RIVM report 601516014/2006

Factsheets for the (eco)toxicological risk assessment strategy of the National Institute for Public Health and the Environment
Part VI

J.W.A. Scheepmaker (ed.)

Contact: J.W.A. Scheepmaker
Expert Centre for Substances
Jacquelne.Scheepmaker@rivm.nl

This investigation has been performed by order and for the account of the Board of Directors of the National Institute for Public Health and the Environment, within the framework of project 601516, Kennislacunes risicobeoordeling.

RIVM, P.O. Box 1, 3720 BA Bilthoven, telephone: +31 302749111; fax: +31 302742971

Authors

Chapter 1: P.J.C.M. Janssen
Chapter 2: W. ter Burg
Chapter 3: S.M.G.J. Pelgrom
Chapter 4: R. Fleuren

Abstract

Factsheets for the (eco)toxicological risk assessment strategy of the National Institute for Public Health and the Environment (RIVM)- Part VI

This report contains four factsheets describing risk assessment methods used at the Centre for Substances and Integral Risk Assessment (SIR) and the Expert Centre for Substances (SEC) of the National Institute for Public Health and the Environment (RIVM). The first three factsheets concern human risk assessment, and the fourth, environmental risk assessment.

The first factsheet, Relevance of changes in selected blood biochemical parameters, deals with biochemical blood parameters, such as the bilrubin level, as related to liver damage in test animals. The toxicological significance of increases in these parameters is evaluated here.

The second factsheet, Strategy for quantitative risk assessment pertaining to skin sensitisation, describes the strategy using the Local Lymph Node Assay (LLNA), a method to assess the potential of substances to bring on hypersensitive reactions. This hypersensitivity is called skin sensitisation. The LLNA test method provides insight into the relationship between the dose and its effect. A strategy for the new approach – quantitative risk assessment – is proposed in relation to skin sensitisation.

The third factsheet deals with the Leydig Cell tumour, one of the three types of tumours that can occur in testicles. The factsheet discusses the relevance of Leydig cell tumour in animals, caused by exposure to chemicals, to human risk assessment.

The fourth factsheet, Proposal for interpreting leaching study data for wood preservatives (biocides), is important for the risk assessment of chemical substances in the environment. The leaching rate of the substance from the wood is crucial in the risk assessment of wood preservatives, since the substance ends up in the environment by means of this route. This factsheet evaluates a few different models and proposes an efficient and simple method for the otherwise complicated establishment of this leaching rate.

Key words: biochemical parameters, liver damage, local lymph node assay, LLNA, skin sensitisation, Leydig cell tumour, testicle tumour, wood preservatives, leaching.

Rapport in het kort

Factsheets voor de (eco)toxicologische risicobeoordelingsstrategie van het Rijksinstituut voor Volksgezondheid en Milieu (RIVM) - Deel VI

Dit rapport bundelt vier ‘factsheets’ over methodieken voor de risicobeoordeling van stoffen bij het Centrum voor Stoffen en Integrale Risicobeoordeling (SIR) en het Stoffen Expertise Centrum (SEC). De eerste drie factsheets behandelen onderwerpen die betrekking hebben op humane risicobeoordeling, de laatste gaat over risicobeoordeling in het mileu.
De eerste factsheet Relevance of changes in selected blood biochemical parameters, gaat in op biochemische bloedparameters die gerelateerd zijn aan leverschade (zoals bv. bilirubine gehalte) in proefdieren. De factsheet evalueert de toxicologische betekenis van toenames in deze parameters.
De tweede factsheet Strategy for quantitative risk assessment for skin sensitisation using the Local Lymph Node Assay (LLNA), gaat over een methode die de potentie van stoffen kan bepalen om overgevoeligheid bij huidcontact te veroorzaken. Het gevoelig raken voor stoffen wordt huidsensibilisatie (skin sensitisation) genoemd. De gebruikte LLNA testmethode geeft inzicht in de relatie tussen de hoogte van de dosis en het uiteindelijke effect. In deze factsheet wordt een strategie voorgesteld voor een kwantitatieve risicobeoordeling van huidsensibilisatie, wat tot nu toe niet gebruikelijk is.
De derde factsheet gaat over Leydigcel tumoren. De Leydig cel tumor is één van de drie typen tumoren die in testikels kunnen voorkomen. De factsheet bediscussieert of de Leydig cel tumor als gevolg van blootstelling aan chemische stoffen bij dieren, relevant is voor de humane risicobeoordeling.
De vierde en laatste factsheet Proposal for the interpretation of leaching study data for wood preservatives (biocides), heeft waarde voor de risicobeoordeling van stoffen in het milieu. Bij de risicobeoordeling van houtverduurzamingsmiddelen speelt de snelheid waarmee het middel uit het hout verdwijnt (uitloging) een cruciale rol, aangezien het middel via deze route in het milieu belandt. Aan de bepaling van deze snelheid zitten momenteel diverse haken en ogen. Deze factsheet behandelt een aantal verschillende modellen en stelt een efficiënte en simpele aanpak voor.

Trefwoorden: biochemische parameters, leverschade, lokale lymfeknoop test, LLNA, huidsensibilisatie, Leydigcel tumor, testikeltumoren, houtverduurzamingsmiddelen, uitloging.

Preface

This report was written within the framework of the project ‘Kennislacunes risicobeoordeling’ (‘*Knowledge gaps in risk assessment’*). The factsheets presented in this report have been reviewed by members of the peer review groups of the Centre for Substances and Integral Risk Assessment (SIR) and the Expert Centre for Substances (SEC), and in some cases additional experts were consulted. The following persons are acknowledged for their contribution: M.E. van Apeldoorn, A.J. Baars, R.B. Beems, J. van Benthem, E. de Boer, J.G.M. van Engelen, B. Hakkert, A.G.A.C. Knaap, F.X.R. van Leeuwen, J.B.H.J. Linders, H. van Loveren, J.J.A. Muller, M.E.J. Pronk, M.T.M. van Raaij, A.P. Verschoor, P.W. Wester, G. Wolterink and M. van Zijverden.

Contents

Samenvatting

In dit rapport worden vier factsheets gepresenteerd die worden gebruikt voor de beoordeling van stoffen bij het Centrum voor Stoffen en Integrale Risicobeoordeling (SIR) en het Stoffen Expertise Centrum (SEC) van het Rijksinstituut voor Volksgezondheid en Milieu (RIVM).

Factsheet Relevance of changes in selected blood biochemical parameters
In proefdierstudies worden standaard monsters genomen waarin de hoogte van bepaalde biochemische parameters worden bepaald. Deze factsheet evalueert de toxicologische betekenis van eventuele toenames in deze biochemische parameters. Deze factsheet beperkt zich tot de biochemische parameters die betrekking hebben op leverschade.

Factsheet Strategy for quantitative risk assessment for skin sensitisation using the Local Lymph Node Assay (LLNA)
Onze huid kan na contact met chemische stoffen overgevoelig raken. Het gevoelig raken voor stoffen wordt huidsensibilisatie (skin sensitisation) genoemd.
Tegenwoordig wordt de LLNA testmethode steeds vaker gebruikt. Deze methode geeft inzicht in de relatie tussen de hoogte van de dosis en het uiteindelijke effect. In deze factsheet wordt een strategie opgezet waarmee kwantitatieve gegevens kunnen worden beoordeeld en gebruikt voor de risicobeoordeling.

Factsheet Leydig Cell tumour
De Leydig cel tumor is één van de drie typen tumoren die in testikels voorkomen. De factsheet Leydig cel tumoren geeft een raamwerk voor het vaststellen van de relevantie van het vóórkomen van tumoren bij dieren voor de humane risicobeoordeling. De factsheet beschrijft het mechanisme achter de vorming van dit type tumoren. Daarnaast wordt ingegaan op de overeenkomsten en verschillen tussen dier en mens voor wat betreft de anatomie en regulering van het hypofyse-hypothalamus-testikel hormoonsysteem. De factsheet definieert de omstandigheden waaronder Leydigcel tumoren kunnen worden beschouwd als niet-relevant voor de humane risicobeoordeling.

Factsheet Proposal for the interpretation of leaching study data for wood preservatives (biocides)
Bij de risicobeoordeling van houtverduurzamingsmiddelen speelt de snelheid waarmee het middel uit het hout verdwijnt een cruciale rol, aangezien het middel vanuit het hout in het milieu belandt. Aan de bepaling van deze snelheid zitten momenteel zowel experimentele als modelmatige haken en ogen. Deze factsheet behandelt een aantal verschillende modellen en stelt een efficiënte en simpele aanpak voor die gebruikt kan worden in de risicobeoordeling als onderdeel van de toelatingsprocedure voor deze stoffen.

Summary

This report presents four factsheets that are used for the risk assessment of substances at the Centre for Substances and Integral Risk Assessment (SIR) and the Expertise Centre for Substances (SEC) of the National Institute for Public Health and the Environment (RIVM).

Factsheet Relevance of changes in selected blood biochemical parameters
In animal experiments clinical chemistry on blood is done in a standard way. In the samples the contents of certain biochemical parameters is determined. This factsheets evaluates the toxicological meaning of increases in certain biochemical parameters. The factsheet restricts the biochemical parameters to those that are involved in liver damage.

Factsheet Strategy for quantitative risk assessment for skin sensitisation using the Local Lymph Node Assay (LLNA)
Dermal exposure to chemicals can result in skin senstisation leading to a hypersensitive state. The LLNA test method is increasingly being used. This method gains an insight in the relation between the dose and its effect. In this factsheet a stratety is being proposed for a quantitative risk assessment of skin sensitization.

Factsheet Leydig Cell tumour
The Leydig cell tumour is one of the three types of tumours that can occur in testicles. The factsheet gives an outline for the determination of the relevance of the occurrence of tumours in animals for the human risk assessment. The facsheet describes the mechanism behind the induction of this type of tumour. Also, the factsheet addresses the similarities and the differences between animals and humans regarding the anatomy and regulation of the pituitary – hypothalamus-testicle hormone system. The factsheet defines the cases in which Leydig cell tumours can be considered to be not relevant for the human risk assessment.

Factsheet Proposal for the interpretation of leaching study data for wood preservatives (biocides)
In the risk assessment of wood preservatives the leaching rate of the substance from the wood is crucial, since the substance end up in the environment through this route. The establishment of this leaching rate is complicated. This factsheet evaluates a few different models and proposes an efficient and simple method which can be used as part of the registration procedure of these substances.

Introduction

One of the main tasks of the Expert Centre for Substances (SEC) and the Centre of Substances and Risk Assessment (SIR) of the National Institute for Public Health and the Environment (RIVM) is to assess the risk of compounds for public health and the environment. The availability of adequate and up-to-date risk assessment methods is of the highest importance to fulfil this task. Some of these methods follow international guidance, but many have been developed within the RIVM during the process of evaluation. These risk assessment methods are not rigid procedures but can be adapted based on new/developing scientific information, possibly triggered by questions from policy makers or by developments in (inter)national organisations.

For specific problems or gaps in the assessment of (eco)toxicological effects, 'factsheets' are written by employees of SEC and SIR in co-operation with experts. These factsheets describe the current assessment strategies of SEC and SIR, and their main aim is to provide a transparent and accessible guidance for issues that are not covered by regular guidance documents. After adoption of the factsheet by the advisory board and the head of the laboratories SEC or SIR all employees of SEC and SIR have to follow the risk assessment method described in the factsheet.

In 2001, the first eight factsheets were published in an RIVM report[1], followed by similar reports in 2002, 2003, 2004 and 2005[2,3,4,5]. These reports can be downloaded via http://www.rivm.nl/bibliotheek/index-en.html. The present report contains four factsheets that were produced in 2005 and 2006 by SIR and SEC:

1. Relevance of changes in selected blood biochemical parameters for the assessment of liver damage
2. Strategy for quantitative risk assessment for skin sensitisation using the Local Lymph Node Assay (LLNA)
3. Leydig Cell tumour
4. Proposal for the interpretation of leaching study data for wood preservatives (biocides)

We hope that by publishing these factsheets, the risk assessment methods followed by RIVM/SEC and RIVM/SIR will become more transparent. The authors of each factsheet have tried to describe the state of the art of their subject.
Remarks, omissions or supplementary information will be appreciated and can be send to Jacqueline.Scheepmaker@rivm.nl and will be passed on to the responsible authors.

[1] Luttik R, Van Raaij MTM, editors. Factsheets for the (eco)toxicological risk assessment strategy of the National Institute for Public Health and the Environment (RIVM). Bilthoven: National Institute for Public Health and the Environment; 2001. Report no. 601516007.
[2] Luttik R, Pelgrom SMJG, editors. Factsheets for the (eco)toxicological risk assessment strategy of the National Institute for Public Health and the Environment. Part II. Bilthoven: National Institute for Public Health and the Environment; 2002. Report no. 601516009.
[3] Luttik R, Van Raaij MTM, editors. Factsheets for the (eco)toxicological risk assessment strategy of the National Institute for Public Health and the Environment. Part III. Bilthoven: National Institute for Public Health and the Environment; 2003. Report no. 601516010.
[4] Smit CE, Van Raaij MTM, editors. Factsheets for the (eco)toxicological risk assessment strategy of the National Institute for Public Health and the Environment. Part IV. Bilthoven: National Institute for Public Health and the Environment; 2004. Report no. 601516012.
[5] Scheepmaker, JWA, Smit CE, Van Raaij MTM, editors. Factsheets for the (eco)toxicological risk assessment strategy of the National Institute for Public Health and the Environment. Part V. Bilthoven: National Institute for Public Health and the Environment; 2004. Report no. 601516013.

1. Relevance of changes in selected blood biochemical parameters for the assessment of liver damage

Factsheet FSV/017, date 01/03/2006

Author: P.J.C.M. Janssen

Contents

1.1 Problem definition

In animal toxicity studies changes in blood biochemistry parameters (e.g. ASAT, ALAT) may be observed. Sometimes these changes are not accompanied by clear histopathological changes or other effects. Whether or not such changes should be viewed as adverse, is to be evaluated in the present factsheet. Only limited literature review was possible in the available time and consequently not all blood biochemistry parameters could be dealt with. Those parameters (from the core battery recommended by the OECD) most intimately linked to liver function were selected for the present factsheet.

1.2 Introduction

In animal experiments clinical chemistry on blood (blood biochemistry) is done on one or several occasions throughout the test period. International regulatory guidelines by OECD, US-EPA, US-FDA etc. provide recommendations on the exact parameters to be determined. The core battery of parameters includes the following enzyme activities: alanine aminotransferase (ALAT or ALT), aspartate aminotransferase (ASAT or AST), alkaline phosphatase (AP or ALP) and gamma-glutamyl transferase (GGT). The primary aim of these determinations is detection of liver damage. However, the activity of these enzymes is widespread among tissues and some patterns of changes reflect damage to other organs rather than to the liver. Another point to be noted is that these enzyme activities are not the only parameters relevant for assessing hepatic function/damage; other blood-biochemical parameters such as bilirubin, bile acids, lipids and plasma proteins and even prothrombin time have to be considered as well. Further, specific tests for liver function exist and when applied in particular cases provide additional diagnostic information on liver function.

The present factsheet will deal with the above enzyme activities only in reference to possible liver damage. First regular liver functions will briefly be outlined. Then the biological variation in enzyme actvities will be adressed including the concept of reference values for these parameters. Finally conclusions will be drawn as to the risk assessment strategy at RIVM on the point of enzyme activities. In order to provide a somewhat more complete picture with regard to liver damage, bilirubin and prothrombin time are also considered in the present factsheet.

A large literature exists on animal and human blood biochemistry. A major source for the present fact sheet was the book Animal Clinical Chemistry, A Primer for Toxicologists by E.O. Evans (ed.) (1996). In addition important information has resulted from human clinical practice, such as embodied in the Guidelines for performace of laboratory tests of liver function and injury of the US National Academy of Clinical Biochemistry (Dufour et al., 2000a, 2000b). On the specific issue of the toxicological interpretation of blood biochemistry results from animals studies, which is the subject of the present fact sheet, only scant information was found. International bodies such as WHO or OECD or national bodies such as US-EPA, ATSDR apparently do not provide guidelines on this subject. The limited literature review that is possible within the present project should be seen as an initial exploration only and the resulting guidelines for dealing with blood biochemistry changes are inevitably of a general and pragmatic nature.

1.3 The liver

1.3.1 Liver functions

In the liver a wide variety of processes occurs. They include (Woodman, 1996):

- synthesis of almost all plasma proteins;
- protein catabolism (formation of urea which is excreted via the kidneys);
- synthesis of carbohydrates from fats and proteins;
- synthesis of ATP from glucose (Embden-Meyerhoff);
- storage of carbohydrates as glycogen;
- contributes to regulation of blood glucose;
- synthesis of lipids such as cholesterol and trigylcerides;
- haem synthesis;
- synthesis of cytochrome P450 detoxifying enzymes;
- synthesis of insulin-like growth factors;
- detoxification of lipid-soluble molecules with subsequent excretion into the bile or into the bloodstream for transport to kidneys.

1.3.2 Liver disease

Liver injury and its diagnosis (for which enzyme activities are a tool) represents a complex field of study. An important distinction is that between predictable liver lesions that show good dose- and time-dependence and cross-species consistency (type I lesions) and unpredictable, idiosyncratic liver lesions (type II). In Health Canada (2004) various forms of liver disease based on experience with human drugs, are discussed.

