	–	–	–

LD_{50}: geometrical or arithmetical (*) mean LD_{50} values (mg/kg)

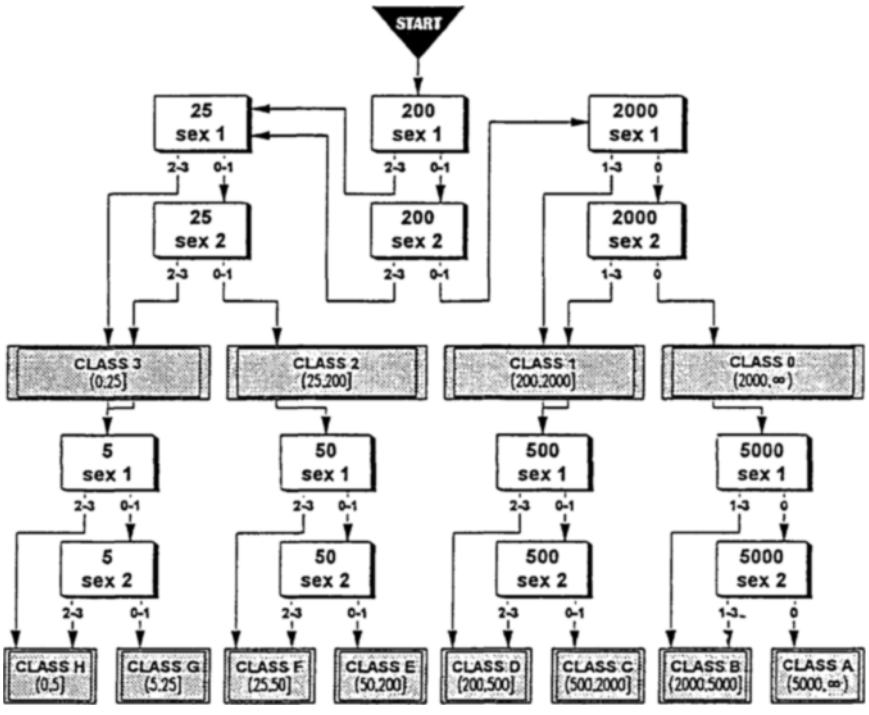


Fig. 1 Test procedure of the ring study with three animals per sex and dose level. Abbreviations: *5, 25, 50, 200, 500, 2000, 5000* doses and/or class limits in mg/kg; *0, 1, 2, 3* number of moribund/dead animals

2000 mg/kg. The results obtained can be allocated to the commonly used classification systems (option 1). Exceptionally, it may be desirable to achieve a further refinement of classification than would be possible after conducting the test with the three fixed dose levels of 25, 200 or 2000 mg/kg. In these cases further testing at additional fixed dose levels of 5, 50 or 500 mg/kg may be considered (option 2). This procedure was agreed upon at the OECD Expert Meeting held in Berlin from January 26 to 28, 1994.

The experimental design of the tests conducted by the participants was as follows: the study was conducted "blind". Due to the broad toxicity profile of the test substances, it was decided for animal welfare reasons to use only the starting dose of 200 mg/kg (Fig. 1). In addition, testing of the next higher dose level of 2000 mg/kg body weight was performed when at the 200 mg/kg level no, or one animal died at each step. The testing of 500 mg/kg body weight was only possible after one or two steps of the 2000 mg/kg level had been tested. The 5000 mg/kg dose level was tested in the study but the participants at the OECD Expert Meeting agreed unanimously not to include this dose level in the official guideline. For reasons of completeness the results of the tests obtained with the use of 5000 mg/kg body weight are documented here. One laboratory decided for reasons of animal welfare not to test the 5000 mg/kg dose level.

All data presented in this paper were evaluated according to the schemes shown in Figs 2 and 3. In principle, three test procedures are available. The first procedure, defined as option 1, but with the additional doses of 5 and 5000 mg/kg, as shown in Fig. 2, is a modified test scheme of a previous publication (Diener et al. 1994). This test design allows classification for all commonly used systems. Figure 3 shows option 2, with subsequent testing at doses of 5, 50 and 500 as well as with the additional dose of 5000 mg/kg. This is the second procedure, allowing classification for systems with limits of 5, 50 and 500 mg/kg. However, this second procedure has the disadvantage that it is not always possible to classify a substance afterwards without repeating the test procedure. This procedure results in a moderately better classification for borderline substances at 50 or 500 mg/kg. The third procedure, also shown in Fig. 3, allows testing of the additional doses of option 2 only after having completed all steps of option 1. This third procedure (referred to as option 1 and 2) also allows allocation to all commonly used classification systems, in general without repeating the test procedure.