Intrinsic hepatotoxicity is predictable, dose- and time-dependent and occurs in most, if not all, subjects exposed to appropriate doses of the causative substance. The lesions are usually readily reproducible in animals although different species may have significantly different susceptibility. The time to onset is generally very short. Some drugs, however, such as heparin, statins and nicotinic acid cause increases in transaminases after a few days of treatment but they do not cause significant liver damage. The mechanism here is unknown but the characteristics suggest a biochemical effect. The time to onset is generally short but longer than for intrinsic toxins.

- Cholestatic injury (blockage of bile excretion) is characterized by an increase in conjugated bilirubin and alkaline phosphatase in the absence of hepatic necrosis. It can have characteristics of either intrinsic toxicity or an idiosyncratic reaction.
- Mitochondria hepatotoxicity is the inhibition of mitochondrial DNA synthesis leading to liver failure. This type of toxicity can either be intrinsic or idiosyncratic. Elevated transaminases may occur.
- Most drugs that have had to be withdrawn because of liver toxicity have the characteristics of idiosyncratic liver necrosis. Better ways of detecting and differentiating these reactions from the other types of liver toxicity are needed. Most of these reactions probably are immune-mediated.
- As will be evident from the above description, enzyme activities provide only a limited view on what may be a complex disease process and as a consequence residual uncertainty will frequently remain as to their interpretation, especially so when there are no concomitant histopathological effects.

1.4 Major biomarkers of liver damage

1.4.1 Enzyme activities

The determination of activities of ASAT, ALAT, AP and GGT in animal experiments has been derived from human clinical medicine. As explained by Woodman (1996), plasma enzyme activities offer only an indirect reflection of tissue damage. The rate of synthesis and plasma clearance are important factors in addition to leakage due to toxic damage. The wide tissue distribution of enzymes makes organ specificity a difficult goal. The advantage of enzyme activities, however, is that they represent a sensitive measure: due to the great difference between normal plasma and normal intracellular levels, the blood levels can rise rapidly as a result of enzyme loss from a relatively small number of cells.

Other enzyme activities than those mentioned above, are also used for identifying liver toxicity. Glutamate dehydrogenase (GLDH) activity is a marker of hepatocellular damage, as is the well-known enzymes lactate dehydrogenase (LDH) and ornithine carbamoyl transferase (OCT). Sorbitol dehydrogenase (SDH) is another such marker. GLDH and OCT may also indicate mitochondrial damage. In the present factsheet only the enzyme activities ASAT, ALAT, AP and GGT will be dealt with.

ALAT and ASAT
As already stated, ASAT and ALAT are widely distributed in cells throughout the body. ALAT is present primarily in cytosol of cells and to a lesser extent in mitochondria. ASAT shows the reverse pattern, i.e. high activity in mitochondria and lower activity in cytosol. Both ALAT and ASAT are present in many tissues, including heart, kidneys, muscles and brain, but levels in liver are highest, hence their use as markers of liver damage. In humans ALAT and ASAT activities in liver are about 7000 and 3000 times that in blood serum (Dufour et al. 2000a). ALAT and ASAT are the prime diagnostic markers for hepatocellular damage. ALAT is the most universal marker of hepatocellular injury across species whereas ASAT is mostly used only in conjunction with ASAT. ASAT activity being highest in mitochondria, its elevation may indicate specific mitochondrial damage in hepatocytes. Due to their short half-lives, presence of ALAT and ASAT in plasma is thought to reflect recent cell damage.

In human medicine the term ‘transaminitis’ has been applied to mild elevations of ALAT and ASAT that occur without other clinical laboratory abnormalities in asymptomatic individuals. When observed after drug exposure these small elevations are not interpretable, Amacher (1998) points out. In human medicine the range of normal values usually is defined as the 97.5 (or 95) percentile cut-off in a population without known disease. In specific groups of patients not known to have liver disease percentages of abnormally testing individuals have been reported to be up to 8% or even 25-30% (Amacher, 1998).

AP and GGT
AP is usually referred to as a single enzyme but actually the common analytical assays cover a variety of alkaline phosphatases. AP is widely distributed in tissues, occurring in the border membranes of the bile canaliculi and on sinusoidal surfaces of the liver, the intestinal mucosa, the osteoblasts of bone, the renal proximal tubules, the placenta and mammary glands. In most species age-related changes of osseous AP are observed reflecting periods of bone growth. Thus, levels early in life are much higher than those at adult age. The distribution of AP in

specific organs shows variety across species. AP is the standard diagnostic marker for cholestasis (blockage of bile excretion). It however has shortcomings in performing this function. The low activity in rat and cat liver, the high intestinal component in rat plasma and the variability in primate plasma activities do not make it an ideal choice.

Alternative enzymes have been used as markers for hepatobiliary damage, most frequently gamma glutamyl transferase (GGT). This enzyme is found primarily in brush border cells of the renal convoluted tubules and on the canicular surfaces of the hepatic parenchymal cells. Despite its relatively high tissue concentrations in the kidney plasma GGT does not appear to alter following renal injury. In hepatic studies plasma GGT is a marker for cholestasis even in rats where plasma GGT levels are normally very low (<2 IU/litre). Even more specifically, GGT in this species has been reported as an indicator of bile duct lesions. In dogs and cats, similarly as in rats, GGT activity in blood is very low (Woodman 1996). In humans, GGT is a well-established marker for cholestasis, considered slightly more sensitive than AP. Lack of specificity, however, reduces its value in humans (in several diseases increased GGT-activities in serum occur without the cause being known).

1.4.2 Bilirubin and prothrombin time

Bilirubin is the breakdown product of haem, the porphyrin-part of hemoglobine. The liver is responsible for both bilirubin conjugation with glucuronic acid and its excretion. The most common measure is total bilirubin but conjugated and unconjugated bilirubin can also be measured seperately (and used for differential diagnosis of obstructive and haemolytic jaundice). In humans bilirubin levels above 70 μmol/litre lead to the clinical picture of jaundice. Cholestasis in humans may result in plasma levels above 340 μmol/litre. Yellow pigmentation is visible in separated plasma already at 30 μmol/litre. Dogs and rats – but not monkeys – have much lower levels of bilirubin in their blood and small increases are significant indicators of hepatic damage. Although increased total bilirubin indicates hepatobiliary damage, the increases should be evaluated critically because extrahepatic factors also influence bilirubin concentrations. Additional indicators of cholestasis may be needed at this point.

Prothrombin time (PT) measures the time for blood plasma to clot after addition of several clotting factors, one of which is prothrombin. Clotting factors are proteins synthesized in the liver and increased PT may point to liver injury. In human medicine PT is a well-established biomarker for hepatic failure. Specifically for acute hepatic injury PT has even been described as the best indicator (along with bilirubin) (Dufour et al., 2000b). Nevertheless changes in PT may also indicate non-hepatic injury, not only specific effects on blood clotting but also digestive disease.
PT as a measure for liver injury in animal toxicity experiments, however, has not been addressed extensively in literature. Woodman (1996) for instance not even lists PT as a measure for hepatotoxicity in animal studies.

1.5 Normal variation in values and interpretation of study results

The variation in clinical biochemistry parameters as encountered in animal and human study results comprises (following Robinson and Evans, 1996):

- biological variation between species, sexes, ages, nutritional conditions, health conditions;
- analytical imprecision (connected with the technical conduct of studies);
- methodological variation (due to differences in study design);
- interactions between the above three.

Enzyme activities are expressed mostly as International Units: one unit is the amount of enzyme that will catalyze the transformation of 1 μmol of substrate per minute. The activities are expressed as IU/L of mIU/ml. The actual value found depends on the biochemical conditions under which the assay is carried out:

- identity and pH of buffer;
- identity and concentration of substrate;
- temperature;
- presence of activators;
- specimen storage.

National and international recommendations on these point vary, which hampers comparisons between studies.

Another factor to influence the result of enzyme activity determinations is the conditions under which blood is collected, i.e. after feeding or fasting, whether or not anaesthetic is used, whether animals are bled randomly and the site of blood collection.

1.6 Normal values, reference values

Normal enzyme activities vary widely between species, sexes and individuals. Even within individuals wide variations occur. ALAT in humans, for instance, shows 45% variation during the day (highest in afternoon) and 10-30% from day to day. Normal bilirubin levels have a considerably smaller range than enzyme activities and the same goes for prothrombin time. In a rat study Carakostas and Banerjee (1990) studied biological and analytical variation in blood biochemical parameters in a total number of 120 rats (60m, 60f) with blood sampling after 3, 6, 9 and 12 weeks. Components of variation in results were expressed as percentage of mean for analytical variation and for intra- and inter-animal variation respectively.

	Components of variation (as % of mean)				r-Ratio[1)]
	Mean	Analytical	intra-animal	Inter-animal	
ALAT	41.22	19.8	7.8	16.4	0.23
ASAT	71.92	24.1	7.9	16.9	0.22
AP	342.1	10.9	17.5	26.2	0.45
GGT-males	0.97	91.2	79.7	0.0	ND
GGT-females	1.34	80.1	60.0	9.0	42.74
Bilirubin-males	0.27	25.7	34.2	3.7	64.66
Bilirubin-females	0.24	28.6	36.7	0.0	ND

1) r-ratio is intra-animal/inter-animal variance

As the authors point out, the large intra-animal variation in GGT and bilirubin probably was due to between-day analytical variation. The latter variation usually is high in rat measurements due to the low serum levels in this species (low analytical precision in rat reference range). As to the r-ratios, note that values smaller than unity indicate variation between animals is greater than within animals. For parameters with low r-ratios, population-based reference ranges (see below) may not be reliable indicators because individuals may have a much narrower range of normal results than the population reference range. Carakostas and Banerjee (1990) point out that particularly ALAT and ASAT have low r-ratios and that therefore for these parameters comparison with pre-test values may be preferable. More in general they stress that protocols as used in typical safety assessment studies have important practical limitations dictated by costs, desired levels of animal use, regulatory guidelines and requirements for other parts of safety assessment studies. Much more elaborate studies would be needed to account for the differences in biological and analytical variations in each blood biochemical test parameter (Carakostas and Banerjee, 1990).

Whithin medical practice judgement on blood biochemistry makes use of the concept of the reference interval, which indicates the range of values observed in healthy individuals. Typically this range is described by the upper and lower reference limits, which are commonly defined as the upper and lower and 95-percentiles respectively of results from healthy individuals tested under specific conditions (for example, fasting, drawn in the morning, with the patient seated for 10–15 min). Reference intervals may be established for different groups, such as for men and women or for children and adults. Partitioning is necessary when data show that results are significantly different between particular subgroups. Although individual laboratories seldom perform extensive studies to establish reference limits, the validity of the limits used must be verified by testing a small number of healthy individuals to assure that the reference limits suggested in studies performed by the manufacturers of methods or reagents or published in the literature are acceptable for the population tested by the laboratory. Within medical practice upper reference limits have been established for cholesterol and glucose but for enzyme activity tests for liver injury similar values are yet to be developed. For developing more universally usable reference limits for enzyme activites harmonization in methods between laboratories is required.

The concept of reference intervals has not been developed as rigorously within experimental animal toxicology. Tables of normal values for experimental values for different animal laboratory species have been published but these are meant only as rough guidelines. These tables however do make clear that parameters of blood biochemistry tend to vary considerably within and between species. In addition there may be variation between strains, ages and, sometimes, sexes and most importantly also between laboraties. The following table, derived from a larger table compiled by the University Animal Care Center of the University of Arizona Center, illustrates inter- and intraspecies variation. As explicitly stated in the original presentation of these data, they represent only indications gathered from a variety of sources; the range limits should not be interpreted as firm boundaries.

	Unit	mouse	rat	Hamster	rabbit	cat	dog	Rhesus
ALAT	IU/L	17-77	35-80	25-70	12-67	4-60	5-110	0-68
ASAT	IU/L	54-298	57-196[1)]	28-140	14-113	0-48	10-47	19-197
AP	IU/L	35-96	16-50	15-160	5-20	15-115	5-250	9-89
GGT[2)]	IU/L	n.a.[7)]	0.0	n.a.	n.a.	"negligible"	"negligible"	23-74
t-bilirubin	Mg/dL	0-0.9	0.2-0.5	0.2-0.8	0.2-0.7	0.0-0.6	0.0-0.7	0.1-0.2
PT	Sec	7-19[3)]	9-15[4)]	n.a.	n.a.	8.6[5)]	6-10[5)]	n.a. [6)]

1) added based on Lillie et al. (1996) and Wolford et al. (1986)
2) added based on Lillie et al. (1996) (rat), Woodman (1996) (cat, dog), Wolford et al. (1986) (monkey)
3) added based on O'Brien & Holmes (1993)
4) for CD rats based on Lang (1993)
5) Normal values as given at: http://www.thepetcenter.com/exa/nv.html
6) Normal values in humans between 11 and 13.65 seconds: http://www.nlm.nih.gov/medlineplus/ency/article/003652.htm)
7) n.a.= not available

Levine (2002) presents some detailed tables. Values for different age ranges for CD rats, F-344 rats, CD-1 and BLAB/c mice, Beagle dogs, New Zealand White rabbits and non-human primates. As explained by Levine the ranges of values presented are indications only.

1.7 Interpreting increases in ALAT, ASAT, AP, GGT, bilirubin and PT

For human medicine specific criteria for liver injury have been defined by international consensus, involving a combination of ASAT, ALAT, AP and total bilirubin. Increases greater than two times the upper limit of normal values are used as cut-off points for diagnosis of liver injury. Increases below this criterium of two times the upper limit are to be designated as a "liver test abnormality" only (inconclusive evidence for liver injury). As already indicated, analogous decision schemes are not available for laboratory animals.

In diagnosing liver disease in humans generally accepted reference intervals are increasingly used. In part such intervals are still in the process of development. For the interpretation of blood biochemistry results from animal toxicity studies no generally accepted reference intervals are available. The tables of normal values for several animal species that have been published in a number of instances, provide only rough, informal guidance and do not represent reliable reference intervals against which blood biochemistry data from individual studies can be judged. Given the biological and analytical variation between species, ages and laboratories, the development of such reference intervals is not easily achieved and would require a large-scale concerted international effort. Probably reference ranges for different ages and different strains would be required. In the absence of acceptable reference ranges a more pragmatic approach is inevitable, one in essence describing the current practice of commense sense professional medical judgement focussed on detecting patterns of consistent abnormalities in the different toxicity studies available for a compound under evaluation. As Evans (1996) crucially points out in his introduction to the book Animal Clinical Chemistry, small group changes in blood biochemical parameters, even within accepted reference ranges, may indicate early toxicity. For ASAT, ALAT in particular it was already noted above that despite their variability they represent remarkably sensitive parameters for tissue damage.

For increases in blood biochemistry parameters in animal toxicity studies, an informal rule of thumb minimum increase of 50% relative to the concurrent control has been suggested as cut-off point for a (potentially) adverse change. For ASAT, ALAT and AP with their high variability this percentage seems appropriate. For GGT with ist low baseline value this percentage may be too high. The parameters bilirubin and PT have relatively low variability and for these a lower minimum percent of change is warranted. A cut-off point of 10% may be considered an appropriate choice here.

The Health Council of the Netherlands (GR 2003), in a report exploring the applicability of the Benchmark Dose-method, calculated observable effect sizes based on the relative standard deviations (RSDs) as derived from previously conducted rat experiments with 20 animals per sex per group. The following critical effect sizes were proposed in in this document GR (2003):

Parameter	CES (%)
PT	5
ALP	30
ASAT	20
ALAT	25

Bilirubin was not included in this evaluation. The above CES-values are based on statistical variances, i.e. they reflect what can be detected statistically. Their toxicological significance has not been established.

Apart from the size of the change in the relevant blood biochemical parameters and its statistical significance, presence or absence of a dose response in the observed increase is crucially important. Since the focus should be on detecting relevant patterns of findings, histopathological findings must always be evaluated carefully. In the liver for evidence, hepatocellular damage (confirming increased ALAT and ASAT as toxicologically relevant) and hepatobiliary changes (confirming increases in AP and GGT as toxicologically relevant) should be evaluated. Even in the absence of histopathological effects, however, increases in the blood biochemical parameters discussed in this factsheet that meet the criteria of statistical significance and/or the 50% rule of thumb, may be indicative of early toxic damage and accordingly be relevant. The presence or absence of dose dependency in the observed response and the presence or absence of similar or other liver effects in other toxicity studies should provide important additional information in the expert judgement that is needed to reach a conclusion.

1.8 Conclusion

Liver injury and its diagnosis (for which enzyme activities are a tool) represents a complex field of study. As will be evident from the foregoing discussion, enzyme activities provide only a limited view on what may be a complex disease process and as a consequence residual uncertainty will frequently remain as to their interpretation, especially so when there are no concomitant histopathological effects.

At the present stage no reliable reference ranges for ALAT, ASAT, AP, GGT, bilirubin and PT in different animal species are available. The following pragmatic criteria for expert judgement on the relevance of observed increases in these parameters are available:

- statistical significance against concurrent control combined with a rule-of-thumb of 50% increase for ASAT, ALAT and AP and 10% for bilirubin, PT and bilirubin as the minimum for biological meaningfulness;
- presence of histopathological effects in the same study;
- consistency of effects across data base for compound under scrutiny;
- general toxicological plausibility of liver effects for compound under scrutiny (comparison with known toxicity of chemical analogues).

Note that expert judgement in the application of these factors is indispensable. The focus is to be on detecting *patterns* of consistent changes.

In annex I a decision scheme shows how the above criteria can be applied.

References

Amacher DE (1998). Serum transaminase elevations as indicators of hepatic injury following the administration of drugs. Regulatory Toxicology and Pharmacology 27, 119-130.

Amacher DE (2002). A toxicologist's guide to biomarkers of hepatic response. Human and Experimental Toxicology 21, 253-262.

Caisey JD and King DJ (1980). Clinical chemical values for some common laboratory animals. Clinical Chemistry 26 (13), 1877-1879.

Carakostas MC (1992). Interpreting clinical laboratory data in safety assessment studies. Toxicologic Pathology 20 (3), 480-483.

Carakostas MC and Banerjee AK (1990). Interpreting rodent clinical laboratory data in safety assessment studies: biological and analytical components of variation. Fundamental and Applied Toxicology 15, 744-753.

Davies DT (1996). Study design and regulatory requirements. In: EO Evans (ed.) Animal Clinical Chemistry – A primer for toxicologists. London, Taylor & Francis 1996.

Danan G. (2002). Presentation of the results of the International Consensus Meeting on Drug-Induced Liver Disorders, DIA Paris 10/2002.

Davies, DT (1992) Enzymology in preclinical safety evaluation. Toxicologic Pathology 20 (3), 501-505.

Dekker, S, de Heer, C, Rennen MAJ (2001). Critical effect sizes in toxicological risk assessment: a comprehensive and critical evaluation. Environmental Toxicology and Pahrmacology 10, 33-52.

Dufour DR, Lott JA, Nolte FS, Gretch DR, Koff RS and Seeff LB (2000a). Diagnosis and monitoring of hepatic injury. I. Performance characteristics of laboratory tests. Clinical Chemistry 46 (no.12), 2027-2049.

Dufour DR, Lott JA, Nolte FS, Gretch DR, Koff RS and Seeff LB (2000b). Diagnosis and monitoring of hepatic injury. II. Recommendations for use of laboratory tests in screening, diagnosis and monitoring. Clinical Chemistry 46 (12), 2050-2068.

Evans EO (1996). General introduction. In: EO Evans (ed.) Animal Clinical Chemistry – A primer for toxicologists. London, Taylor & Francis 1996.

Evans EO (1996). General enzymology. In: EO Evans (ed.) Animal Clinical Chemistry – A primer for toxicologists. London, Taylor & Francis 1996.

Fraser CG (2004). Inherent biological variation and reference values. Clinical Chemistry Laboratory Medicine 42 (7), 758-764.

Hall RL (1992). Clinical pathology for preclinical safety assessment: Current global guidelines. Toxicologic Pathology 20 (3), 472-476.

Health Canada (2004). Recommendations from the Scientific Advisory Panel Sub-groups on Hepatotoxicity: Hepatotoxicity of health products. Draft dated May 6, 2004.

Henny J and Petersen PH (2004). Reference values; from philosophy to a tool for laboratory medicine. Clinical Chemistry Laboratory Medicine 42 (7), 686-691.

Levine, BS (2002). Animal Clinical Pathology. In: MJ Derelanko, MA Hollinger (eds.) Handbook of Toxicology – Second edition. Boca Raton etc. CRC Press, 2002.

Lillie LE, Temple NJ and Florence LZ (1996). Reference values for young normal Sprague Dawley rats: weight gain, hematology and clinical chemistry. Human and Experimental Toxicology 15, 612-616.

Lindena J, Büttner D and Trautschold I (1984). Biological, analytical and experimental compontents of variance in a long-term stduy of plasma constituents in rats. Journal of Clinical Chemistry and Clinical biochemistry 22 (1), 97-104.

Matsuzawa T, Nomura M and Unno T (1993). Clinical pathology reference ranges of laboratory animals. Journal of Veterinary Medical Science 55 (3), 351-362.

NACB (2000). Laboratory Guidelines for Screening, Diagnosis and Monitoring of Hepatic Injury (Published Guidelines). At: http://www.nacb.org/lmpg/Hepatic_LMPG_PDF.stm

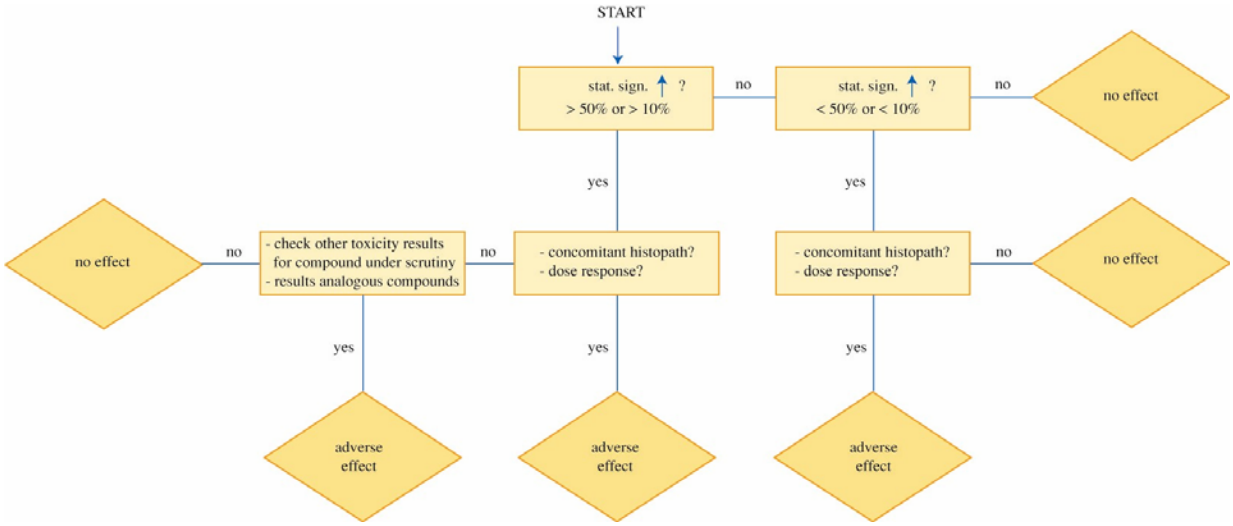
Robinson J and Evans EO (1996). Preanalytical and analytical variables. In: EO Evans (ed.) Animal Clinical Chemistry – A primer for toxicologists. London, Taylor & Francis 1996.

Wolford ST, Schroer RA, Gohs, FX, Gallo PP, Brodeck M, Falk HB, Ruhren R (1986). Reference range data base for serum chemistry and hematology values in laboratory animals. Journal of Toxicology and Environmental Health 18 161-188.

Woodman DD (1996). Assessment of hepatotoxicity. In: EO Evans (ed.) Animal Clinical Chemistry – A primer for toxicologists. London, Taylor & Francis 1996.

Internet:
http://www.charite.de/ch/neuro/expneuro/resources/pdf/chemical.pdf
http://www.ahsc.arizona.edu/uac/invest/business/chemistry.pdf

Annex Ia: Decision scheme.



2. Strategy for quantitative risk assessment for skin sensitisation using the Local Lymph Node Assays (LLNA)

Factsheet FSV/019, date 02/02/2006

Author: W. ter Burg

Contents:

2.1 Introduction

Dermal exposure to chemicals can result in skin sensitisation leading to a hypersensitive state (also referred to as delayed type IV hypersensitivity). When a skin sensitized subject is re-exposed to the same chemical this may lead to allergic contact dermatitis (ACD). In 2000, the number of new contact eczema patients was estimated to be 420,000 in the Netherlands of which approximately 30,000 cases can be attributed to skin sensitisation due to chemical exposure (Coenraads et al. 2003). For this reason, it is important to identify the hazards and risks from chemicals to induce skin sensitisation (Kimber et al., 2002).
To this end, as early as 1935 guinea pig assays are used to assess the ACD potential of chemicals. The Guinea Pig Maximisation Test (GPMT) has been considered to be the preferred method for predicting skin sensitisation. Next to the GPMT the Buehler Assay (BA) is the other guinea pig assay. The general concept on skin sensitisation encompassed the assumption that for induction of skin sensitisation no threshold exists. A substance was considered either a skin sensitizer or a non-sensitizer. Current understandings on the mechanism of action of skin sensitisation have led to a change in perspective. Although not yet proven, there is reason to believe that a threshold exists for skin sensitisation as in other immunisation processes.
GPMT and BA predict the ability of a substance to sensitize the skin considerably well. The interpretation of results are rather qualitative than quantitative and disregards any threshold (Dean et al., 2001). Conversely, the murine local lymph node assay (LLNA) does provide quantitative data which are used to obtain a dose-response relationship. The decision criteria whether a chemical is considered a skin sensitizer in LLNA is based on a threshold level (a three-fold increase in lymph node cell proliferation). At present, the LLNA is solely used to assess the ability of a substance to sensitize the skin (EC, 2003). In addition to its use in classification and labelling, the quantitative data obtained in LLNA could also be used to perform a quantitative risk assessment (QRA). However, up until now QRA is not performed for skin sensitising substances. For this reason research has been conducted to propose quantitative risk assessment methodology based on data from LLNA.
In this fact sheet it will be discussed whether or not the LLNA is suitable as foundation for QRA. General mechanism of sensitisation and principles of the assays will be discussed. Furthermore, a strategy is set up for QRA for skin sensitisation.

2.2 Background information

2.2.1 General mechanism of skin sensitisation

Allergic contact dermatitis is an immunological mediated cutaneous reaction to a certain chemical. The process of skin sensitisation and resulting ACD proceed in two phases: the induction phase and the elicitation phase. During the induction phase a subject is exposed to an inducing contact allergen sufficient to provoke an immune response leading to hypersensitivity (EC, 1996, Kimber et al., 2002). Due to this hypersensitive state an accelerated and more extensive secondary immune response will be elicited when the subject is subsequently exposed to the same or cross reactive chemical contact allergens (Divkovic et al., 2005). At the site of contact a cutaneous inflammatory reaction will take place, which is recognized clinically as ACD (Kimber et al., 2002).

The cellular skin sensitisation process is not yet fully elucidated; on the other hand, current understanding on the skin sensitisation process provides a proper insight on the general

mechanism. In order for a chemical to act as a skin sensitizer the chemical must first be able to cross the skin barrier. A chemical should have certain characteristics to pass the stratum corneum and gain access to the viable epidermis (Kimber et al., 2001). Then, chemicals must be able to stimulate certain T-cells by acting as a hapten. As stated by Landsteiner and Jacobs (cited in Divkovic et al., 2005): 'small organic molecules can only become a sensitizing entity when bound to a skin protein. Sensitizing chemicals are too small to be recognized by the immune system and need to be protein bound in order to elicit an immune response'. From this statement it was hypothesized that chemical allergens would bind to proteins to form hapten-conjugates (Divkovic et al., 2005).
Up until now it is known that skin sensitisation is dependent on the initiation of specific T-lymphocyte responses, which are provoked in lymph nodes. These local lymph nodes drain the site of exposure where a chemical sensitized the skin. In these local lymph nodes the antigen presenting cells (APC) accumulate. Briefly, Langerhans cells (LC) present in the epidermis can recognize and internalize hapten-protein conjugates. Regulated by local cytokines LC migrates as a dendritic cell (DC) towards the local lymph nodes, where they present the antigen on their surface with major histocompatibility complex (MHC) molecules of class II. From this point onwards the LC or DC is an antigen presenting cell (APC) which initiates activation of naïve T-cells (Grabbe and Schwarz, 1998; Kimber et al., 2002, 2002; Divkovic et al., 2005; Ryan et al., 2005) (see figure 2-1 for structural mechanism, adopted from US EPA 2004).
Clones of these naïve T-cells will circulate as memory cells through the bloodstream and/or reside in the skin. Upon re-exposure to the allergen, these memory cells infiltrate the site of contact and will release factors (such as cytokines) which attract inflammatory cells leading to a hypersensitive response (elicitation phase, see figure 2-1). The contacted skin will become erythematous and swollen which is recognized as ACD (Arts et al., 2005).

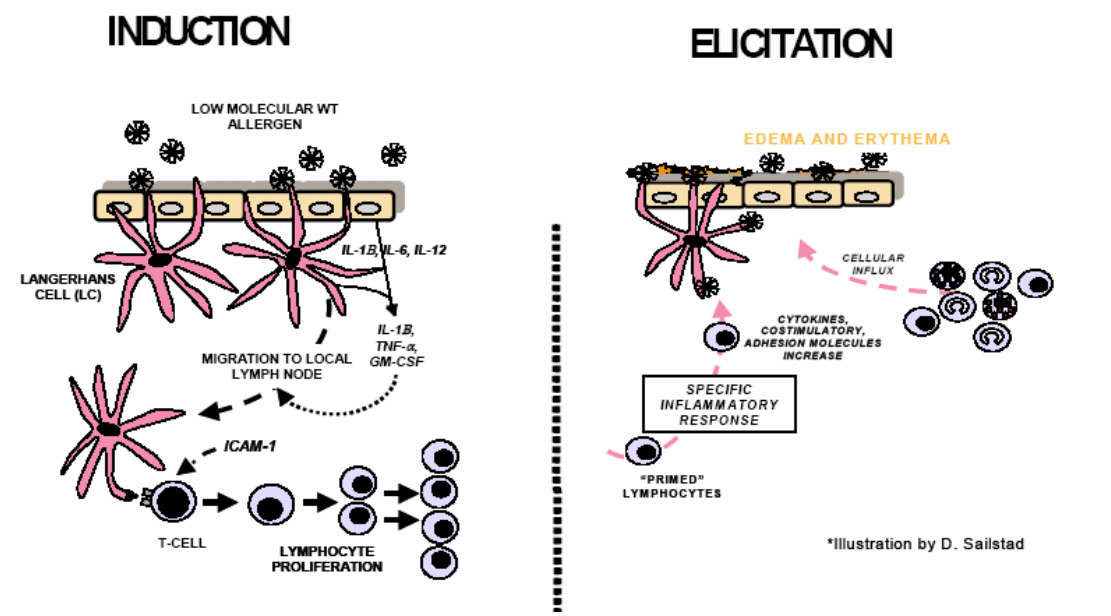


Figure 2-1: structural overview of skin sensitisation and elicitation.

The biological process of skin sensitisation (as described above) indicates that the response is related to the dose administered. The greater the level of exposure to contact allergen, the more vigorous will be the induced response and sensitisation acquired. Therefore, the potency is defined as a function of the amount of chemical required for the acquisition of skin sensitisation (Kimber et al., 2002). The table below displays a number of parameters which have influence on the potency of a certain chemical inducing skin sensitisation (Table 1

adopted from Griem et al., 2003). The potency is, however, partly determined by chemical-specific properties. Some of these chemical-specific properties are molecular weight (Arts et al., 2005), lipophilicity, ability to form hydrogen bounds, and ability to form haptens. These chemical characteristics could well influence the potency of a chemical to induce skin sensitisation. The other actors are the host-specific parameters and the availability of T-lymphocytes. However, the latter two will not play a role in differences in potency determined in guinea pig assays or LLNA, because experimental settings will be the same.

Table 2-1: Overview of parameters affecting potency of skin sensitization.

Parameter	Explanation
Skin penetration	Only after passing the skin barrier a chemical can interact with cells of the immune system and elicit a sensitisation; the penetration depends on the chemical itself (size, lipophilicity, and reactivity) and on circumstances (skin hydration, location of affected skin, presence of solvents that promote penetration).
Protein binding	The chemical can only be recognized by T-lymphocytes after binding covalently to soluble proteins or membrane proteins. Inefficiently binding leads to lower local concentrations.
Metabolism	Some sensitizing chemicals (prohapten) are not protein reactive as they come, but have to be metabolized. Efficiency of metabolism depends on enzyme expression and, eventually, genetic polymorphism.
Efficiency of uptake by LC	Only hapten bound to soluble proteins, membrane proteins or to damaged/dead skin cells can be taken up by LC and is thus available for presentation to T-cells.
Induction and maturation of LC	LC must be induced to leave the skin and to migrate to draining lymph node and to mature (e.g. upregulation of costimulatory membrane molecules) into DC; this activation can be caused by the sensitizing chemical itself (cytotoxic, irritative effect) or by circumstantial influences, such as physical injury, (chemical) irritation, or UV radiation.
Presentation of haptenated peptide-MHC complexes	In order to activate T-lymphocytes, one or more kinds of haptenated peptides must be cut out of the haptenated protein in a partial degradation by proteases and the haptenated peptide must fulfil the peptide binding motif allowing it to bind to a class II MHC molecule, so that it can be presented as a haptenated peptide-MHC complex at the DC surface.
Foreignness of haptenated peptide-MHC complexes	T-lymphocytes must be available that carry a T-cell receptor specific for the presented complex of peptide-MHC molecule; suitable T-lymphocytes may be lacking due to genetic polymorphism in the T-cell receptor gene segments or due to immunological tolerance (inactivation or deletion of T-lymphocytes with certain T-cell receptors.

2.2.2 Description of test methods

The GPMT and the BA both have been described in EC (B.6) and OECD (406) guidelines. In the document was stated that the GPMT is favoured over BA. The LLNA is described in EC B.42 which is equivalent to the OECD TG 429 (2002). In addition, tests in humans are also addressed in risk assessment and are shortly described below. The US EPA also describes these tests in their health effect test guidelines: OPPTS 870.2600 Skin sensitisation. These test are in detail described in above mentioned documents, therefore the tests will only be discussed briefly.

2.2.2.1 Guinea Pig Maximisation Test (OECD 406)

Healthy young male or female albino guinea pigs (females should be nulliparous and non-pregnant) are initially exposed to the test substance by intradermal injections and/or epidermal application (induction phase). The application of the test substance can be co-administered with an adjuvant (commonly: Freunds Complete Adjuvant, FCA). The procedure includes

application of three injections with FCA alone, the test substance alone, and FCA and test substance simultaneously. At day 7 (6-8) after the intradermal application, skin treatment is repeated with a filter paper fully loaded with the test substance in a proper vehicle. The filter paper is occluded and held in contact with the clipped skin for 48 hours. Control groups are treated similarly, but are only exposed to the adjuvant and vehicle.
After a rest period of 10 to 14 days the challenge takes place (elicitation phase). A patch or chamber is loaded with either test substance or vehicle and placed on the flanks of the animals for 24 hours. Approximately 24 and 48 hours after removal of the patch the treated areas will be observed for effects. All skin reactions resulting from induction or elicitation should be observed and recorded according to the grading scale of Magnusson/Kligman. A substance is considered positive when 30% of the induced animals react to the challenge with substance.