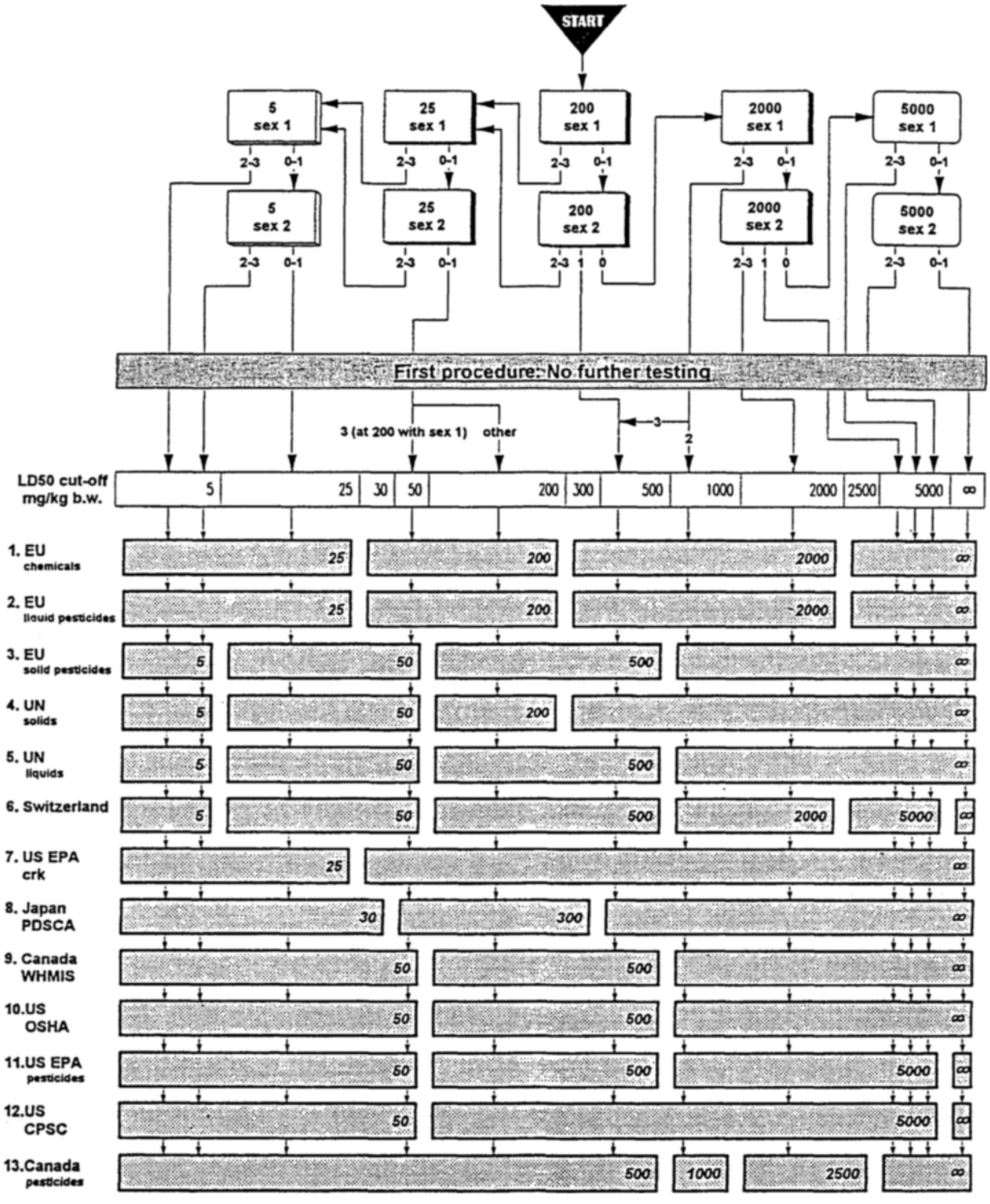
The experimental results were also predicted by biometrical methods. For prediction of the correct classification probabilities and the animal numbers for each substance, the formulae and the procedures published by Diener et al. (1994) were used but taking into account the changes proposed at the OECD Expert Meeting. For all calculations LD_{50}- and β-values are needed. The LD_{50} values were taken from the literature and then determined as a geometrical mean or, if not possible, as an arithmetical mean of the gender with the lowest values. These numbers are used as reference values to define the term "correct classification". In contrast to the LD_{50} values, slopes were only available for very few substances in the literature. Therefore all slopes were derived from the study. In order to estimate the LD_{50} values and the β-slopes of the 20 substances, the test results of all nine laboratories were used. The values were then calculated for each substance by using the maximum likelihood method based on the probit model. A computation is only possible when there are at least two doses which result in different mortality rates not equal to 0% and 100%.

Results

Validation of the ring study tests

The test procedure of the international validation study is shown in Fig. 1, and it was designed in such a way that the currently used classification systems would be covered. The experimental results obtained for the classification of classes A to H are shown in Table 2. Due to the small class size in some instances (for example, a factor of 2 for the two limits of class F), an allocation to more than two toxicity classes would be expected and this occurred with seven substances (nos 4–8, 15 and 18). Five substances were allocated to only one class (nos 1, 3, 14, 19 and 20) and the remaining seven substances to two classes.

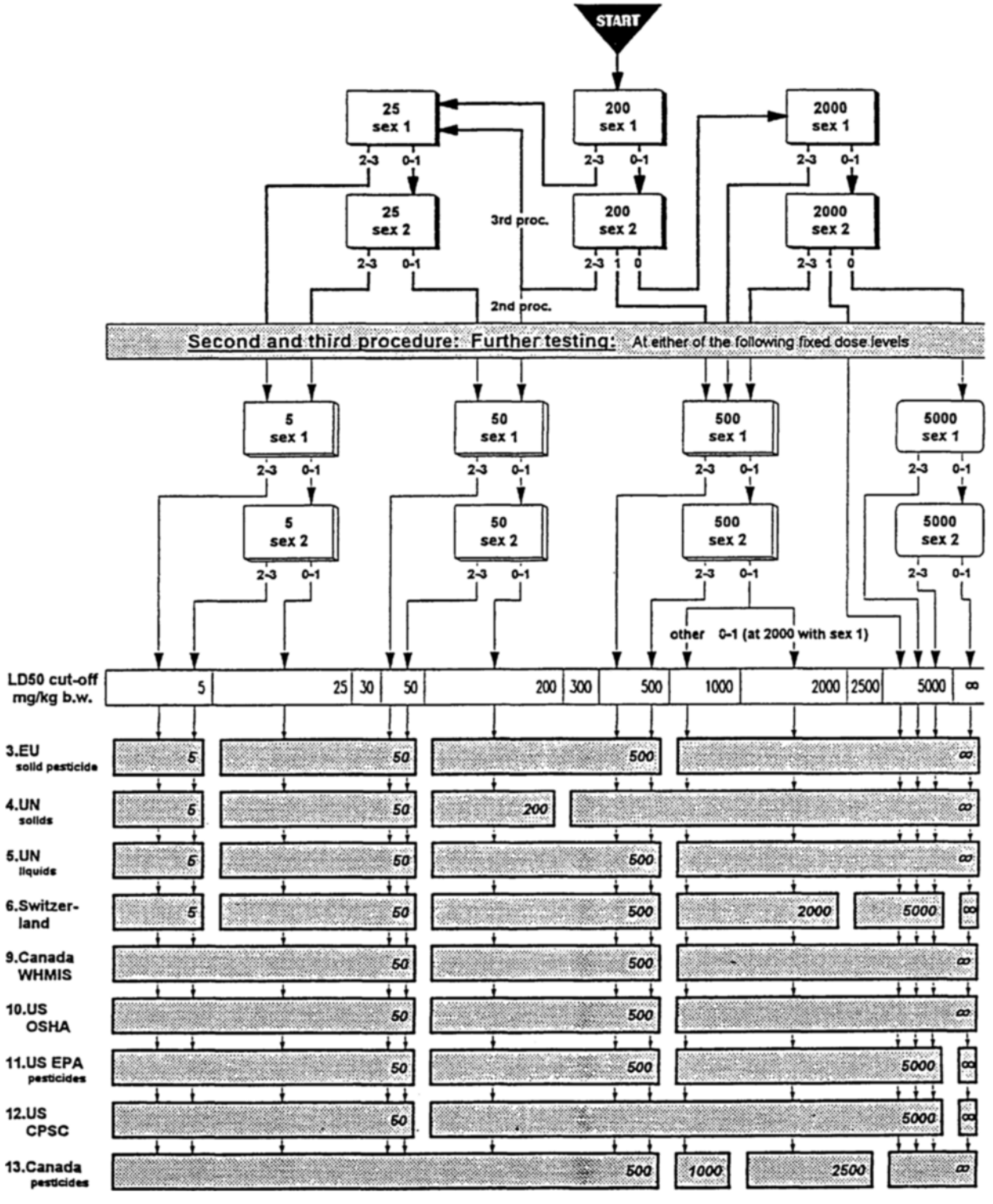
Compared with the experimental results, the biometrical results are not very different. These differences are essen-



tially caused by the small class ranges, the differences between the LD_{50} values of the literature and those of the ring study for some substances and also by the rather small number of participating laboratories.