2.2.2.2 Buehler Assay (OECD 406)

Healthy young male or female albino guinea pigs are exposed to test substances and/or vehicles. Exposure is performed with a test patch system topically applied to the test area on the skin (induction). After one and two weeks the same application is carried out as on day zero. After approximately four weeks the challenge takes place by an occlusive patch (elicitation phase) with either the test substance or vehicle to the posterior untreated flank of the animal. The treated and control animals will be observed for effects. All skin reactions resulting from induction or elicitation should be observed and recorded according to the grading scale of Magnusson/Kligman. A substance is considered positive when 15% of the induced animals react to the challenge with substance.

2.2.2.3 Human tests

Next to the animal studies, there are also human (volunteer) studies, such as the human repeat-insult patch test (HRIPT) and the Human Maximisation Test (HMT). In both HRIPT and HMT humans are exposed during induction and challenge to determine whether a substance can sensitize the skin. In patch tests subjects are exposed once to known allergens in order to test whether the subject is sensitized to a certain substance. In fact, elicitation of skin sensitisation is screened in human patch tests. Furthermore the exposures are relatively high to provoke a reaction in the human subject. Data obtained from human patch tests therefore cannot be directly used in QRA. On the other hand, results from human volunteer studies should be considered in risk assessment processes as background information (Kimber et al., 2001).

2.2.2.4 Local Lymph Node Assay (OECD 429)

For the LLNA, young adult female mice of strain CBA/Ca or CBA/J are used. The females should be nulliparous and non-pregnant. Other strains may be used when sufficient evidence and data are generated to demonstrate that specific strain differences do not exist. The test procedure is as follows: on the first day 25 µl of test substance, vehicle or positive control is applied to the dorsum of each ear. Doses are selected from the concentrations series 100%, 50%, 25%, 10%, 5%, 2.5%, 1%, 0.5% etc. Preferred positive substances are hexyl cinnamic aldehyde and mercaptobenzothiazole. On day two and three the application procedure carried out on day one is repeated (induction phase). On the sixth day the animals are injected with 250 µl phosphate-buffered saline containing 20 µCi of ^{3}H-methyl thymidine or 2 µCi of ^{125}I-Iododeoxyuridine with 10^{-5} M-fluorodeoxyuridine into all mice via the tail vein. Five hours later the mice are sacrificed and auricular draining lymph nodes are excised. Single cell suspension of lymph node cells (LNC) is then prepared. With help of β-scintillation counting incorporation of ^{3}H-methyl thymidine or ^{125}I-Iododeoxyuridine is measured as disintegrations per minute (DPM). A mean DPM per group can be calculated (pooled data) or a DPM per animal can be calculated (non-pooled). The results of LLNA are presented as SI (Stimulation

Index). The SI is determined by the ratio of the proliferation in the treated groups to that in the vehicle control groups, where the average SI for vehicle group is set at one.
An indication of the quality of the LLNA should be presented alongside of the test results. This may be presented in the form of the test results from a positive control which should be chosen such that the induction is clear but not excessive and not exceed a half year before the test was performed. Furthermore the positive control should elicit a positive response with a stimulation index (SI) > 3 over the negative control group (vehicle). Supported by historical data the arbitrary choice was made to use an SI of 3 as indicator whether a test substance is a positive skin sensitizer or not (Dean et al., 2001).

2.2.3 Factors of influence on LLNA

2.2.3.1 Type of vehicle

The type of vehicle may influence the outcome in both guinea pig assays and LLNA. First of all the vehicle has to dilute the chemical properly and is not allowed to react with a chemical. Secondly, the vehicle should sufficiently adhere to the skin or should be properly occluded for an acceptable period of time during treatment. Hydrophillic chemicals should therefore be diluted into a vehicle system that wets the skin, but does not run off easily. Aqueous vehicles or runny liquids should be avoided (US EPA 2003). In order of preference, recommended vehicles are: acetone/olive oil (4:1), *N*,*N*-dimethylformamide, methyl ethyl ketone, propylene glycol, and dimethyl sulphoxide (Dean et al.,2001; US EPA, 2003). However olive oils may pose a problem in the LLNA. It was reported by Montelius et al. (1996) that SI levels of at least 16 were derived when tested at 100% and at least 2.9 when tested as acetone/olive oil (4:1). The 'control' group was in this case untreated animals. The author concluded that acetone/olive oil (4:1) should not be used as vehicle.

2.2.3.2 False positives and negatives

Although the LLNA was judged to be a suitable testing method for identifying potential hazardous skin sensitizers, it displayed during the evaluation a number of false positives. The chemicals which tested false positive were strong skin irritants. Strong irritants disrupt the structure of the epidermis. Irritants seem to have promoter effects on the sensitisation of chemicals (Felter et al 2002; 2003). Skin irritants induce an inflammatory reaction. Inflammatory reactions can lead to non-specific cell activation of dendritic cells, the migration of these dendritic cells, and finally causing a non-specific mediated stimulation of proliferation of LNC (Vohr and Ahr 2005). For skin sensitisation, only the cell-specific LNC proliferation due to chemical exposure is of interest. The fact that sodium lauryl sulphate (SLS; non-sensitizing irritant) was not able to influence the ability of DNCB (a potent skin sensitizer) to penetrate the skin, but did stimulate LNC proliferation in regard to DNCB tested alone, proved that irritants may promote sensitisation. Irritant pre-treatment may be used to increase the sensitivity of the LLNA for detection of weak sensitizers (De Jong et al., 2002).
When measuring the LNC proliferation without taking into account the irritant characteristics of a chemical thus leads to 'false' positive reactions. In addition, some chemicals that tested 'false' positive in LLNA were non-sensitizers in humans. No distinction, however, is made between an antigen-specific immune response and non-specific inflammatory reaction in LLNA (Dean et al., 2001; Vohr and Ahr, 2005).
In the GPMT the same problem occurs and will not indefinitely provide correct results. The LLNA was validated for testing of irritants (Dean et al., 2001). As there are no proper alternatives at present both the LLNA and/or GPMT can be considered for testing skin irritant for their ability to induce skin sensitisation. The LLNA is preferred over GPMT, because LLNA provides quantitative data.

There are other clinical outcomes that can be considered, besides the proliferation of LNC. These alternative endpoints include ear swellings, lesions, abrasive skin, caused by irritant. Another endpoint is the measurement of antigen responding cells by flow cytometry, thereby discriminating between contact sensitizers and non-specific proliferation (Dean et al., 2001; Kimber et al., 2002). However, these alternative endpoints are not validated yet and should be further investigated.

On the other hand the LLNA sometimes fails to identify skin sensitizers. In general, this is of greater concern for regulators. False negative results were found with some weak sensitizing agents and metals, e.g. nickel, undoubtedly a contact allergen (Dean et al., 2001; Kimber et al., 2002; Vohr and Ahr, 2005). Metals do not respond at all in LLNA and therefore ICCVAM stated that metals cannot be tested for skin sensitisation in the LLNA (Dean et al., 2001).
Guinea pig assays also displayed difficulties in identifying metals as potential skin sensitizers, for example nickel. One has to bear in mind that also the GPMT and BA show false positive and negative results.

2.2.4 Hazard identification

Both guinea pig assays and LLNA are used to identify skin sensitizers in the framework of EU classification and labeling. In guinea pig assays the animals are exposed and challenged to a substance, thereby actually measuring the ability to elicit an immune reaction in a sensitized animal. A substance is considered a positive skin sensitizer in the guinea pig assays when a minimum fraction of animals show skin lesions after treatment. Basically these tests will only provide a yes or no answer to whether a substance is able to sensitize the skin. Scaling the potency in these tests is limited when a single dose regime is followed, but possible when multiple doses are administered. The potency is then determined by the fraction of animals responding and the concentration of the substance administered. The potency categorization for GPMT and BA are displayed in an EU document (EC, 2003).
Data obtained from the LLNA are better suited to assign skin sensitizers into specific potency categories. Unlike the guinea pig assays the LLNA focuses on induction of sensitization only. Furthermore the LLNA provide objective and quantitative endpoints which can be used to obtain a dose-response relationship. As mentioned above a substance is a positive responder when a three-fold increase in LNC proliferation is observed compared to the control group. The concentration of the test substance, expressed as weight percentage in vehicle, at SI = 3 is the EC3 value. The lower the EC3 value, the more potent the substance as is displayed in an EU document (see also EC, 2003; Basketter et al., 2005).
As the guinea pig assays did not provide useful data for quantitative risk assessment (QRA); risk assessors were unable to conduct QRA for skin sensitising chemicals. The LLNA, however, does provide quantitative data which can be used in QRA. Up until now QRA is not performed for skin sensitizers. Performing a QRA for skin sensitising substances is desired, because it adds quantitative value to a characterisation of risk.

2.3 Approaches for quantitative risk assessment

2.3.1 Determining point of departure

As mentioned before, the GPMT and/or BA are not suited for QRA, because no quantitative data is provided. For this reason research was conducted to use the LLNA as foundation for QRA, because it can provide quantitative data. The endpoint measured in the LLNA is the LNC proliferation which directly links to the dose administered. Because multiple dose groups are

used, the LLNA test provides a dose-response relationship. From this obtained dose-response relationship an EC3 value can be calculated for the chemical. The EC3 value is the concentration of the chemical that is required to induce a threefold increase in LNC proliferation (response of SI = 3) compared to the vehicle/control. The three-fold increase is regarded the decision criteria whether a chemical is a skin sensitizer or not. Thus the derived EC3 value and belonging effect (LNC proliferation / skin sensitisation) is considered toxicologically relevant in risk assessment. The EC3 value can therefore be used to derive the point of departure in QRA for induction of skin sensitisation.
The EC3 value derived from LLNA is expressed in weight percentages. In quantitative risk assessment it is preferred to express the EC3 value as an area dose, because it was revealed that the induction of skin sensitisation is dependent on the dose per skin area (Kimber et al., 2001; Farage et al., 2003; Griem et al 2003). The area dose, expressed as $\mu g/cm^2$, is calculated by multiplying the concentration in LLNA (in weight %) with a conversion factor of 250 (for details see Griem et al., 2003).

With help of linear interpolation (ruler based method) an EC3 value can be determined from obtained data (Basketter et al., 1999a; 1999b; 2000; Kimber et al., 2001; 2002). Considering only the data points immediately above and below the SI = 3, provides an estimate of the EC3 value which is heavily reliant on these two data points.
More sophisticated models can also be applied to calculate an EC3 value as was shown by Van Och et al. (2000) where curve fitting analyses were applied. A family of nested models was used to fit the data obtained from LLNA (for details see Van Och et al., 2000). The uncertainty in the estimate of EC3 is assessed by bootstrapping so that confidence intervals can be derived.

Using the curve fitting method is preferred over the ruler based method, because the entire dataset obtained from LLNA is taken into account. Furthermore, this method can provide estimations on sensitizing concentrations not addressed in the concentration range. This may be useful for very weak sensitizers where an SI = 3 was not yet reached. In addition, determining the 90% confidence interval indicates that there is only a 5% probability that positive reactions do occur below the lower confidence limit. To be on the safe side the lower confidence limit of the derived EC3 value will be used as the point of departure.

However, when pooled data are provided, determining a confidence interval is not possible. In this case it is suggested to apply an uncertainty factor. The degree of the extrapolation factor is determined by observing the data from a study from Van Och and colleagues (2000). In this study the authors applied the curve fitting method on 10 substances where EC3 values and confidence intervals were calculated. The log distributions of the confidence intervals showed that deviations from the average were not higher than half order of magnitude (half order of magnitude on a log-scale equals to a factor three on a normal scale). For this reason an uncertainty factor of 3 will be used to derive the point of departure. When curve fitting is not available the ruler-based method can be used. It is advised to apply the same extrapolation factor of 3 on the derived EC3 value.

2.3.2 Methods for extrapolation

QRA can be performed following two methods. One method is to use extrapolation factors to derive a safe dose for humans directly from the point of departure. Another approach is to scale the potency of sensitizing chemicals. In the latter method the scaling of potency is used to derive default NELs (No effect levels) for humans.

2.3.2.1 Use of assessment factors

From the point of departure, a safe dose for humans can be determined for induction of skin sensitisation since there is reason to believe that a threshold exists for skin sensitisation. Griem and colleagues (2003) made an effort in deriving an 'acceptable non-sensitizing area dose' (ANSAD) by using an uncertainty/extrapolation factor approach.
Three extrapolation factors were suggested by Griem; these were interspecies differences, intraspecies differences, and a correction factor for exposure duration exposure.

- Interspecies differences

For interspecies differences however it was argued to use a factor of 3 instead of the classical factor of 10. With regard to skin penetration, rodents tend to display a considerably higher skin penetration for most chemicals compared to humans. The use of murine data is therefore considered conservative since three- to tenfold higher penetration have been reported (Griem et al., 2003). Furthermore, the biological process that takes place for the immune system to respond to sensitizers is a local effect where similar processes are considered across mammalian species.

- Intraspecies differences

No arguments are provided to deviate from the default factor of ten usually applied in toxicology; therefore a factor of 10 is used.

- Correction factor for exposure duration

De Jong and colleagues (in prep.) studied the effect of repeated exposure to formaldehyde donors on local lymph node responses. The animals were exposed to levels below EC3 value. The authors observed that the animals were still sensitized after repeated exposure even though exposure was below EC3. This supports the need for an extrapolation factor for repeated exposure. For this reason, a default extrapolation of ten used in toxicology is also used to correct for exposure duration. Hence a factor of 10 is used.

Based on the arguments above, Griem proposed to apply an extrapolation factor of 300 (3x10x10) on the EC3 value to derive an ANSAD. Note that this derived safe dose for humans is based on induction of skin sensitisation and not on elicitation! Furthermore, the intraspecies extrapolation factor does not account for previous sensitized individuals.
A different approach which can be used in QRA is the use of the Margin of Safety (MOS) or Margin of Exposure (MOE) procedure. The desired MOS is based on the extrapolation factors. In case of the example set by Griem et al. (2003) the desired MOS would be 300. The point of departure from LLNA divided by the potential exposure from the sensitizing chemical should not be lower than 300.

Lower concentrations are usually needed to initiate a skin reaction in previous sensitized individuals. To this end Griem and colleagues (2003) also developed a procedure to derive an 'acceptable non-elicitation area dose' (ANEAD). In theory it is expected that there exists a relation between induction and elicitation, because several factors influencing sensitizing potency are also relevant in elicitation (see also table 1). A relationship was found by the authors between the ratio of sensitisation threshold/elicitation threshold and induction threshold (see formula). Hereby data from human patch tests (elicitation thresholds) and from LLNA is combined (induction thresholds).
Using the point of departure the author proposes a variable sensitisation-elicitation extrapolation factor of 10^(0.84 * log(NOEL induction threshold [µmol/cm^2]) + 1.81) and the extrapolation factors described above (300) to derive the ANEAD.

Log (ratio sensitisation/elicitation threshold) = 0.84 * log (induction threshold [μmol/cm^2]) + 1.81

The authors further note that data from human studies can also be used to derive an ANEAD; by applying a factor of 100 to the elicitation NOEL (10 for intraspecies differences and 10 for repeated exposure).

Another possible method to derive an elicitation threshold is to use the MET (Minimum Elicitation Threshold) method (US EPA, 2004). It uses data from the results of tests in previous sensitized individuals. However in regard to risk assessment this method has not been discussed to what criteria should be used. An example of MET method is provided by US EPA where Nethercott (1994) calculated the 10% MET (concentration at which 10% of study group responded) based on results of a study in previously sensitized individuals. No discussion of the relevance of the 10% MET was provided.

2.3.2.2 Scaling potencies

The second method for applying QRA for skin sensitisation is the use of potency scaling. An important step in risk assessment is the evaluation and categorization of potencies of chemicals able to induce human health effects, in this case skin sensitisation.
The derived EC3 values can be compared to one another. The higher an EC3 value for a certain chemical, the lower the potency of that chemical to sensitize the skin. Kimber and colleagues (2002) suggest that the relative potency can be categorized into five categories as follows: 1) EC3 values below 0.1% (volume percentage of substance in vehicle) as the very strong sensitizers, 2) EC3 values between 0.1-1% as strong sensitizers, 3) EC3 values between 1-10% as moderate sensitizers, 4) EC3 values between 10-100% as weak sensitizers, and 5) EC3 values of greater than 100% which are non-sensitizers in the LLNA.
In the EU document for classification and labelling, however, a more simplified categorisation for potency is used. EC3 values below or equal to 0.2% are considered extreme sensitizers, EC3 values between 0.2-2% are strong sensitizers and above 2% chemicals are considered to be moderate skin sensitizers (EC, 2003). The content of this EU document was published in an article by Basketter et al. (2005).