Out of a total of 180 tests, 175 (97.2%) tests could be used for all classification results (Fig. 1). In five cases (2.8%) the dose of 5000 mg/kg was not tested. In one test a starting dose of 25 mg/kg was used with substance no. 3

Fig. 2 First procedure of the ATC method with three animals per sex and dose level. Abbreviations: *5, 25, 50, 200, 500, 2000, 5000* dose and/or cut-off value in mg/kg; *0, 1, 2, 3* number of moribund/dead animals. For correct classification not necessary to test: two doses of 5: systems 1, 2, 7, 8, 9, 10, 11, 12, 13; two doses of 25: system 13 and in the case of 3 (at 200 with sex 1) systems 9, 10, 11, 12; two doses of 5000: systems 1, 2, 3, 4, 5, 7, 8, 9, 10, 13; two doses of 2000: systems 4, 7, 8; second dose of 2000: systems 3, 5, 9, 10. *EU* European Union, *UN* United Nations, *US EPA* United States Environmental Protection Agency, *crk* community right tn know, *PDSCA* Poisonow and Deleterious Substances Control Act, *WHMIS* Workplace Hazardous Materials Information System, *OSHA* Occupational Safety and Health Administration, *CPSC* Consumer Protection Safety Commission



and a starting dose of 2000 mg/kg was used with two tests with substances nos 9 and 20.

Altogether, 2565 animals were used and the mean was 14.3 animals per test. The total number of moribund/dead animals was 736 and on the average 4.1 animals were moribund/dead per test. Twenty-two animals were killed in a moribund status.

Fig. 3 Second and third procedure of the ATC method with three animals per sex and dose level. Abbreviations: *5, 25, 50, 200, 500, 2000, 5000* dose and/or cut-off value in mg/kg; *0, 1, 2, 3* number of moribund/dead animals. For correct classification not necessary to test: two doses of 5: systems 9, 10, 11, 12, 13; two doses of 25: system 13; two doses of 5000: systems 3, 4, 5, 9, 10, 13; two doses of 2000: systems 12 (only second procedure), and 3, 4, 5, 9, 10; two doses of 50: system 13. *Thick line:* part of option 1 and 2; thin line: part of option 2 (with 5000 mg/kg)

Inter-laboratory results were identical in 72% of the tests. A comparison between the classified LD_{50} values of the literature and the classes of the ATC method showed an identical classification of 65%.

Validation of the ring study tests for the international classification systems

The following results will be presented: classification of the test substances and numbers of animals used and of moribund/dead animals. For the classification systems of the EU (chemicals and liquid pesticides) and likewise of the EU (solid pesticides) and UN (liquids) the evaluation and results for these systems will be shown in detail for the 20 tested substances (Tables 3 and 4). The data obtained for the other classification systems with these test schemes (Figs 2 and 3) will be tabulated in a summarized form (Table 5).

Classification systems of the EU for chemicals and liquid pesticides

The classes for these two systems are (0,25], (25,200], (200,2000] and (2000,∞) mg/kg, defined as class 3 to class 0, respectively.

The evaluation scheme is shown in Fig. 4 and Table 3 lists the experimental and biometrical results.

The experimental results of the tests demonstrate that eight substances (nos 1–3, 12–14, 19 and 20) were allocated to the predicted class by all laboratories. Eleven substances (nos 4, 6–11, 15–18) spanned two toxicity classes. The overwhelming majority of the laboratories ranked these substances to the predicted classes and only with substances nos 4, 8 and 10 did the majority of test results result in allocation to a less toxic (nos 4 and 8) or to a more toxic class (no. 10). Only phenylthiourea (no. 5) spanned three classes.

Identical interlaboratory results were obtained, with the exception of the above mentioned deviations.

The mean number of animals used increased with decreasing toxicity of the tested substances, ranging from 5.7 animals per test to no more than 12.0 animals per test. Mortality occurred in all tested animals with the most toxic substances but it never exceeded a mean number of 5.9 animals per test. Almost no deaths were seen with the least toxic substances.

The biometrical results are in good agreement with the experimental results. Notable deviations from the experimental results were detected in two cases: for substance no. 4a more frequent biometric allocation to the correct class occurred in comparison to the experimental results and with substance no. 5a correct classification was predicted for 88% compared to 44% of the test results.

The test results obtained for the 20 tested substances in the national validation study for the classification system of the EU for chemicals (Schlede et al. 1992) were compared with the results of the international study with the same classification system according to Fig. 4. Allocation to the assigned toxicity class occurred in 88% and 78%, respectively.

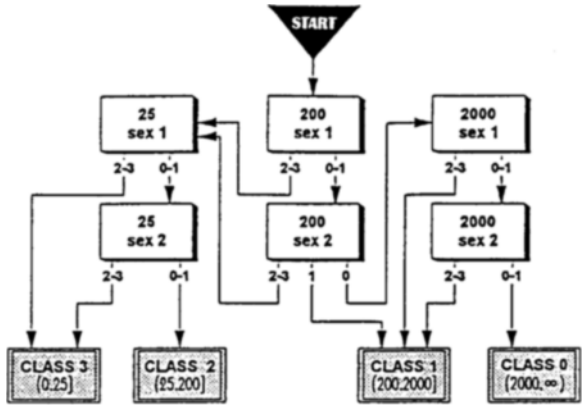


Fig. 4 The ATC test procedure for the classification system of the EU for chemicals and liquid pesticides with three animals per sex and dose level. Abbreviations: *25, 200, 2000* doses and/or class limits in mg/kg; *0, 1, 2, 3* number of moribund/dead animals

Classification systems of the EU for solid pesticides and of the UN for liquids

The class limits for these two systems are (0,5], (5,50], (50,500] and (500,∞] mg/kg, defined as class 3 to class 0, respectively.

Figure 5 shows the three different evaluation procedures. In Table 4 the experimental and biometrical results obtained for these three procedures are presented.

Altogether, the experimental classification results for the three procedures are not very different. Only with the first procedure substantial deviations were found in comparison to the second and third procedures with substances nos 9, 12 and 13. In these cases allocation to a more stringent toxic class occurred more frequently with the first procedure. For all three procedures one toxic class was assigned for substances nos 1, 3, 16, 19 and 20 and another substance, phenylthiourea (no. 5), spanned three toxicity classes. In general, for the other substances the ranking to two neighbouring classes occurred for all procedures. Exceptions are substances nos 9–11, 14, 15, 17 and 18, where the allocation to only one class increased when the second and/or the third procedure was used. Interlaboratory results varied only slightly for the three procedures. As expected, the number of animals used increased with the use of the second and/or third procedure but the highest number did not exceed 15 animals per tested substance. In contrast, the number of moribund/dead animals varied, depending on the choice of the procedure and the toxicity profile of the tested substances. The number of moribund/dead animals was identical or increased for substances nos 1–11. On the other hand, for substances nos 12–20 of class 0, with respect to the second procedure, very few or even no moribund/dead animals were found.

Table 3 Experimental and biometrical results per substance for the classification systems of the EU (chemicals and liquid pesticides)

Allocated class, number and name of the substance	Number of laboratories per class								ATC classification (%)				Mean number of animals			
	Class 0		Class 1		Class 2		Class 3		≙ LD_{50}		identical		used		moribund/dead	
	exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.
Class 3																
1 Aldicarb	–	–	–	–	–	–	9	9.0	100	100	100	100	6.0	6.1	5.9	5.7
2 Parathion	–	–	–	–	–	–	9	9.0	100	100	100	100	6.0	6.0	5.9	6.0
3 Thiosemicarbazide	–	–	–	–	–	–	9	9.0	100	100	100	100	5.7	6.0	5.6	5.9
4 Indomethacin	–	–	–	–	5	0.2	4	8.8	44	98	56	98	8.0	6.4	4.4	5.6
5 *N*-Phenylthiourea	–	–	3	0.1	2	1.0	4	7.9	44	88	44	88	8.7	7.3	4.2	5.1
Class 2																
6 Mercury(II)oxide	–	–	–	–	8	3.2	1	5.8	89	36	89	64	9.3	7.8	3.7	5.1
7 Sodium arsenite	–	–	–	–	8	6.5	1	2.5	89	72	89	72	8.7	8.6	3.8	4.3
8 Bis(tributyltin)oxide	–	–	7	1.3	2	7.2	–	0.5	22	80	78	80	9.0	9.3	3.0	3.1
9 Acrylamide	–	–	1	0.8	8	8.2	–	–	89	92	89	92	9.0	9.4	3.1	2.5
Class 1																
10 Cadmium chloride	–	–	3	3.6	6	5.4	–	–	33	40	67	60	9.7	8.9	2.8	2.4
11 Caffeine	–	–	8	5.4	1	3.6	–	–	89	60	89	60	9.0	8.5	2.8	2.4
12 Aniline	–	–	9	9.0	–	–	–	–	100	100	100	100	9.0	9.0	3.0	3.0
13 Ferrocene	–	0.1	9	8.9	–	–	–	–	100	99	100	99	9.0	9.3	2.9	2.7
14 Sodium salicylate[a]	–	0.4	6	5.6	–	–	–	–	100	93	100	93	9.0	9.8	2.5	2.5
15 Acetanilide[a]	3	1.2	5	6.7	–	0.1	–	–	63	84	63	84	10.1	9.8	1.6	2.3
Class 0																
16 Acetonitrile[a]	5	3.2	3	4.8	–	–	–	–	63	40	63	60	10.9	10.9	0.8	2.0
17 Butylated hydroxyanisole[a]	5	3.4	3	4.6	–	–	–	–	63	43	63	57	10.5	10.8	1.1	1.9
18 *N,N*-Dimethylformamide[a]	7	4.7	1	3.3	–	–	–	–	88	59	88	59	11.3	11.1	0.4	1.6
19 Quercetin dihydrate	9	9.0	–	–	–	–	–	–	100	100	100	100	12.0	12.0	0.0	0.0
20 Ethylene glycol	9	9.0	–	–	–	–	–	–	100	100	100	100	11.3	12.0	0.0	0.0

[a] Not feasible to evaluate all test results due to differences in the testing and evaluation procedure (see Figs 1 and 2)

With respect to the biometrical results, no noteworthy deviations were detected when compared to the experimental results.

All currently used classification systems

The class limits for all systems are shown in Figs 2 and 3. Examples of the evaluation procedure are shown in Figs 4 and 5. Table 5 shows the experimental and biometrical results. For direct comparison the results of the above mentioned systems and of the ring study are included. Identical class systems (nos 1 and 2, 3 and 4, 9 and 10; see also Fig. 2) are not covered separately.

Due to the specific class limits, an evaluation for all three procedures is not sensible or feasible for all systems. This is the case for systems nos 1, 2, 7, 8 and the ring study.

For the first procedure the experimental results for a correct classification ranged between 63% (no. 6) and 93% (no. 7). The majority of the test results were in the range of 80%. In general, with the second and third procedure higher rates of correct classification were produced. With the exception of no. 6 they ranged from 81 to 90%. Worth mentioning were the deviations from the correct classification to a less stringent class with classification system no. 8, and for the second and third procedure for systems nos 11 and 12. A deviation to a more stringent class was found more frequently for systems 3, 4, 6, 9–11 and 13 for the results obtained with the first procedure.

For the currently used classification systems, identical interlaboratory results ranged from 82 to 94%.

On average, fewer than ten animals were used, 6.9 animals being the lowest and 14.0 animals being the highest mean number. The mean number of moribund/dead animals ranged from 2.0 to 4.3 animals.

Very good agreement could be demonstrated between the experimental and biometrical results.

Biometric determination of the LD_{50} values and of the slopes

Figure 6a demonstrates the distribution of the 20 test substances with respect to their characteristic parameters LD_{50} and β as estimated from the ring study. Due to the test results for substances nos 1, 19 and 20 the LD_{50} values and their slopes could not be determined. The remaining substances 2–18 are randomly located for β from 0.5 to 6 and for LD_{50} from 5 to 20000 mg/kg. The correlation between the LD_{50} values of the literature and the LD_{50} values derived from the ring study is shown in Fig. 6b. The correlation coefficient of 0.956 is, as expected, very large. Little difference was observed for some of the test substances. With substances 3, 9, 10 and 13 the test results produced more stringent toxicity than the classified LD_{50} values of the literature and the opposite occurred especially with substances 5, 8 and 18.