To use the potencies in QRA, Gerberick et al. (2001) suggested yet another categorization of potency for risk assessment. Gerberick used the correlation between EC3 values (in area dose) and the clinically observed NOAEL in human tests to set default NOELs for humans. A weight of evidence approach using all available potency data from both animal models as human experience are then used to determine the appropriate potency category (Gerberick et al., 2001; Felter et al., 2003). The author emphasises that the classification scheme has been adopted for the facilitating of identifying a default NOEL. Consequently the classification does not necessarily represent the best approach of classification of skin sensitizing potency. Further the author provides no information on the magnitude of the database which is used as foundation for the classification. Griem and colleagues (2003), however, observed a correlation close to 1 between human NOELs and the EC3 value from LLNA for approximately 30 chemicals.
Briefly, six categories ranging from non-sensitizing, for which no NOEL is derived, to potent skin sensitizers with a default NOEL of 1 μg/cm^2 were determined. When an EC3 value from LLNA falls within a range the lower limit of that range will be considered as a default in QRA. In table 2-2 below the categories are listed according to Gerberick et al. (2001).

Table 2-2: classification of NOELs for sensitisation potency according to Gerberick et al. (2001).

Sensitization potency classification: default NOELs for use in quantitative risk assessment (QRA)			
LLNA EC3	Sensitization potential	Experimental human NOEL	Default NOEL for use in QRA
NC[a]	non-sensitizing	NC[a]	NA[b]
>10,000 μg/cm^2	extremely weak	>10,000 μg/cm^2	10,000 μg/cm^2
1000–10,000 μg/cm^2	weak	1000–10,000 μg/cm^2	1000 μg/cm^2
100–1000 μg/cm^2	moderate	100–1000 μg/cm^2	100 μg/cm^2
10–100 μg/cm^2	strong	10–100 μg/cm^2	10 μg/cm^2
≤10 μg/cm^2	potent	≤10 μg/cm^2	1 μg/cm^2

From the derived NOEL for humans, obtained from extrapolations or potency based risk assessment, a sensitisation reference dose (S-RfD) can be calculated. In QRA sensitisation uncertainty factors (SUFs) are applied to calculate S-RfD. The areas generally considered are: inter-individual response variability, vehicle matrix differences, and variation in product use patterns not described in exposure assessment (e.g. whether or not occluded). As a default a factor of 10 is suggested for each of the areas (Gerberick et al., 2001; Felter et al., 2002; 2003; Farage et al., 2003; US EPA, 2004). The reader is referred to Felter et al. (2002) for a more detailed description of SUFs. However, the SUFs as proposed by Felter (2002) should be addressed under exposure assessment and should not be considered to derive a NOEL for humans.

These methods of using default NOELs are not preferred over the method published by Griem et al (2003) (see 1.3.2.1), because it is very rough in the sense that it disregards information available. Ideally information from both LLNA and human assays should be advised to provide a strong weight of evidence for classification. In this ideal situation where the EC3 value and human data are available the potency scaling method can be used.

2.4 Sensitive groups

In principle all individuals can be sensitized by chemicals leading to ACD. Individuals with skin disease or subjects who suffered from previous inflammatory reactions in the skin are more prone to become sensitized when exposed. The skin barrier is in those cases damaged leading to a higher skin penetration of a chemical. Furthermore individuals who already have (multiple) skin allergies could constitute a subpopulation with a higher susceptibility against other chemicals (Felter et al., 2002; Griem et al., 2003). Another sensitive group is individuals who are immunohypersensitive. Skin sensitization may then occur in an earlier stage and more intense (Fourtanier et al., 2005). There are no reasons to deviate from the general factor of 10 to protect the sensitive human.

2.5 RIVM risk assessment strategy

Based on the information gathered in this document we conclude that the data obtained from LLNA can be used for a quantitative risk assessment. The LLNA provides a dose-response relationship, which enables a quantitative risk assessment for skin sensitizers (Griem et al., 2003).

However, it is acknowledged that LLNA does not identify metals as sensitizers of the skin. In the validation study by ICCVAM it was concluded that metals should not be tested in LLNA (Dean et al., 2001). Hence, the LLNA cannot be used for metals.
Furthermore, the QRA for skin sensitisation is based only on induction and not on elicitation. Current understanding does not yet allow a quantitative risk assessment for elicitation. Although research has been conducted to obtain a safe dose for previous sensitized individuals (Griem et al., 2003; US EPA, 2004), methods are still not validated or accepted.
In the case that a point of departure could not be derived from LLNA a QRA for skin sensitisation cannot be performed.

2.5.1 Point of departure

Although an SI of 3 as the decision criterium is set by arbitrary choice, it has stood time and proved to be accurate in assigning positive skin sensitizers (Dean et al., 2001). For this reason the LLNA test can be used for QRA and the EC3 value from LLNA is considered toxicologically relevant.
There are three methods to derive a point of departure which depends on the study design or the kind of information provided. Data obtained from the LLNA can either be non-pooled or pooled. A third situation may occur when only an EC3 value was found in literature, but in which no data from the test was included.

1. From non-pooled data an EC3 value with confidence intervals can be derived by using a curve fitting method (dose-response modelling). The lower confidence limit of the EC3 value should be used as point of departure in QRA.
2. When pooled data are provided, a determination of the confidence interval is not possible. A point estimate of the EC3 value can be made by using the curve fitting method. In those cases where a lower confidence limit cannot be determined, an uncertainty factor of 3 will be used on the obtained EC3 value to derive the point of departure (EC3 value divided by 3).
3. In the case only an EC3 value is available the point of departure should be derived by dividing the EC3 value with a factor 3.

The point of departure is preferably determined by non-linear regression methods, as suggested by Van Och et al. (2000), where dose-response relationships were fitted to the data. Furthermore, the curve fitting method is preferred over the ruler based method because all data from the LLNA is regarded.

2.5.2 Conversion factor from weight percentage to area dose

Since it was revealed that the induction of skin sensitization is dependent on the dose per skin area (Kimber et al., 2001; Farage et al., 2003; Griem et al., 2003) concentrations in percentiles will be converted to moles or grams per cm^2. To derive the area dose in $\mu g/cm^2$ the administered concentration must be multiplied by a factor 250 (see Griem et al., 2003 for details).

2.5.3 Assessment factors

To derive a safe dose for skin sensitisation from the point of departure it was decided to use assessment factors. This extrapolation method, as proposed by Griem et al. (2003), includes the use of three assessment factors for interspecies and intraspecies differences and for the correction for exposure duration. The assessment factors are set at 3 for interspecies, 10 for

intraspecies, and 10 for correction for exposure duration, which equals to a total assessment factor of 300. These assessment factors should be used unless proper argumentation is provided to deviate from these default extrapolation factors (see 1.3.2.1). To derive the safe dose for skin sensitisation the point of departure should be divided by a factor of 300.

A schematic overview of the risk assessment strategy is provided in figure 2-2. When using the MOS/MOE procedure a value of minimum 300 is desired. Note that the intraspecies extrapolation factor used to account for sensitive human does not include previous sensitised individuals.

The use of assessment factors is preferred over the potency based method. It must be mentioned that the potency based QRA is international widely promoted and much effort is invested to develop a large database which then can be used in potency based QRA.

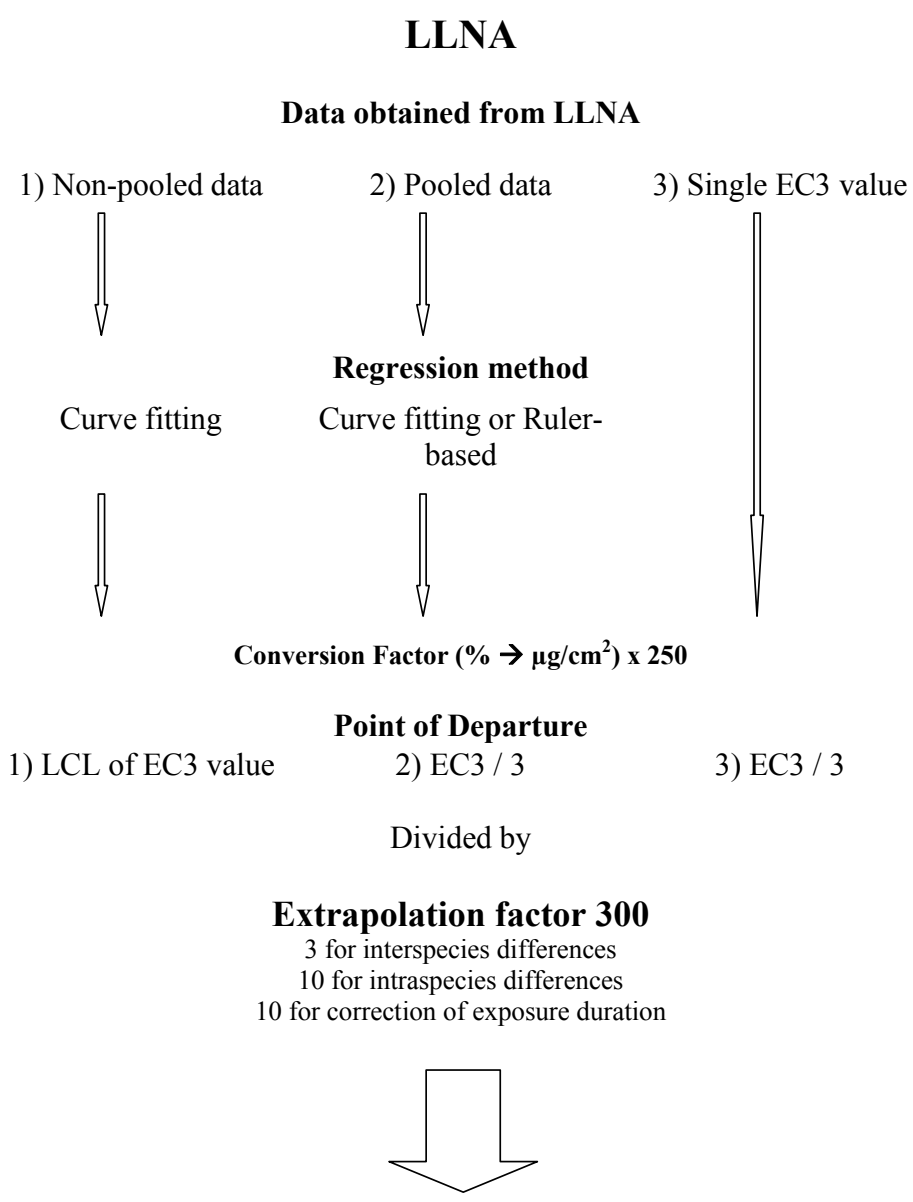


Figure 2-2: A schematic overview of the risk assessment strategy for skin sensitisation concerning data obtained from the LLNA.

Concluding remark: this factsheet is not prepared in order to propose a standard risk assessment for the endpoint of skin sensitisation in routine regulatory frameworks. However, in those case where a QRA for skin sensitisation is needed or desired this fact sheet describes the method most appropriate.

2.5.4 Examples

In this paragraph two examples will be provided on how to derive a safe dose for humans from data obtained in the LLNA.

2.5.4.1 Example 1

Data from Van Och et al. (2000) concerning the compound Phthalic anhydride:
Step 1: Non- pooled data was obtained from LLNA. The authors used the curve fitting method to derive the EC3 value and the confidence interval. The resulting EC3 value and confidence interval was: 0.357 (0.226-0.560) %.
Step 2: The point of departure is the lower confidence limit of the EC3 value: 0.226%.
Step 3: Conversion from weight % to area dose (x 250): 56.5 $\mu g/cm^2$ = point of departure.
Step 4: Divide the point of departure by the assessment factor of 300 (56.5 / 300) = 0.188 $\mu g/cm^2$.
The safe dose for Phthalic anhydride for humans is thus: 0.188 $\mu g/cm^2$

2.5.4.2 Example 2

The second example concerns data from Betts et al. (2005) where Disperse blue 106 was tested in LLNA (data from experiment 2 is used):
Step 1: Pooled data was obtained from the LLNA. The authors used the ruler-based method to derive the EC3 value which was 0.012%.
Step 2: Point of departure is EC3 value divided by uncertainty factor 3: 0.012% / 3 = 0.004 %.
Because no confidence interval could be derived a factor of 3 is used to account for uncertainty from the data.
Step 3: Conversion from weight % to area dose (x 250): 1 $\mu g/cm^2$ = point of departure.
Step 4: Divide the point of departure by the assessment factor of 300 (1 / 300) = 0.0033 $\mu g/cm^2$
The safe dose for Dispersion blue 106 for humans is thus: 0.0033 $\mu g/cm^2$.

References

Arts JHE, Mommers C, and De Heer C. 2006. Dose-response relationships and threshold levels in skin- and respiratory allergy. Critical reviews in Toxicology 36(3): 219-51.

Basketter DA, Lea LJ, Dickens A, Briggs D, Pate I, Dearman RJ, and Kimber I. 1999a. A comparison of statistical approaches to the derivation of EC3 values from Local Lymph Node Assay dose responses. J. Appl. Toxicol. 19: 261-266.

Basketter DA, Lea LJ, Cooper K, Stocks J, Dickens A, Pate I, Dearman RJ, and Kimber I. 1999b. Threshold for classification as a skin sensitizer in the Local Lymph Node Assay: a statistical evaluation. Food Chem. Toxicol. 37: 1167-1174.

Basketter DA, Blaikie L, Dearman RJ, Kimber I, Ryan CA, Gerberick GF, Harvey P, Evans P, White IR, and Rycroft RJG. 2000. Use of the local lymph node assay for the estimation of relative contact allergenic potency. Contact Derm. 42: 344-348.

Basketter DA, Andersen KE, Liden C, Van Loveren H, Boman A, Kimber I, Alanko K, and Berggren E. 2005. Evaluation of the skin sensitizing potency of chemicals by using the

existing methods and considerations of relevance for elicitation. Contact Derm. 52: 39-43.

Betts CJ, Dearman RJ, Kimber I, and Maibach HI. 2005. Potency and risk assessment of a skin-sensitizing disperse dye using the local lymph node assay. Contact Derm. 52: 268-272.

Dean JH, Twerdok LE, Tice RR, Sailstad DM, Hattan DG, and Stokes WS. 2001. ICCVAM evaluation of the murine local lymph node assay. II Conclusions and recommendations of an independent scientific peer review panel. Regul. Toxicol. Pharmacol. 34: 258-273.

De Jong WH, Tentij M, Spiekstra SW, Vandebriel RJ, and Van Loveren H. 2002. Determination of the sensitising activity of the rubber contact sensitisers TMTD, ZDMC, MBT and DEA in a modified Local Lymph Node Assay and the effect of sodium dodecyl sulfate pretreatment on local lymph node responses. Toxicology. 176: 123-134.

De Jong WH, De Klerk A, Ter Beek M, Veenman C, and Van Loveren H. (In prep.) Effect of prolonged repeated exposure to formaldehyde donors with doses below the EC3 value on local lymph node responses.

Divkovic M, Pease CK, Gerberick GF, and Basketter DA. 2005. Hapten-protein binding: from theory to practical application in the *in vitro* prediction of skin sensitization. Contact Derm. 53: 189-200.

European Commission. 1996. B.6. Skin sensitization. In: Dir 96/54/EC.

European Commission. 2003. Meeting of the Sensitisation Expert Group, Ispra 4-6 November 2002. ECBI/81/02 Rev. 2; 8 January 2003.

European Commission. 2004. B.42. Skin sensitization: Local Lymph Node Assay. In: Dir 2004/73/EC.

Farage MA, Bjerke DL, Mahony C, Blackburn KL, and Gerberick GF. 2003. Quantitative risk assessment for the induction of allergic contact dermatitis: uncertainty factors for mucosal exposures. Contact Derm. 49: 140-147.

Felter SP, Robinson MK, Basketter DA, and Gerberick F. 2002. A review of the scientific basis for uncertainty factors for use in quantitative risk assessment for the induction of allergic contact dermatitis. Contact Derm. 47: 257-266.

Felter SP, Ryan CA, Basketter DA, Gilmour NJ, and Gerberick GF. 2003. Application of the risk assessment paradigm to the induction of allergic contact dermatitis. Regul. Toxicol. Pharmacol. 37: 1-10.

Fourtanier A, Moyal D, maccario J, Compan D, Wolf P, Quehenberger F, Cooper K, Baron E, Halliday G, Poon T, Seed P, Walker SL, and Young AR. 2005. Measurement of sunscreen immune protection factors in humans: a consensus paper. J. Invest. Dermatol. 125: 403-409.

Gerberick GF, Robinson MK, Felter SP, White IR, and Basketter DA. 2001. Understanding fragrance allergy using an exposure-based risk assessment approach. Contact Derm. 45: 333-340.

Grabbe S and Schwarz T. 1998. Immunoregulatory mechanisms involved in elicitation of allergic contact hypersensitivity. Immunol. Today. 19: 37-44.

Griem P, Goebel C, and Scheffler H. 2003. Proposal for a risk assessment methodology for skin sensitization based on sensitization potency data. Regul. Toxicol. Pharmacol. 38: 269-290.

Kimber I, Basketter DA, Berthold K, Butler M, Garrigue JL, Lea L, Newsome C, Roggeband R, Steiling W, Stropp G, Waterman S, and Wiemann C. 2001. Skin sensitization testing in potency and risk assessment. Toxicol. Sci. 59: 198-208.

Kimber I and Dearman RJ. 2002. Allergic contact dermatitis: the cellular effectors. Contact Derm. 46: 1-5.

Kimber I, Dearman RJ, Basketter DA, Ryan CA, and Gerberick GF. 2002. The local lymph node assay: past, present and future. Contact Derm. 47: 315-328 review.

Ryan CA, Gerberick GF, Gildea LA, Hulette BC, Betts CJ, Cumberbatch M, Dearman RJ, and Kimber I. 2005. Interactions of contact allergens with dendritic cells: Opportunities and challenges for the development of novel approaches to hazard assessment. Toxicol. Sci. 88 (1): 4-11.