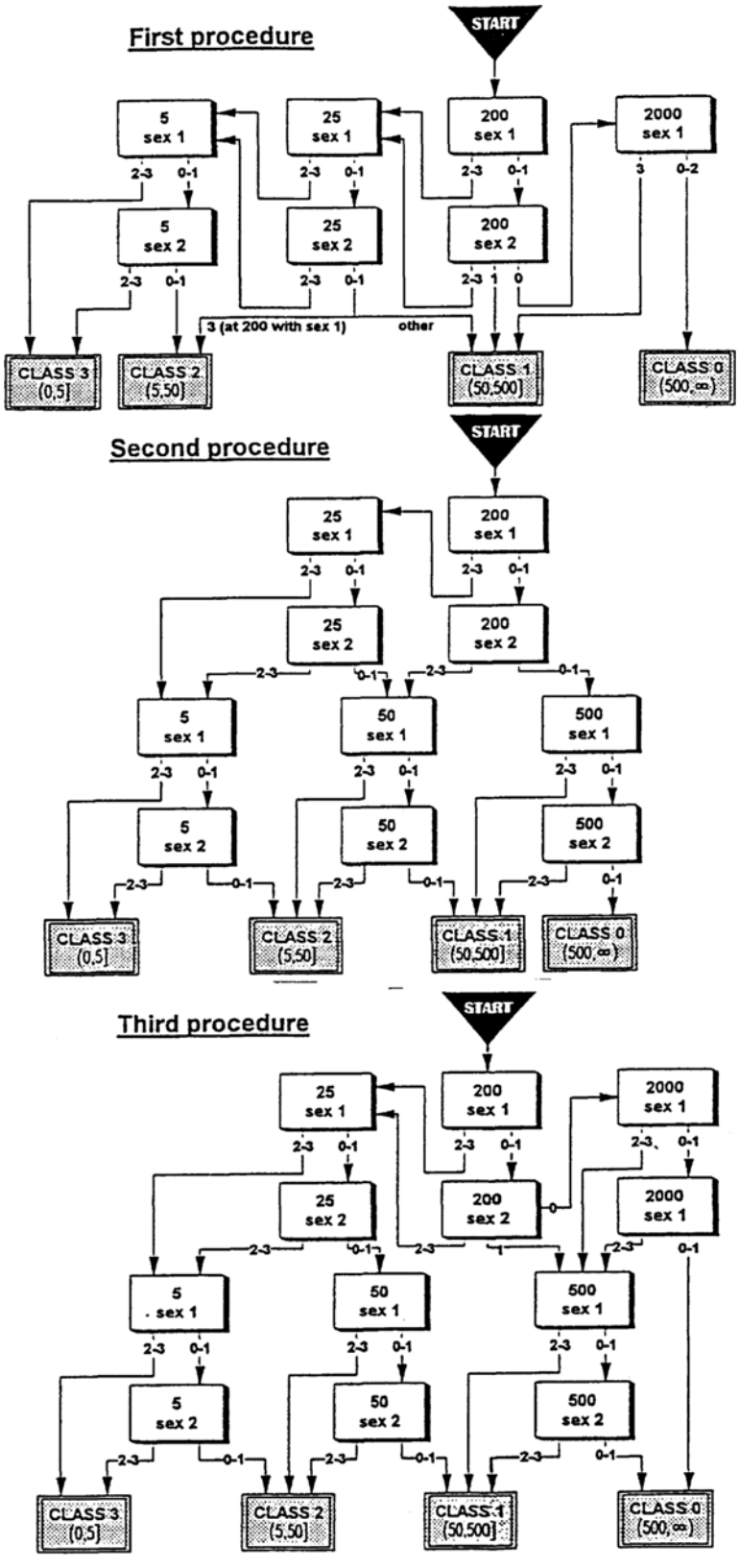


Fig. 5 First, second and third procedure of the ATC method for the classification systems of the EU (solid pesticides) and of the UN (liquids) with three animals per sex and dose level. For abbreviations see Fig. 4

Table 4 Experimental and biometrical results per substance for the classification systems of the EU (solid pesticides) and of the UN (liquids)

Allocated class, number and name of the substance	Procedure (1-3)	Number of laboratories per class								ATC classification (%)				Mean number of animals			
		Class 0		Class 1		Class 2		Class 3		≙ LD_{50}		identical		used		moribund/dead	
		exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.
Class 3																	
1 Aldicarb	1	–	–	–	–	–	0.4	9	8.6	100	96	100	96	9.0	9.7	8.9	8.3
	2	–	–	–	–	–	0.4	9	8.6	100	96	100	96	9.0	9.7	9.0	8.3
	3	–	–	–	–	–	0.4	9	8.6	100	96	100	96	9.0	9.7	8.9	8.3
2 Parathion	1	–	–	–	–	3	1.0	6	8.0	67	89	67	89	10.7	10.0	7.9	8.4
	2	–	–	–	–	3	1.0	6	8.0	67	89	67	89	10.7	10.0	7.9	8.4
	3	–	–	–	–	3	1.0	6	8.0	67	89	67	89	10.7	10.0	7.9	8.4
Class 2																	
3 Thiosemicarbazide	1	–	–	–	–	9	9.0	–	–	100	100	100	100	11.7	12.0	5.7	6.0
	2	–	–	–	–	9	9.0	–	–	100	100	100	100	11.7	12.0	5.7	6.0
	3	–	–	–	–	9	9.0	–	–	100	100	100	100	11.7	12.0	5.7	6.0
4 Indomethacin	1	–	–	2	–	7	8.0	–	1.0	78	89	78	89	10.7	12.1	4.4	6.5
	2	–	–	1	–	8	8.0	–	1.0	89	89	89	89	13.3	12.2	5.7	6.5
	3	–	–	1	–	8	8.0	–	1.0	89	89	89	89	13.3	12.2	5.7	6.5
5 *N*-Phenylthiourea	1	–	–	5	0.6	2	3.2	2	5.2	22	36	56	58	10.7	11.5	5.0	6.9
	2	–	–	4	0.2	3	4.1	2	4.7	33	46	44	52	11.3	11.4	5.4	7.0
	3	–	–	4	0.2	3	3.7	2	5.2	33	41	44	58	12.7	11.9	6.3	7.2
6 Mercury(II)oxide	1	–	–	1	0.1	8	8.9	–	–	89	99	89	99	10.0	11.6	3.7	5.1
	2	–	–	3	0.1	6	8.9	–	–	67	99	67	99	13.7	12.9	5.6	6.1
	3	–	–	3	0.1	6	8.9	–	–	67	99	67	99	14.3	12.9	5.6	6.1
7 Sodium arsenite	1	–	–	–	0.1	9	8.9	–	–	100	99	100	99	9.3	10.2	3.8	4.3
	2	–	–	3	0.4	6	8.6	–	–	67	95	67	95	13.3	12.9	5.3	6.2
	3	–	–	3	0.4	6	8.6	–	–	67	95	67	95	13.3	12.9	5.3	6.2
Class 1																	
8 Bis(tributyltin)oxide	1	–	–	9	6.8	–	2.2	–	–	100	75	100	75	9.0	9.6	3.0	3.1
	2	2	–	7	6.8	–	2.2	–	–	78	75	78	75	11.7	13.1	2.8	4.3
	3	2	–	7	6.6	–	2.3	–	–	78	74	78	74	14.3	14.6	5.0	4.5
9 Acrylamide	1	–	–	2	6.6	7	2.4	–	–	22	73	78	73	9.0	9.4	3.1	2.5
	2	–	–	9	9.0	–	–	–	–	100	100	100	100	14.3	13.9	3.1	2.8
	3	–	–	9	9.0	–	–	–	–	100	100	100	100	14.7	15.1	3.4	2.9
10 Cadmium chloride	1	–	–	7	8.4	2	0.6	–	–	78	93	78	93	9.7	8.9	2.8	2.4
	2	–	–	9	9.0	–	–	–	–	100	100	100	100	13.0	12.0	3.1	3.2
	3	–	–	9	9.0	–	–	–	–	100	100	100	100	15.0	13.8	3.4	3.5
11 Caffeine	1	1	–	7	8.7	1	0.3	–	–	78	97	78	97	9.0	8.5	3.1	2.4
	2	–	–	9	9.0	–	–	–	–	100	100	100	100	10.0	11.0	3.2	3.3
	3	–	–	9	9.0	–	–	–	–	100	100	100	100	13.0	12.8	5.8	4.0
Class 0																	
12 Aniline	1	–	0.4	9	8.6	–	–	–	–	0	4	100	96	9.0	9.0	3.0	3.0
	2	8	8.4	1	0.6	–	–	–	–	89	93	89	93	12.0	11.9	0.3	0.7
	3	8	8.4	1	0.6	–	–	–	–	89	93	89	93	15.0	14.9	3.3	3.6
13 Ferocene	1	1	4.3	8	4.7	–	–	–	–	11	48	89	52	9.0	9.0	2.9	2.4
	2	5	8.9	4	0.1	–	–	–	–	56	99	56	99	11.3	12.0	1.4	0.2
	3	5	8.9	4	0.1	–	–	–	–	56	99	56	99	14.3	15.2	4.3	2.9
14 Sodium salicylate	1	6	6.4	3	2.6	–	–	–	–	67	72	67	72	9.0	9.0	2.0	2.0
	2	9	9.0	–	–	–	–	–	–	100	100	100	100	12.0	12.0	0.0	0.1
	3	9	9.0	–	–	–	–	–	–	100	100	100	100	15.0	15.4	2.0	2.6
15 Acetanilide	1	6	6.5	3	2.5	–	–	–	–	67	72	67	72	9.0	8.7	1.6	1.7
	2	8	7.7	1	1.4	–	–	–	–	89	85	89	85	12.0	11.8	0.1	1.2
	3	9	7.8	–	1.2	–	–	–	–	100	87	100	87	14.0	14.6	1.8	3.1
16 Acetonitrile	1	9	8.3	–	0.7	–	–	–	–	100	92	100	92	9.0	9.0	0.8	1.2
	2	9	8.9	–	0.1	–	–	–	–	100	99	100	99	12.0	12.0	0.0	0.3
	3	9	8.9	–	0.1	–	–	–	–	100	99	100	99	13.3	14.4	0.8	2.2
17 Butylated hydroxyanisole	1	7	8.1	2	0.9	–	–	–	–	78	90	78	90	8.3	8.9	0.7	1.2
	2	8	8.7	1	0.3	–	–	–	–	89	97	89	97	11.7	11.9	0.2	0.5
	3	9	8.8	–	0.2	–	–	–	–	100	98	100	98	13.7	14.2	1.1	2.2
18 *N,N*-Dimethyl-formamide	1	8	8.2	1	0.8	–	–	–	–	89	91	89	91	8.7	8.8	0.4	0.9
	2	9	8.7	–	0.3	–	–	–	–	100	97	100	97	11.7	12.0	0.0	0.6
	3	9	8.9	–	0.1	–	–	–	–	100	98	100	98	13.0	13.5	0.4	1.8
19 Quercetin dihydrate	1	9	9.0	–	–	–	–	–	–	100	100	100	100	9.0	9.0	0.0	0.0
	2	9	9.0	–	–	–	–	–	–	100	100	100	100	12.0	12.0	0.0	0.0
	3	9	9.0	–	–	–	–	–	–	100	100	100	100	12.0	12.0	0.0	0.0
20 Ethylene glycol	1	9	9.0	–	–	–	–	–	–	100	100	100	100	8.3	9.0	0.0	0.0
	2	9	9.0	–	–	–	–	–	–	100	100	100	100	11.3	12.0	0.0	0.0
	3	9	9.0	–	–	–	–	–	–	100	100	100	100	11.3	12.0	0.0	0.0

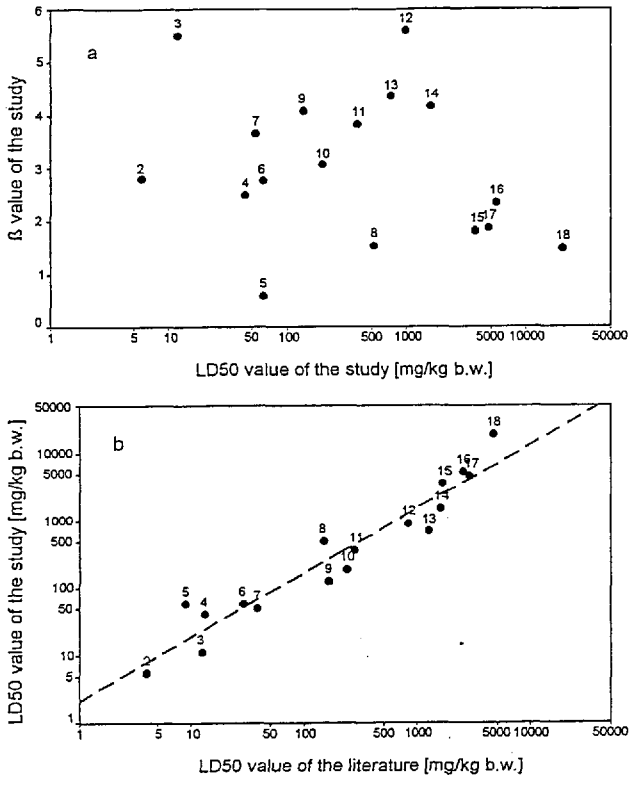


Fig. 6a, b LD_{50} and β values derived from the ring study (**a**) and comparison of the LD_{50} values of the literature and of the ring study with regression line (**b**). For identification of substances see Table 3

Signs of toxicity and autopsy findings

Toxic signs reported as reduced locomotor activity, piloerection, abdominal position, dyspnoea and hunched posture showed the best reproducibility during the time course of observations.

In general, autopsy findings revealed no differences in comparison to the findings reported in the literature or the results of the national validation study. Specific findings on indomethacin (ascites and hydrothorax) and on cadmium chloride (atrophy of the testes) that were reported in the national validation study (Schlede et al. 1992) were corroborated in this study.

Both the signs of toxicity and the autopsy findings are in general agreement with the data obtained in the national validation study (Schlede et al. 1992).