US EPA. 2003. Health effects test guidelines: OPPTS 870.2600 Skin sensitization.

US EPA. 2004. FIFRA scientific advisory panel background document May 4-6, 2004 meeting – Proposed Hazard identification methodology for assessment of dermal sensitization risk. 1-27.

Van Och FMM, Slob W, De Jong WH, Vandebriel RJ, and Van Loveren H. 2000. A quantitative method for assessing the sensitizing potency of low molecular weight chemicals using a local lymph node assay: employment of a regression method that includes determination of the uncertainty margins. Toxicol. 146: 49-59.

Van Och FMM, Vandebriel RJ, Prinsen MK, De Jong WH, Slob W, and Van Loveren H. 2001. Comparison of dose-responses of contact allergens using the guinea pig maximization test and the local lymph node assay. Toxicol. 167: 207-215.

Vohr H-W and Ahr HJ. 2005. The local lymph node assay being too sensitive? Arch. Toxicol. 12: 721-728.

3. Leydig cell tumour

Factsheet FSV/012-01 update of FSV/012-00, date 16-02-2006

Author: S.M.G.J. Pelgrom

Contents

3.1 Introduction and aim

Spontaneous testicular tumours are commonly observed in some rat strains (Bär, 1992; Prentice and Meikle, 1995; Bosland, 1996) and in dogs (Peters and Sluijs, 1996) and mice (Chandra and Frith, 1992; Clegg et al., 1997). The tumours can be divided in three types: Sertoli cell tumour, seminoma and Leydig cell tumour (interstitial cell tumour) (Peters and Sluijs, 1996), (Cook et al., 1999). The main function of Leydig cells (LCs) is the production of testosterone.
Leydig cell tumours (LCTs) are mostly benign and are observed especially in older animals (Bär, 1992; Prentice and Meikle, 1995; Peters and Sluijs, 1996).
Some uncertainty exists about the true occurrence of Leydig cell adenomas in man, although occurrence seems to be rare and restricted primarily to white males (Clegg et al., 1997). The known risk factors for LC adenomas are restricted to heritable disorders of the endocrine system (congenital adrenal hyperplasia and androgen insensitivity syndrome), Klinefelter's syndrome, Androgen Insensitivity Syndrome (AIS, Quigly et al., 1995 in Document ECBI/61/03 Classification and labelling, Draft, 2003), familial male precocious puberty (FMPP, Shenker et al., 1993 1995 in Document ECBI/61/03 Classification and labelling, Draft, 2003) and cryptorchidism with related testicular atrophy (Schottenfeld et al., 1980; Prentice and Meikle, 1995; ref. 104 in Clegg et al., 1997).
LCs are a common target of compounds tested in rodent carcinogenicity bioassays (Cook et al., 1999). Doubts have been raised about the relevance of chemical-induced LCTs for human risk assessment (Bär, 1992; Alison et al., 1994; Prentice and Meikle, 1995; Clegg et al., 1997; Cook et al., 1999). In this factsheet, the toxicological relevance of a chemical-induced increase in LCTs in laboratory animals for human risk assessment will be discussed.

3.2 General information and mechanisms of induction

In the rat the distinction between LC hyperplasia and adenoma (tumour) is, like other endocrine tumours in the rat, arbitrarily based on size (Cook et al.(no date); Clegg et al., 1997; Cook et al., 1999). There has been much discussion concerning the exact size that should be used to discriminate between hyperplasia and adenoma. Two arbitrary criteria have been developed for rodents (Cook et al., 1999):
According to the National Toxicology Program (NTP), an aggregate of LCs smaller than the diameter of a seminiferous tubule is classified as being focal hyperplasia (Boorman et al., in Cook et al., 1999). A mass of LCs greater than that of a seminiferous tubule is classified as a tumour. Masses of this size generally produce some compression of adjacent tubules.
Guidelines recommended by The Society of Toxicologic Pathologists for standarization of diagnosis propose three tubular diameters as the arbitrary separation of focal LC hyperplasia from LC neoplasia (McConell et al., in Cook et al., 1999). In addition, the proliferative focus must also have morphological features consistent with LCT. These features may include evidence of autonomous growth by symmetrical peripheral compression of adjacent seminiferous tubules, evidence of cellular pleomorphism, and development of a typical endocrine sinusoidal vascular network.
To date, the debate continues without a generally accepted size criterium (Cook et al., 1999). In toxicology safety assessment studies, almost none of the references use the pathological criteria to distinguish between LC hyperplasia (LCH) and LCT (Cook et al., 1999). It is considered that the transition from LCH to LCT is part of a continuous spectrum of change, and both LCHs and

LCTs may be routinely pooled in the interpretation and analysis of carcinogenicity studies (Prentice and Meikle, 1995).
In contrast to rodents, no size criterion is used to distinguish between hyperplasia and tumour in humans (Mostofi and Davis, 1990 in Cook et al., 1999).

A number of chemicals, including many non-DNA-reactive compounds, have been shown to increase the incidence of LC hyperplasia and tumours in chronic studies in certain strains of rats, and occasionally in mice and dogs (Prentice and Meikle, 1995). Morphologically, there appears to be no difference between spontaneous and chemically induced LCTs (Cook et al. (no date); Prentice and Meikle, 1995). For many of the compounds that induce LCH and/or LCT the mechanism of action is not known. However, there are some generalisations that can be made regarding the mode of action or chemical activity, primarily involving effects on the hypothalamic-pituitary-testicular axis (HPT-axis).

The regulation of testosterone production by LCs can be divided into extratesticular (via pituitary) and intratesticular (paracrine: by seminiferous tubules, and autocrine: by LC products such as estradiol and testosterone) (Ewing, 1992). The primary sites in the regulation of testosterone production in rats and humans are the hypothalamus and the pituitary (see Figure 3-1) (Cook et al., 1999; De Groot et al., 1995b).
In order to understand how chemicals may induce LCTs by interrupting the HPT-axis, it is necessary to understand the regulation of the HPT-axis. The hypothalamus secretes gonadotropin-releasing hormone (GnRH), which stimulates the secretion of luteinizing hormone (LH) by the pituitary. LH maintains testosterone levels in homeostasis (ref 17 in Cook et al. (no date); Freeman, 1991; Prentice and Meikle, 1995; Cook et al., 1999) and LH binds to Leydig cells and activates adenylate cyclase to increase cAMP levels. Increased cAMP levels stimulate testosterone biosynthesis causing testosterone levels to rise in the bloodstream. Testosterone exerts a negative feedback on the release of GnRH and LH from the hypothalamus and pituitary, respectively (see Figure 1) (Cook et al. (no date); Cook et al., 1999).
An additional feedback control mechanism of LH involves the steroid estradiol. Testosterone is converted to estradiol via the enzyme aromatase, a process which is commonly referred to as aromatization. Aromatase activity is found in adipose tissue, liver, skeletal muscle and testis. In males, the majority of estradiol synthesis occurs in adipose tissue. Similarly to testosterone, increasing blood estradiol levels will attenuate LH secretion.

Figure 3-1 also illustrates mechanisms by which chemicals can disrupt the HPT-axis and produce LCTs in rats. In animals treated with some LC tumour-inducing agents, an increased level of circulating LH was observed, leading to the current concept that these adenomas arise due to chronic LH elevation (Clegg et al., 1997; Cook et al. (no date); Bär, 1992).
Increases in LH level were not seen in all studies of chemicals for which the proposed mode of action calls for elevated LH. Compensation may occur to restore homeostasis, such as induction secondary to seminiferous tubule damage, paracrine involvement, events related to peroxisome proliferation and increased steroid clearance due to enzyme induction (Clegg et al., 1997).

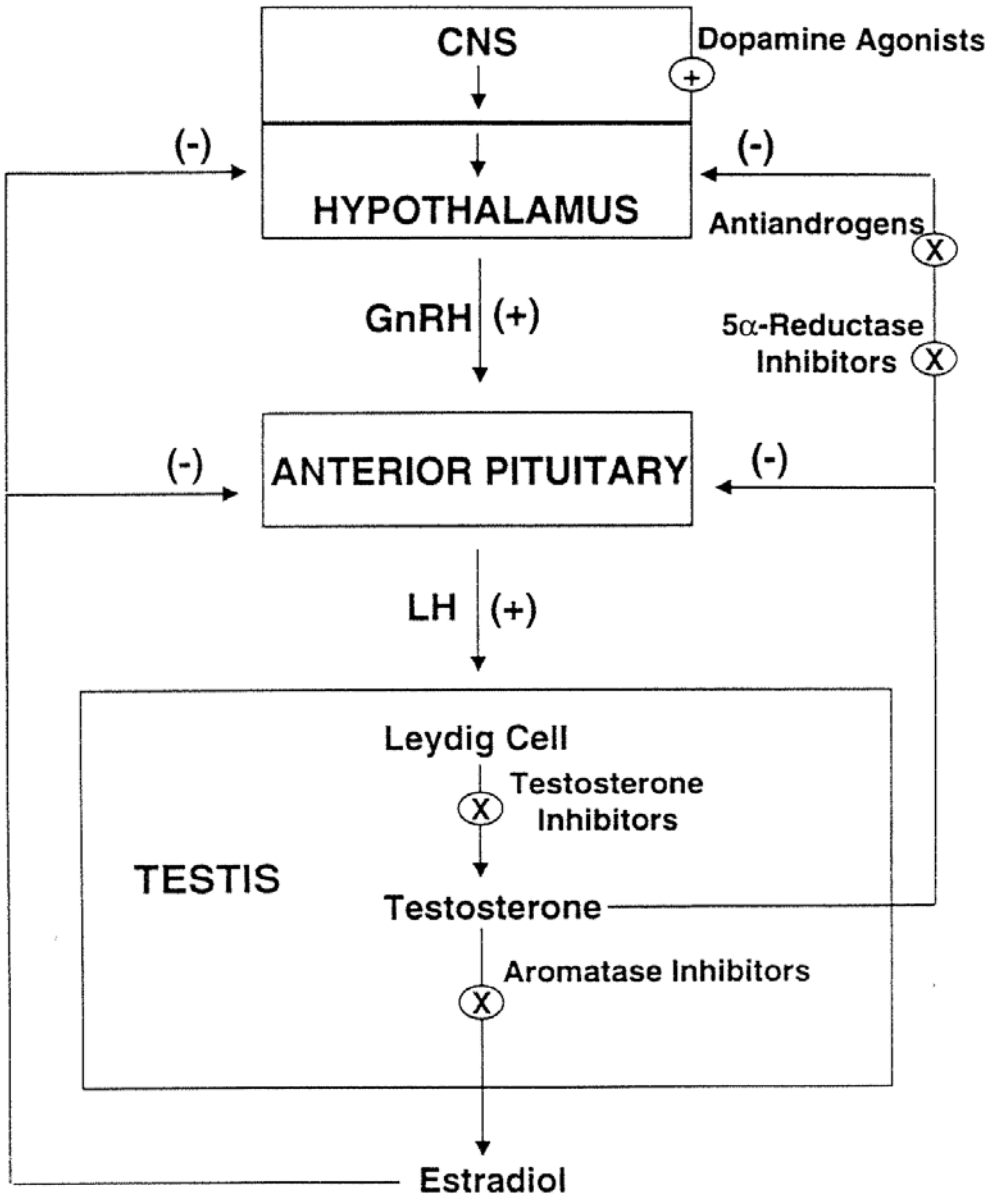


Figure 3-1. Regulation of the HPT-axis and control points for potential disruption. Symbols: (+) feedback stimulation; (-) feedback inhibition; ⊕ receptor stimulation; ⊗ enzyme or receptor inhibition (Cook et al. (no date); Cook et al., 1999).

In Tables 1 and 2 of the Addendum, compounds are listed that produce LC hyperplasia and/or LCTs in rats, mice or dogs. Because of the implications in risk assessment, DNA-reactive and non-DNA-reactive compounds are considered separately (Clegg et al., 1997).

3.2.1 Direct DNA reactive mechanisms

Most of the genotoxic compounds that induce LCTs also induce adenomas and/or carcinomas at other sites (see Table 2 in the Addendum; Cook et al., 1999), and LCT is one of the endpoints to be considered (Clegg et al., 1997)

3.2.2 Non-DNA reactive mechanisms

The non-genotoxic compounds are subdivided by their mode of action, chemical activity, chemical class or as other (Table 1 in the Addendum) (Cook et al., 1999).
There are 7 modes of action that appear to affect hormonal control of LC activity (Clegg et al., 1997; Cook et al., 1999).

5. Androgen receptor antagonism: competition with testosterone and dihydrotestosterone (DHT) for binding to the androgen receptor. This competition reduces the net androgenic signal to the hypothalamus and pituitary resulting in an increase in LH with a concomitant elevation of testosterone.
6. Testosterone biosynthesis inhibition: decrease in testosterone levels increases LH levels in rats, resulting in the development of LCTs.
7. 5α-Reductase inhibition: blocks the conversion of testosterone to DHT. DHT amplifies the androgenic signal through several mechanisms: (i) unlike testosterone, DHT cannot be aromatized to estrogen and thus its effect is purely androgenic and (ii) DHT binds to the androgenic receptor with greater affinity and stability than testosterone. Hence, 5α-reductase inhibitors decrease DHT levels, which reduce the net androgenic signal received by the hypothalamus and pituitary and thereby causes a compensatory increase in LH levels. 5α-reductase inhibitors induce LCTs in mice and LC hyperplasia in rats.
8. Aromatase inhibition: blocks the conversion of testosterone to estradiol, resulting in a decrease in estradiol and an increase in LH levels. (In some chronic studies, dogs have been reported to be more sensitive than rats for the development of LC hyperplasia in response to aromatase inhibition (Clegg et al., 1997; Walker and Nogués, 1994).
9. Estrogen agonism: induces LCTs in certain strains of mice (Alderly Park outbred, BALB/c, Strong A) but not in rats (Cook et al. (no date). These strain differences have been attributed to 2 factors: (i) estrogen increase LH levels in mice, but decrease LH levels in rats and (ii) estrogen may directly stimulate LC proliferation via a paracrine mechanism (Clegg et al.,1997; Cook et al., 1999).
10. GnRH agonism: the only documented examples of a non-LH type mechanism that can induce LCTs. GnRH agonists induce LCTs in rats by binding to GnRH (or LHRH, luteinizing hormone releasing hormone) receptors on Leydig cells. Because Leydig cells from mice, monkey and human do not contain GnRH (LHRH) receptors, these species are believed to be not susceptible to tumour induction by this class of compounds (Cook et al. (no date); Clegg et al., 1997; Donaubauer et al., 1987; Hunter et al., 1982; Prentice and Meikle, 1995).
11. Dopamine agonism: decreases serum prolactin levels in rats, which causes downregulation of LH receptors on LCs, and thus a decrease in testosterone production. Decreased testosterone results in an increased LH level. This effect has not been reported for any other species than rat (and decreased prolactin does not decrease the number of LH receptors on human LC). An alternative mechanism has been proposed, namely that dopamine agonists increase GnRH levels that subsequently increases LH levels. The relative contribution of these two mechanisms toward the development of LCTs remains to be determined (Cook et al., 1999).

3.2.3 Other possible mechanisms

Since many compounds are able to induce LC hyperplasia and/or tumours by mechanisms other than the HPT-axis disrupting mechanisms described above, many studies have focussed on additional modes of action for LCT induction, although the contribution of the suggested mechanisms needs to be elucidated. Additional factors with LCT-inducing potential are discussed below.

1. In addition to the direct effects of elevated levels of LH, proto-oncogene activation appears to be also a consequence of LH stimulation of LC, and might provide a common underlying mechanism (Clegg et al., 1997).
2. Testosterone production is also suppressed by glucocorticoids in rats (elevated after stress) (Cook et al., 1999). It is proposed that stress (indicated by an increase in serum corticosteroid levels), related to individual caging, particular among males, directly impairs testosterone synthesis and produce LC atrophy, which leads to a feedback increase in the synthesis of LH by the anterior pituitary (Nyska et al., 1998). This phenomenon has been observed in toxicity studies with rats. (note: rats in inhalation and dermal toxicity studies are singly caged, whereas in feed or gavage studies, rats are group-caged (Nyska et al., 1998).
3. While it is clear that normal functioning of the LCs is dependent on an appropriate endocrine environment within the testis, primarily that provided by LH stimulation, the paracrine environment also plays an important role (seminiferous damage, IGF-1, TGFα, IL-1), (Ewing, 1992; Prentice and Meikle, 1995; Cook et al., 1999).
4. Several peroxisome proliferators have been shown to induce LCTs in rats, without inducing peroxisomes in the Leydig cells (whereas peroxisome proliferation was observed in the liver) (Clegg et al., 1997; Cook et al. (no date).
5. Besides cAMP, other second messenger systems (arachidonic acid, leukotrienes, calcium/calmoduline, chloride ions, free radicals (ref. 28 in Clegg et al., 1997; Cook et al., 1999) may also be involved the induction of LC hyperplasia (Clegg et al., 1997).

3.3 Normal values and natural variation

The spontaneous incidence of LCTs in animals is species, strain and age dependent (Chandra and Riley, 1994; Clegg et al., 1997; Cook et al. (no date); Cook et al., 1999). See Tables 3 and 4 in the Addendum.
The comparison of spontaneous incidences between animal and man is tenuous because 1) the diagnosis in animals is from histological examination from animal experiments (standard observation) while in man (with complaint) from palpation, and 2) some rat strains are (very) sensitive (high LCT incidences). In spite of these differences, the data in Table 3-1 suggest a substantially lower occurrence of LCT in humans compared to rats (Clegg et al., 1997).