Discussion

Classification

Conventional acute oral toxicity testing – the LD_{50} test- requires at least 20–30 animals. When no substance-related death is anticipated a limit test is carried out with the use of ten animals (five females and five males; OECD 1981a, 1987; Commission of the European Communities 1992a). The numerical values of the LD_{50} test results are mainly used for allocation to defined toxicity classes, the so-called classification system. These allocations in turn result in specific regulatory preventive measures, such as labelling, specific packaging requirements and probable restrictions on use and distribution. The need for conducting LD_{50} tests with high numbers of animals for regulatory purposes or for further toxicological studies is generally not necessary (Zbinden and Flury-Roversi 1981). Instead of determining a numerical value, toxicity classes with the use of few animals can be estimated as a range in which the true LD_{50} is located. This approach is justifiable and meets the aspects of human health protection in a similar or even better manner than LD_{50} testing.

The ATC method was designed and evaluated biometrically and experimentally (Roll et al. 1986, 1989; Schlede et al. 1992; Diener et al. 1994, 1995) to meet the need for human health protection and at the same time to cover aspects of animal welfare. The ATC method allows the use of three fixed starting doses (25, 200 and 2000 mg/kg) and only three animals of each sex per step. For animal welfare reasons and for direct comparison and due to the "blind" conduct of the study, the participants decided to use only the starting dose of 200 mg/kg for all 20 tested substances. The results presented in this paper therefore refer only to this starting dose. Nevertheless, the use of the additional starting doses of 25 or 2000 mg/kg would have produced identical classification results because they are in fact independent of the starting dose (Diener et al. 1994 1995). Classification is possible after having tested the above mentioned doses (with the exception of class limit of 5 mg/kg) or to use additional fixed doses of 5, 50 and 500 mg/kg. The differences between these two possibilities with respect to classification results are considered not to be substantial, with an increase in correct classifications after additional testing for some of the classification systems. This is shown in Table 5 for the evaluation procedures. Thus, the necessity to test the additional dose levels of 5, 50 or 500 mg/kg is mainly given for the 5 mg dose level.

In general, the substances were allocated to one toxicity class or to some extent to an adjacent class (Tables 3 and 4). *N*-Phenylthiourea showed a distribution of results among three classes (Tables 3 and 4). Similar difficulties were encountered with this substance in the international validation study of the fixed dose procedure (Van den Heuvel et al. 1990). However, in the national validation study of the ATC method, five out of six tests resulted in allocation to one class (Schlede et al. 1992). This variability may be explained by the low estimated β of 0.6 for this substance (Fig. 6a) and taking into account that the estimated LD_{50} value of the study deviates significantly from the LD_{50} value given in the literature (Fig. 6b).

The conformity of the interlaboratory results for all commonly used classification systems ranged from 82 to 94% (Table 5) and for individual classification systems from 44% (for *N*-phenylthiourea) to 100% (Tables 3 and 4). A very good agreement with the biometrical results is obvious. This high degree of reproducibility – with the exception of *N*-phenylthiourea – further demonstrates the reliability of the ATC method.

Table 5 Experimental and biometrical results for all classification systems. For abbreviations see Fig. 3

Nr.	Classification system	Procdure	ATC classification (%)								Mean number of animals			
			≙ LD_{50}		more toxic		less toxic		identical		used		moribund/dead	
			exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.	exp.	bio.
1.	EU chemicals,													
2.	EU liquid pesticides	1	78	79	9	18	13	3	84	83	9.1	8.9	2.9	3.2
3.	EU solid pesticides	1	73	81	22	18	5	1	86	87	9.4	9.7	3.1	3.4
4.	UN liquids	2	86	93	5	6	9	1	87	93	11.9	11.9	3.8	3.3
		3	87	93	4	6	9	1	88	94	13.2	13.2	3.8	4.1
5.	UN solids	1	81	87	9	11	10	2	89	88	7.9	8.3	2.3	2.8
		2	82	88	6	10	12	2	88	90	8.9	9.1	2.5	3.0
		3	82	88	6	10	12	2	88	90	9.1	9.3	2.5	3.0
6.	Switzerland	1	63	72	26	25	11	3	82	79	10.9	11.1	3.5	3.7
		2	78	84	9	14	13	3	82	86	13.8	13.6	4.2	4.2
		3	78	83	9	14	13	3	82	86	14.0	13.9	4.3	4.2
7.	US EPA crk	1	93	94	1	5	6	1	94	96	6.9	6.8	2.0	2.4
8.	Japan PDSCA	1	79	90	1	2	20	9	91	91	6.9	6.8	2.0	2.4
9.	Canada WHMIS	1	76	85	19	15	5	0	88	90	6.9	6.8	2.3	2.2
10.	US OSHA	2	90	97	3	3	7	1	90	97	11.0	9.5	3.6	2.8
		3	90	97	3	3	7	0	90	97	12.2	11.7	3.6	3.7
11.	US EPA pesticides	1	70	84	22	15	8	1	85	89	8.0	8.3	2.5	2.5
		2	86	96	3	3	11	1	87	96	11.4	11.2	3.8	3.7
		3	86	96	3	3	11	1	87	96	12.8	12.4	3.9	3.8
12.	US CPSC	1	87	96	6	3	7	1	92	96	8.1	8.3	2.5	2.5
		2	87	96	0	1	13	2	90	96	8.4	8.4	3.7	3.5
		3	89	98	0	2	11	1	93	98	11.2	10.7	3.4	3.5
13.	Canada pesticides	1	75	75	23	24	2	1	88	83	6.9	6.8	2.3	2.3
		2	81	79	16	20	3	1	88	88	8.7	8.5	2.7	2.6
		3	81	79	16	20	3	1	88	88	8.7	8.5	2.7	2.6
14.	Ring study	–	65	65	17	31	19	3	72	81	14.3	14.1	4.1	4.2

Number of experimental animals

The number of experimental animals depends on the true LD_{50} and β of a substance, the choice of the starting dose, on the classification system and the evaluation procedure.

Two extremes are possible: in one case three animals are treated with a dose of 25 mg/kg and two or three animals die; no further testing is necessary, such as for EU chemicals. In a second case (e.g., UN liquids), three animals are treated at each step, also using a starting dose of 25 mg/kg, and the last tested step would be the second 2000 mg/kg dose level with the first procedure and two additional steps at 500 mg/kg with the third procedure (Fig. 5). This would result in the use of 18 animals with the first procedure and six more animals with the third procedure. This kind of testing procedure is very unlikely, but it cannot be dismissed. Under practical conditions and as tested in this international validation study, a starting dose of 200 mg/kg is recommended for animal welfare reasons when there is no information on the substance to be tested.

In general, with the use of the first procedure the smallest number of animals are used. However, with both the second and third procedures often more animals are used without significantly increasing the results of a correct classification (Tables 4, 5).

The lowest mean number of used animals was 6 or less (Table 3, substances 1–3) and the highest mean number was 15 (Table 4, substance 14). This is a reduction of 80 and 50% in the number of used animals when compared to the LD_{50} test (30 animals) or a corresponding reduction of 70 and 25% when in an LD_{50} test 20 animals are used. On average, two to four steps with the use of 6–12 animals are required for a reliable allocation to the commonly used classification systems. This is shown as mean numbers in Table 5 which ranged from 6.9 (systems 7, 8, and 13) to the highest mean number of 14.0 for system 6. Also, when the limit test with 2000 mg/kg is carried out with the ATC method, six animals are used instead of ten animals with the classical limit test. Altogether, this is an impressive reduction in the number of animals used in comparison to the LD_{50} test and the classical limit test.

Number of moribund/dead animals

As mentioned for the number of experimental animals, the number of moribund/dead animals also depends on the true LD_{50} and β of a substance, the choice of the starting dose, on the classification system and on the evaluation procedure.

The number of moribund/dead animals per step ranges from zero to three. In this study for individual classification systems the lowest mean number was zero (substances 19 and 20 in Tables 3 and 4) and the highest mean number was 9.0 (see substance 1 in Table 4). For all commonly used classification systems the lowest mean number was 2.0 (Table 5 for systems 7 and 8) and the highest mean number was 4.3 animals (Table 5 for system 6). In contrast, in an LD_{50} test at least 10 to more than 15 dead animals would be expected.

Biometrical results

All experimental results were also evaluated biometrically in order to validate further the ATC method. In general there is excellent agreement between the biometric and experimental results. Deviations were considered to be negligible, especially with respect to the mean numbers of animals (used and moribund/dead). The variability in the classification results is mainly caused, in addition to random effects, by the different LD_{50} values of some substances, e.g. *N*-phenylthiourea, indomethacin and bis(tributyl)tinoxide. Here the experimental results are based on the estimated LD_{50} values of the ring study, while the biometrical results are based on the LD_{50} values of the literature (Fig. 6b).

Detailed descriptions of the biometrical considerations of the ATC method have been published (Diener et al. 1994, 1995).

Signs of toxicity and autopsy findings

Another important part of the ATC method is the careful observation of toxic signs, their onset and their duration after treatment. This gives additional information on the toxicity profile of a substance that could be used for other toxicological studies. The further evaluation of all these data, including autopsy findings, is now in progress and is sponsored by a grant from the German government. The object of this evaluation is to obtain, if possible, criteria that would result in the avoidance of death of animals as an endpoint after a single oral treatment.

Final remarks

In conclusion, it can be emphasized that the ATC method is a validated alternative to the LD_{50} test. This method allows a substance to be ranked to all commonly used classification systems in a similar or even better manner than the LD_{50} test. It uses substantially fewer animals and results in significantly fewer moribund/dead animals. Thus the ATC method also subjects fewer animals to pain and distress. Under the aspects of animal welfare this is considered to be a great achievement. Acceptance of the ATC method as an official test guideline of the OECD is expected in summer 1995.

At present, international efforts are being made to harmonize the different classification systems currently in use. Also, if international harmonization with the use of only one classification system can be reached, the ATC method will be a reliable alternative to the LD_{50} test. The ATC method has been successfully tested in two large ring studies. In addition, it is based on biometric evaluations predicting the accuracy of a correct classification within adequately separated fixed dose levels.

Acknowledgements The authors are greatly indebted to the participants for their co-operation and the financial assistance provided by their institutes. This study was supported by a grant from the German government (BMFT, grant No. 0318956A) and it was conducted under the patronage of the OECD, Paris.

References

Bruce RD (1985) An up-and-down procedure for acute toxicity testing. Fundam Appl Toxicol 6: 151–157

Commission of the European Communities (1992a) Annex to Commission Directive 92/69/EEC of 31 July 1992 adapting to technical progress for the seventeenth time Council Directive 67/548/EEC on the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances: Annex V, B 1 Acute toxicity (oral). Off J Eur Comm L 383: 110–112

Commission of the European Communities (1992b) Annex to Commission Directive 92/69/EEC of 31 July 1992 adapting to technical progress for the seventeenth time Council Directive 67/548/EEC on the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances: Annex V, B.1 bis Acute toxicity (oral) – fixed dose method. Off J Eur Comm L 383: 113–116

Deichmann WB, LeBlanc TJ (1943) Determination of the approximate lethal dose with about six animals. J Ind Hyg Toxicol 25: 415–417

Diener W, Siccha L, Mischke U, Kayser D, Schlede E (1994) The biometric evaluation of the acute-toxic-class-method (oral). Arch Toxicol 68: 599–610

Diener W, Mischke U, Schlede E, Kayser D (1995) The biometrical evaluation of the OECD modified version of the acute-toxic-class method (oral). Arch Toxicol 69: 729–734

Kennedy GL, Ferenz RL, Burges BA (1986) Estimation of acute oral toxicity in rats by determination of the approximate lethal dose rather than the LD_{50}. J Appl Toxicol 6: 145–148

Lorke D (1983) A new approach to practical acute toxicity testing. Arch Toxicol 54: 275–287

OECD (1981a) Guidelines for testing of chemicals No. 401: acute oral toxicity. OECD, Paris

OECD (1981b) Decision of the council concerning mutual acceptance of data in the assessment of chemicals. Annex 2, OECD Principles of Good Laboratory Practice. OECD, Paris

OECD (1987) Guidelines for testing of chemicals no. 401: acute oral toxicity. OECD, Paris

OECD (1992) Guidelines for testing of chemicals no. 420: acute oral toxicity – fixed dose method. OECD, Paris

Roll R, Höfer-Bosse Th, Kayser D (1986) New perspectives in acute toxicity testing of chemicals. Toxicol Lett Suppl 31: 86

Roll R, Riebschläger M, Mischke U, Kayser D (1989) Neue Wege zur Bestimmung der akuten Toxizität von Chemikalien. Bundesgesundheitsblatt 32: 336–341

Schlede E, Mischke U, Roll R, Kayser D (1992) A national validation study of the acute-toxic-class method – an alternative to the LD_{50} test. Arch Toxicol 66: 455–470

Van den Heuvel MJ, Clark DG, Fielder RJ, Koundakjian PP, Oliver GJA, Pelling D, Tomlinson NJ, Walker AP (1990) The international validation of a fixed-dose procedure as an alternative to the classical LD_{50} test. Food Chem Toxicol 28: 469–482

Yamanaka S, Hashimoto M, Tobe M, Kobayashi K, Sekizawa J, Nishimura M (1990) A simple method for screening assessment of acute toxicity of chemicals. Arch Toxicol 64: 262–268

Zbinden G, Flury-Roversi M (1981) Significance of the LD_{50}-test for the toxicological evaluation of chemical substances. Arch Toxicol 47: 77–99