Table 3-1: LCT incidences

Species	**LCT incidence (%)**
Rat	0.8 - 100
Mouse	0.1 – 2.5
Dog	0.03 - 16
Human	$0.3 \times 10^{-6} – 3 \times 10^{-6}$

Rat
In Sprague-Dawley, Osborne-Mendel and Brown-Norway rats, LCTs are generally very rare, whereas their incidence reaches 90-100% in 18-24 month-old F344 rats (Weisburger et al., 2002). Wistar-derived strains exhibit variable incidences ranging from less than 10% to nearly 100% (Bär, 1992; Chandra et al., 1992; Clegg et al., 1997). Histologically, about 90% of LCTs in rats are considered to be benign (Chandra and Riley, 1994; Prentice and Meikle, 1995). In the remaining 10% malignant tumours, distant organ metastasis is extremely rare (Chandra and Riley, 1994). Tennekes et al. (2004) observed no significant time-related changes in the incidence of LCT's in the testes in any laboratory rat strain.
Mouse
Spontaneous incidence of LCTs in mice older than 18 months is considerably lower than in rats, and ranges from 0.1 to 2.5% (Chandra and Frith, 1992 and ref. 4 and 5 in Clegg et al., 1997).

Dog
The spontaneous incidence of LCTs in beagle dogs at 7.75 years is 6.3%, and a further 8% showed Leydig cell hyperplasia (James and Heywood, 1979; Cook et al., 1999), whereas the reported incidence for dogs in general ranges from 0.034% to 16%. In dogs, the tumours are mostly benign, and metastases are observed in only 2% (Peters and Sluijs, 1996).
Human
The estimated incidence of LCTs is 0.3 to 3 per million (Bär, 1992; Cook et al., 1999). In man, LCTs are distributed equally in various age groups (Schottenfeld et al., 1980; Prentice and Meikle, 1995). The incidence of LCTs appears to vary by ethnic background, where the highest incidence is seen in white males (Schottenfeld et al., 1980; Cook et al., 1999). An important risk factor is cryptorchidism, which occurs in greater frequency in white men (Schottenfeld et al., 1980; Bosland, 1996). Prenatal exposure to estrogens, specifically DES, has been shown to be a risk factor for cryptorchidism although not in all studies (Bosland, 1996).
Children with LCT are more likely to demonstrate a definite clinical syndrome and may develop isosexual precocious pseudopuberty or feminization. Adult men may show feminization but often demonstrate no symptoms other than a testicular mass. Children with LCTs usually show a high urinary excretion of 17-ketosteroids, while adults have normal values, unless the tumour in the adult is metastatic. Probably all feminizing LCTs produce estrogens (high urinary estrogen excretion and high spermatic vein estradiol to testosterone ratio) (Freeman, 1991).

3.4 Sensitive species/groups

Interspecies
Although in literature high incidences of natural occurring LCTs are described for both rats and dogs, studies on mechanisms of LCT induction are usually restricted to rats (and mice), and also most carcinogenity studies are performed with rodents. Therefore, most information is available on rats. Information on interspecies differences are mainly restricted to rats and humans.
Although the anatomy and regulation of the HPT-axis are generally comparable between rats, mice, dogs, monkeys and humans, there are some differences, which may play a role in the differences in sensitivity of induction of LCTs in these species. The interspecies differences include:

1. Rats lack sex hormone binding globulin (SHBG). In man, about 95% of testosterone in peripheral blood is bound to SHBG, which retards its metabolism and clearance (Cook et al., 1999). Because the ratio between bound and free (bioavailable) testosterone is

kept in balance, it is relatively difficult in man to perturb the peripheral levels of testosterone in any short-term way. In contrast to man, the rat has no peripheral SHBG and thus the blood levels of testosterone can potentially be altered more rapidly (Cook et al., 1999). The half life of circulating LH in humans is in excess of 100 minutes, while in the rat the half life is 5 to 10 minutes (Cook et al., 1999).

2. Rats have a greater LH receptor number per Leydig cell than humans (human LCs contain approximately 1500 LH receptors/cell, whereas rat LCs contain approximately 20000 receptors/cell, a 13-fold difference between rats and humans) (Prentice and Meikle, 1995; Clegg et al., 1997; Cook et al., 1999).
3. Rats have GnRH receptors on their LCs, in contrast to mice, monkeys and humans (Prentice and Meikle, 1995; Clegg et al., 1997).
4. Prolactin modulates LH receptor levels on LCs in rats, but not in humans (Clegg et al., 1997).
5. Rats and humans appear to respond different to exogenous hCG (human Chorionic Gonadotropin, a hormone equivalent to LH in its action on LCs), with rat LC showing hyperplasia and human LC showing hypertrophy (Prentice and Meikle, 1995; Clegg et al., 1997).
6. Mice and monkeys appear to be less sensitive than rats to androgen receptor antagonists. There is no information available for humans (Clegg et al., 1997).
7. 5α-reductase inhibitors induce LC hyperplasia in rats and LC hyperplasia and LCTs in mice. There is no information available for humans (Clegg et al., 1997).
8. Estrogen agonists induce LCTs in mice by mechanisms that do not appear to be present in rats. No information is available for humans (Clegg et al., 1997).
9. Histologic and electron microscopic features of malignant LCTs in rats resemble that of human LCT, except for the presence of crystalloids of Reinke which are present in 40% of the human LCTs, and are not observed in rats (Chandra and Riley, 1994; Cook et al., 1999). The function of Reinke's crystals is unknown (Naughton et al., 1998). The presence of Reinke crystals is unique for human LC and the Australian wild bush rat (Cook et al., 1999).
10. Testosterone levels decline with age in most strains of rats as well as in humans. In rats, this decrease is probably secondary to declining LH levels. This is in contrast to the situation in man, where LH levels tend to increase with age, presumably related to decreasing testosterone levels (Cook et al., 1999).

All of these differences may contribute to the observed higher incidence (and probably greater sensitivity for induction) of LCH/LCT in rodents when compared with human.

Intraspecies

Apart from strain differences in the rat and the observed higher incidence of LCTs in older dogs (in absence on information in the literature on the mechanisms of induction of LCTs in older dogs), there are no indications for specific sensitive groups within a species.
Human variation is described in '3.3 Normal values and natural variation.'

3.5 Assessment and RIVM/SIR strategy

Known human incidence of clinically detectable LCTs in the general population is very low. However, occurrence of LC hyperplasia is unlikely to be detected in men of any age since routine autopsy typically does not include microscopic evaluation of the testes (Clegg et al.,

1997; Cook et al., 1999). The true occurrence of LCTs in humans may be higher than currently thought, but is nonetheless significantly lower than in rats (Prentice and Meikle, 1995; Cook et al., 1999).
The pathways for the regulation of the HPT-axis of rats and humans are qualitatively similar. It appears that there is evidence that suggests that human LCs are, quantitatively, less sensitive than rats in their proliferation response to LH and hence their sensitivity to chemically induced LCTs (Cook et al., 1999). Nevertheless, the available evidence for most mechanisms of action is insufficient to conclude that they are not relevant for humans. For LCT induction there are, based on the current state of knowledge, only two mechanisms for which it can be concluded that they are not relevant to humans: GnRH and dopamine agonists. Because testicular GnRH and Prolactin receptors are either not expressed or expressed only at a very low level in humans, the induction of LCTs in rats by GnRH and dopamine agonists would appear to be not relevant to humans. However, the relevance to humans of the remaining five mechanisms of action cannot be ruled out (Cook et al., 1999).

Generally, only in cases where a species-specific mechanism is involved in the induction of LCTs, and there is sufficient evidence that the mechanism of LCT induction is not relevant in humans, the increases of LCTs in the testspecies are not considered in the risk assessment.
In circumstances in which the mechanism of induction is unknown, it should be assumed that humans are potentially susceptible (Clegg et al., 1997). This view is in line with the view of the EPA and Classification and Labelling on LC hyperplasia/neoplasia (EPA Reregistration Eligibility Decision (RED)., 1998; EPA, 1996; EPA, 1998; Risk Assessment Forum U.S. Environmental Protection Agency, 1997; Document ECBI/61/03 Classification and Labelling, Draft, 2003). In 2004, the Specialised Experts Meeting of the European Chemicals Bureau (ECB) limited their advice to the Working Group on Classification and Labelling to non-genotoxic compounds and in this context agreed that findings of Leydig cell tumours in rats or mice would normally lead to classification for carcinogenicity in Category 3. No classification is needed for substances causing LCT's in rats by perturbating the HPT axis by a mechanism that is proven not relevant for humans: dopamine agonists and GnRH agonists. Data on LCT's in F344 rats or other strains having comparably high spontaneous LCT rate are considered normally not informative (ECBI/08/04 Add.4, 2004).

The distinction between LC hyperplasia and LC tumour is based on size. There are no uniform used/accepted size criteria although the criteria of the Society of Toxicologic Pathologists are frequently used. Since it is considered that the transition from LC hyperplasia to LCT is a continuous spectrum of change, both hyperplasia and neoplasia are considered in the risk assessment. For the risk assessment, the relevance of increased LCTs in laboratory animals is considered separately for genotoxic and non-genotoxic compounds.

1. Only in cases where a species-specific mechanism is involved in the induction of LCTs, and there is sufficient evidence that the mechanism of LC induction is not relevant in humans (in rat GnRH and dopamine agonists), the compound inducing the tumour is considered as carcinogenic for the species concerned, but not for humans (see above).
2. Genotoxic compounds.
   Most of the genotoxic compounds that induce LCTs also induce adenomas and/or carcinomas at other sites and LCT is one of the endpoints to be considered in the risk assessment. Even when only LCT are observed, it cannot be excluded that these substances induce tumours at other sites in humans. Therefore, the LCTs are considered to be relevant for human risk assessment.
3. Non-genotoxic compounds.

An increase in LCT number has to be statistically significant and dose-related for considering in the risk assessment as evidence of carcinogenicity. A statistically significant increase may also occur in the high dose only. Accompanying LC hyperplasia contributes to the weight of evidence of the neoplastic response. However, a compound is not designated carcinogenic based on increased hyperplasia only.

Expert consultation is needed:
- when the increase in LC hyperplasia/neoplasia is the critical effect for determining the NOAEL of a study;
- when a non-significant increased incidence in LCTs, dose-related and/or accompanied by hyperplasia, is observed in species/species strains, which have a low spontaneous incidence of LCTs (see the Addendum);
- when only a significantly (dose-related) increased incidence of hyperplasia is observed;
- when the LCT incidence in the control group deviates from the normal background incidence in that species/species strain, and the LCT numbers observed in the dosed groups may pose a problem.

Note: there is a different approach on the relevance of increased LCT incidences in F344 rats or other strains having comparably high spontaneous LCT rate between RIVM strategy and the view of the Specialised Experts of the working Group on Clasification and Labelling.

References

Alison RH, Capen CC and Prentice DE. 1994. Neoplastic lesions of questionable significance to humans. Toxicologic Pathology 22(2): 179-186.

Bär A. Significance of Leydig cell neoplasia in rats fed lactitol or lactose. 1992. Journal of the American College of Toxicology 11(2): 189-207.

Bosland MC. 1996. Hormonal factors in carcinogenesis of the prostate and testis in humans and animal models. Prog. Clin. Biol. Res. 394: 309-352.

Chandra M. and Frith CH. 1992. Spontaneous neoplasms in aged CD-1 mice. Toxicology Letters 61: 67-74.

Chandra M. and Riley MGI. 1994. Rarely occurring spontaneous metastasizing testicular tumors in rats. Histopathologic and ultrastructural features. Exp. Toxic. Pathol. 46: 155-161.

Chandra M, Riley MGI and Johnson DE. 1992. Spontaneous neoplasms in aged sprague-dawley rats. Arch. Toxicol. 66: 496-502.

Clegg ED, Cook JC, Chapin RE, Foster PMD and Daston GP. 1997. Leydig cell hyperplasia and adenoma formation: mechanisms and relevance to humans. Reproductive Toxicology 11(1): 107-121.

Cook JC, Frame SR, and Obourn, JD. (no date) Leydig cell tumors. Sipes, IG, McQueen, ChA and Gandolfi A. Comprehensive Toxicology. volume 10 ed.
Notes: http://127.0.0.1/toc/TOC1014.HTM

Cook JC, Klinefelter GR, Hardisty JF, Sharpe RM and Foster PMD. 1999. Rodent Leydig cell tumorigenesis: a review of the physiology, pathology, mechanisms, and relevance to humans. Critical Reviews in Toxicology 29(2): 169-261.

Donaubauer HH, Kramer M, Krieg K, Mayer D, Von Rechenberg W, Sandow J and Schütz E. 1987. Investigations of the carcinogenicity of the LH-RH analog Buserelin (HOE 766) in

rats using the subcutaneous route of administration. Fundamental and Applied Toxicology 9: 738-752.
EPA, Federal Register. Part IV 40 CFR Parts 180 and 185. Revocation of pesticide food additive regulations; final rule.C. Iprodione. [Web Page]. Accessed 1996 Jul 29. Notes: Fedreal Register/ Vol. 61, no. 146/ Rules and Regulation, pp. 39527-39540
EPA, Pesticide Fact Sheet. Tralkoxydim, conditional registration. [Web Page]. Accessed 1998 Dec 4.
EPA Reregistration Eligibility Decision (RED). Iprodione [Web Page]. Accessed 1998 Nov. Notes: EPA738-R-98-019
Ewing LL. 1992. The Leydig cell. In: Reversibility in testicular toxicity assessment. Scialli AR and Clegg ED (eds.). CRC Press, Boca Raton, FL. pp. 87-126.
Freeman DA. 1991. Steroid hormone-producing tumors of the adrenal, ovary, and testes. Endocrinol. Metab. Clin. North Am. 20: 751-766.
Hunter MG, Sullivan MHF, Dix CJ, Aldred LF and Cooke BA. 1982. Stimulation and inhibition by LHRH analogues of cultured rat Leydig cell function and lack of effect on mouse Leydig cells. Molecular and Cellular Endocrinology 27: 31-44.
James RW and Heywood R. 1979. Age-related variations in the testes and prostate of beagle dogs. Toxicology 12: 273-279.
Naughton CK, Nadler RB, Basler JW and Humphrey PA. 1998. Leydig cell hyperplasia. Review. British Journal of Urology 81: 282-289.
Nyska A, Leininger JR, Maronpot RR, Haseman JK and Hailey JR. 1998. Effect of individual versus group caging on the incidence of pituitary and Leydig cell tumors in F344 rats: proposed mechanism. Medical Hypotheses 50: 525-529.
Peters MAJ and Sluijs FJ. 1996. Testistumoren bij de hond: een literatuuroverzicht. Tijdschrift Voor Diergeneeskunde 121(2): 36-38.
Prentice DE and Meikle AW. 1995. A review of drug-induced Leydig cell hyperplasia and neoplasia in the rat and some comparison with man. Human and Experimental Toxicology 14: 562-572.
Risk Assessment Forum U.S. Environmental Protection Agency. 1997. Special report on environmental endocrine disruption: an effects assessment and analysis. [Web Page]. Available at: http://www.epa.gov//reg5ogis/risk/endocrin.pdf. Notes: EPA/630/R-96/012
Schottenfeld D, Warshauer ME, Sherlock S, Zauber AG, Leder M and Payne R. 1980. The epidemiology of testicular cancer in young adults. American Journal of Epidemiology 112(2): 232-246.
Tennekes H, Kaufmann W, Dammann M and van Ravenzwaay B. 2004. The stability of historical control data for common neoplasms in laboratory rats and the implications for carcinogenic risk assessment. Reg. Tox Pharmacol. 40: 293-301.
Walker UJ and Nogués V. 1994. Changes induced by treatment with aromatase inhibitors in testicular Leydig cells of rats and dogs. Exp. Toxic. Pathol. 46: 211-213.
Weisburger JH, Rivenson A, Reinhardt J, Pittman B and Zang E. 2002. On the occurrence of Leydig cell tumors in the F344 rat. Cancer Lett. 128(2): 213-216.
ECBI/61/03. oct. 2003. A “Thought Starter” for developing an agreed position on the relevance of leydig cell tumours in rats to humans. Draft Document UK.
ECBI/08/04. Ispra meetings 2004 (January 22-23 ECBI/08/04 Rev.2; April 2-23 ECBI/08/04 Add.4) Commision Working Group of Specialised Experts in the fields of of Carcinogenicity, Mutagenicity and Reprotoxicity. 2. Leydig Cell Tumours.

Addendum

Table 1 Nongenotoxic compounds that produce LC hyperplasia or adenomas in rats, mice or dogs

Compound (CAS number)	Species (strain)[a]	LC response	Other tumor sites	Ref.
A. Classified by Mode of Action				
1. Androgen receptor antagonists				
Bicalutamide (90357-06-5)	Rat (Wistar)	Adenoma	Thyroid	Iswaran et al., 1998
Cimetidine (51481-61-9)	Rat (Wistar)	Adenoma	None	Leslie et al., 1981; Sivelle et al., 1982, Brimblecombe and Leslie, 1984
Fenvalerate (51630-58-1)	Rat (SD)	Adenoma (equivocal)	None	Parker et al., 1984; Eil and Nisula, 1990
Flutamide (13311-84-7)	Rat	Adenoma	None	PDR, 1995d; Viguier-Martinez et al., 1983a,b; Cook et al., 1993
Linuron (330-55-2)	Rat (CD)	Adenoma	None	Cook et al., 1993
Procymidone (32809-16-8)	Rat (Osborne-Mendel)	Adenoma	None	Hosokawa et al., 1993a,b; Murakami et al., 1995
Vinclozolin (50471-44-8)	Rat	Adenoma	None	Gray et al., 1994; Wong et al., 1995
Zanoterone (107000-34-0)	Dog (Beagle)	Hyperplasia	None	Juniewicz et al., 1990
2. 5α -Reductase inhibitors				
Finasteride (98319-26-7)	Rat (CD)	Hyperplasia	None	PDR, 1995m; Prahalada et al., 1994, George et al., 1989
	Mouse (CD-1)	Adenoma	None	PDR, 1995m; Prahalada et al., 1994
3. Testosterone biosythesis inhibitors				
Calcium Channel Blockers (see Section B)	Rat	Adenoma	None	PDR, 1995c,i,l; Roberts et al., 1989
Cimetidine (51481-61-9)	Rat (Wistar)	Adenoma (equivocal)	None	Leslie et al., 1981; Brimblecombe and Leslie, 1984; Morita et al., 1990

Table 1 continued

Ethanol (64-17-5)	Rat (CD)	Adenoma (equivocal)	None	Cheever et al., 1990; Widenius et al., 1989
Lansoprazole (103577-45-3)	Rat (CD)	Adenoma	None	Fort et al., 1995; Meikle et al., 1994
Lead acetate (301-04-2/15347-57-6)	Rat (Wistar)	Adenoma (equivocal)	None	Zawirska and Medras, 1968
Metronidazole (443-48-1)	Rat (Wistar)	Adenoma	Pituitary	Rustia and Shubik, 1979,
Vinclozolin (50471-44-8)	Rat	Adenoma	None	Ronis et al., 1994; Gray et al., 1994; Wong et al., 1995
4. Aromatase inhibitors				
Formestane (566-48-3)	Dog (Beagle)	Hyperplasia	None	Junker-Walker and Nogues, 1994
Letrozole (112809-51-5)	Dog (Beagle)	Hyperplasia	None	Junker-Walker and Nogues, 1994
5. Dopamine agonists/enhancement of dopamine levels				
Mesulergine (64795-35-3)	Rat (Wistar)	Adenoma	None	Prentice et al., 1992; Dirami et al., 1996
Norprolac (87056-78-8)	Rat (CD)	Adenoma	None	Roberts et al., 1993
Oxolinic acid (14698-29-4)	Rat (Wistar)	Adenoma	None	Yamada et al., 1994a,b; Yamada et al., 1995a,b
6. GnRH agonists				
Buserelin (57982-77-1)	Rat (Wistar)	Hyperplasia	None	Donaubauer et al., 1987
Histrelin (76712-82-8)	Rat	Adenoma	Pituitary	PDR, 1995o
Leuprolide (74381-53-6)	Rat	Adenoma	Pituitary	PDR, 1995h
Nafarelin (76932-56-4)	Rat	Adenoma	Pituitary	PDR, 1995p

Table 1 continued

B. Grouped by Chemical Activity

1. Antihypertensives				
Guanadrel (22195-34-2)	Rat	Adenoma	None	PDR, 1995f
Hydralazine (86-54-4)	Rat	Adenoma	Liver	PDR, 1995a
2. Calcium channel blockers				
Felodipine (72509-76-3)	Rat	Adenoma	None	PDR, 1995l
Isradipine (75695-93-1)	Rat (CD)	Adenoma	None	Roberts et al., 1989; PDR, 1995c
Lacidipine (103890-78-4)	Rat (CD)	Adenoma	None	Hamada and Futamura, 1998
Nimodipine (66085-59-4)	Rat (Wistar)	Adenoma	None	PDR, 1995i
3. Fungicides				
Procymidone (32809-16-8)	Rat (Osborne-Mendel)	Adenoma	None	Hosokawa et al., 1993a,b; Murakami et al., 1995
Vinclozolin (50471-44-8)	Rat	Adenoma	None	Wong et al., 1995, Ronis et al., 1994
Folpet (133-07-3)	Rat (CD)	Adenoma	Thyroid	Quest et al., 1993
4. Goitrogens				
Ethylenethiourea (ETU) (96-45-7)	Rat	Adenoma (equivocal)	None	Gak et al., 1976
6-*n*-Propyl-2-thiouracil (PTU) (51-52-5)	Rat (SD)	Hyperplasia	None	Mendis-Handagama and Sharma, 1994
5. Peroxisome proliferators				
Ammonium Perfluorooctanoate (C8) (3825-26-1)	Rat (CD)	Adenoma	Liver pancreas	Sibinski, 1987; Cook et al., 1992; Cook et al., 1994
Clofibrate (637-07-0)	Rat (Alderley Park)	Adenoma	Liver pancreas	PDR, 1995b
Diethylhexylphthalate (DEHP) (117-81-7)	Rat (SD)	Adenoma	None	Berger, 1995
Gemfibrozil (25812-30-0)	Rat (CD)	Adenoma	Liver pancreas	Fitzgerald et al., 1981
HCFC-123 (306-83-2)	Rat (CD)	Adenoma	Liver pancreas	Malley et al., 1995
Methylclofenapate (21340-68-1)	Rat (Alderley Park)	Adenoma	Liver pancreas	Tucker and Orton, 1995
Perchloroethylene (PCE) (127-18-4)	Rat (F344)	Adenoma (equivocal)	Leukemia kidney skin	Clarke and Ragan, 1986
Trichloroethylene (TCE) (79-01-6)	Rat (SD)	Adenoma	Leukemia kidney	Maltoni et al., 1988
Wyeth-14,643 (50892-23-4)	Rat (CD)	Adenoma	Liver pancreas	Cook et al., 1994

Table 1 continued

Compound (CAS number)	Species (strain)[a]	LC response	Other tumor sites	Ref.
C. Grouped by Chemical Class				
1. Fluorochemicals				
HCFC-123 (306-83-2)	Rat (CD)	Adenoma	Liver pancreas	Malley et al., 1995
HCFC-133a (75-88-7)	Rat (Wistar)	Adenoma	None	Longstaff et al., 1984
HFC-134a (811-97-2)	Rat (Wistar)	Adenoma	None	Collins et al., 1995
HCFC-141b (1717-00-6)	Rat (CD)	Adenoma	None	Turnbull et al., 1994
2. Nitroaromatics and related compounds				
p-Nitrochlorobenzene (100-00-5)	Rat (SD)	Adenoma	None	Schroeder and Daly, 1984
Nitroglycerine (55-63-0)	Rat (CD)	Adenoma	Liver	Ellis et al., 1984; PDR, 1995j
2,4-Toluenediamine (95-80-7)	Rat (F344)	Adenoma	Liver pancreas	Cardy, 1979
3. Organochlorines				
o,p'-DDD	Rat	Adenoma (equivocal)	None	Lacassagne and Hurst, 1965
o,p'-DDT (789-02-6)	Rat (Osborne-Mendel)	Adenoma (equivocal)	Liver	Fitzhugh and Nelson, 1947
4. Sugars				
Lactose (63-42-3)	Rat (Wistar)	Adenoma	None	Sinkeldam et al., 1992
Lactitol (585-86-4)	Rat (Wistar)	Adenoma	None	Sinkeldam et al., 1992
Tara gum	Rat (F344)	Adenoma (equivocal)	None	Melnick et al., 1983

Table 1 continued

Compound (CAS number)	Species (strain)[a]	LC response	Other tumor sites	Ref.
D. Unclassified				
Boric Acid (10043-35-3)	Mouse (B6)	Hyperplasia	None	Dieter, 1994
Carbamazepine (298-46-4)	Rat (SD)	Adenoma	None	PDR, 1995r
Felbamate (25451-15-4)	Rat	Adenoma	Liver	PDR, 1995e
Flecainide (54143-55-4)	Rat (Wistar)	Adenoma	None	Case et al., 1984
Indomethacin (53-86-1)	Rat (SD)	Adenoma	None	Goerttler et al., 1992
Isopropanol (67-63-0)	Rat (F344)	Adenoma	None	Burleigh-Flayer et al., 1997
JP-4	Rat (F344)	Adenoma	Kidney	Bruner et al., 1993
d-Limonene (5989-27-5)	Rat (F344)	Adenoma	Kidney	Jameson, 1990
MTBE (1634-04-4)	Rat (CD, F344)	Adenoma	None	Belpoggi et al., 1995
Nicotine (54-11-5)	Rat (F344)	Hyperplasia (equivocal)	None	Thompson et al., 1973
Oxazepam (604-75-1)	Rat	Adenoma	Liver thyroid prostate	PDR, 1995n

[a] Strain is included when specified in citation.

Table 2 Genotoxic compounds that produce LC hyperplasia or adenomas

Compound (CAS number)	Species (Strain)[a]	LC response	Other tumor sites	Ref.
1. Alkylating agents				
N-Nitrosobis-(2-oxopropyl)amine (BOP) (60599-38-4)	Rat (Wistar)	Adenoma	Prostate, vas deferens, coagulating glands, liver	Pour, 1986
1,3-Butadiene (106-99-0)	Rat (CD)	Adenoma	Pancreas	Owen et al., 1987
BrdU (59-14-3)	Rat (LIO)	Adenoma	Kidney, intestine	Anisimov, 1995
BrdU + X-rays (59-14-3)	Rat (LIO)	Adenoma	Prostate, kidney, adrenal cortex, hematopoietic system, thyroid	Anisimov and Osipova, 1993
3-Chloro-2-methylpropene (563-47-3)	Rat (F344)	Adenoma	Forestomach, kidney	NTP, 1986
Cycasin (14901-08-7)	Rat (ACI)	Adenoma	Intestine, liver, kidney	Fukunishi et al., 1971
Dibromochloropropane (DBCP) (96-12-8)	Humans	Hyperplasia	None	Cortes-Gallegos et al., 1980
Diethylnitrosoamine (DEN) (55-18-5)	Mice (RF)	Adenoma	Lung, liver, forestomach	Clapp, 1973
Dimethylnitrosoamine (DMN) (62-75-9)	Rat (Wistar)	Adenoma	Liver, hematopoietic system	Arai et al., 1979; Terao et al., 1978
Isoprene (78-79-5)	Rat (F344)	Adenoma	None	Melnick et al., 1994
8-Methoxypsoralen (298-81-7)	Rat (F344)	Adenoma	Kidney, Zymbal's, lung	Dunnick, 1989
Methy-CCNU (33073-59-5)	Mice (SJL/J)	Hyperplasia	None	Yegana et al., 1988
Nitrosoethylene (NEU) (759-73-9)	Rat (Outbred)	Adenoma	Nervous system	Ird and Smirnova, 1983
Streptozotocin (18883-66-4)	Rat (SD)	Adenoma	Pancreas, kidney, liver	Okawa and Doi, 1983
2. Base substitutions				
5-Azacytidine (320-67-2)	Rat (F344)	Adenoma	Hematopoietic system, skin, lung, kidney	Carr et al., 1988; Carr et al., 1984
Vidarabine (5536-17-4)	Rat	Adenoma	Liver, kidney, intestine, thyroid	Griffith, 1988
3. Metals				
Cadmium (7440-43-9)	Rat (Wistar/F344)	Adenoma	Lung, prostate, hematopoietic system	Bomhard et al., 1987; Waalkes and Rehm, 1992; Waalkes et al., 1997
4. Radiation				
X-Irradiation	Rat (Long-Evans)	Adenoma	None	Lindsay et al., 1969
5. Others				
Ethylene dichloride + Disulfiram (107-06-2/97-77-8)	Rat (SD)	Adenoma	Liver, skin	Cheever et al., 1990

[a] Strain is included when specified in citation.

Table 3 Incidence of LCTs of the testes in male rats for studies terminated at 24 months

Strain	Diagnosis	Incidence	(%)	Ref.
Wistar	Adenoma	87/1249	7.0	Bomhard and Rinke, 1994
Crl:CD®BR	Adenoma	37/721	5.1	Hansen, 1993
Sprague-Dawley	Adenoma	16/349	4.6	Lewis, 1993
F344/DuCrj	Adenoma	537/569	94.4	Iwata et al., 1991
Crl:CD®BR	Adenoma	38/585	6.5	McMartin et al., 1992
Sprague-Dawley	Adenoma	11/1340	0.82	Chandra et al., 1992
	Carcinoma	1/1340	0.07	
Wistar	Adenoma	27/685	3.94	Walsh and Poteracki, 1994
F344	Adenoma	39,253/51,230	76.6	Mitsumori and Elwell, 1988
Crl:CD®BR	Adenoma	59/1260	4.68	Lang, 1992
	Carcinoma	1/1260	0.08	
CDF(F344)/CrlBR	Adenoma	741/946	78.3	Lang, 1990
Crl:CD®SD	Adenoma	29/496	5.8	IRI, 1995b

Table 4 Incidence of LCTs of the testes in male mice for studies terminated at 24 months

Strain	Diagnosis	Incidence	(%)	Ref.
Swiss CD-1®	Adenoma	—	≤1.0	Engelhardt et al., 1993
Crl:CD-1® BR	Adenoma	13/524	2.48	Lang, 1995
IRC Crj:CD-1®	Adenoma	8/891	0.9	Maita et al., 1988
CD-1®	Adenoma	14/725	1.9	Chandra and Frith, 1992
	Carcinoma	1/725	0.14	
B6C3F1	Adenoma	169/46,752	0.4	Mitsumori and Elwell, 1988
Crl:CD-1® Swiss	Adenoma	8/400	2.0	IRI, 1995a

4. Proposal for the interpretation of leaching study data for wood preservatives (biocides)

Factsheet FSM/011, 30-8-2006

Author: R.H.L.J. Fleuren

Contents:

4.1 Introduction

When performing a risk assessment for wood preservatives, results from leaching studies are crucial, since they are the most important input parameter for almost all emission scenarios.
In a leaching study, a piece of wood (treated with a wood preservative) is submerged in water. Next, the water is analysed for the amount of wood preservative that was leached from the wood (mostly in $mg.m^{-2}$) within a certain time interval and based on these results a leaching rate, also called flux (mostly in $mg.m^{-2}.d^{-1}$), can be calculated. This flux is used as input parameter for emission scenarios.
Even though the leaching test is relatively simple to perform, the interpretation of the results is considered difficult. This is caused by a variety of factors.
Firstly, the leaching rate depends strongly on wood species, preparation, impregnation method and is influenced by several environmental conditions (e.g. temperature, pH).
Secondly, even if these parameters are under control, the resulting leaching rate curve poses another problem. In most cases, the leaching rate shows an exponential decline with a high flux in the beginning and a progressively slower flux as time passes.
This factsheet mainly focuses on the latter and to address this problem, several questions were raised for this report:

1. Which part of the leaching rate curve is relevant for risk assessment?
   Is it the worst-case approach using the flux from the first couple of days or the 'equilibrium' flux (which does not really exist) at the end of the leaching study?
2. How does one calculate the average flux over a certain period?
   Should one take the average of the measured fluxes, use the measured cumulative leaching (in $mg.m^{-2}$ wood) from the leaching study divided by the number of days or calculate the flux by integration of the cumulative leaching? Is the flux a necessary parameter or can emissions be calculated using another approach?
3. Can short term results be extrapolated to long-term emissions?
   The standard duration for a leaching test is 64 days, but shorter and longer durations can also occur. Are these results in agreement?
4. Which model should be used to describe the leaching curve?
   For integration of the curve (see question 2 above) or extrapolation to other time spans (see question 3 above) the results have to be fitted to a leaching curve. Several models are suggested, but their advantages and disadvantages are unknown.

4.2 Proposed models in literature

Several models that are able to describe the leaching flux of wood preservatives can be found in literature. In this paragraph, some of these models will be described and their advantages and disadvantages will be addressed.

4.2.1 Models based on fluxes

The models in this paragraph are all based on fluxes (=leaching rates). Models based on cumulative leaching will be addressed in paragraph 4.2.2.

1. Berbee, R.P.M. (1989):
Using regression analysis on leaching data for CCA wood preservatives, Berbee calculated the hyperbole that fitted the data best. This resulted in the following model:

$$\ln F(t) = a \times \ln(t) + b \quad (1)$$

which can be rewritten to:

$$F(t) = e^{b} \times t^{a} \quad (2)$$

in which:

F(t)	= flux (or leaching rate)	[mg.m^{-2} wood.day^{-1}]
e	= natural logarithm	[-]
t	= time	[days]
b	= constant that is estimated by the model	[-]
a	= constant that is estimated by the model	[-]

Based on the above, the model has to estimate two parameters (a and b). Although the model fitted the data from Berbee well, some critical remarks can be made about this model:

- It is an empirical model that does not give a conceptual description of the leaching process which is assumed to be a diffusion based process (see paragraph 4.2.2).
- Extrapolation to longer time periods is dangerous/difficult, because the cumulative leaching (in mg.m^{-2} wood) will approach infinity when the time span (t) goes to infinity. This would result in a leached amount of wood preservative that is far greater than the amount that was actually put into/onto the wood in the first place.
- Because of the (natural) logarithms, an experimentally determined flux of 0 mg.m^{-2} wood.day^{-1} cannot be entered into equation 1. Berbee resolved this problem by entering the detection limit of the analytical method, but this results in an overestimation of the flux.

2. Empirical equations can also be found in literature, for instance by Brooks (1995, 1997a, 1997b, 1998 and 2003) as summarized by Dickey (2003).

Brooks conducted literature reviews of leaching studies in aquatic environments and derived empirical equations to estimate the leaching rates for some wood treatment systems as a function of time and other conditions.

CCA (chromated copper arsenate):

- Flux = exp(-0.048*t)*0.51exp(0.02*S)*(0.65*(0.8462+ln (0.71*R))) (3)

ACZA (ammoniacal copper zinc arsenate):

- Flux = 1908.6*exp(-0.429*t-0.383*pH) in freshwater (4)
- Flux = 32.5*exp(-1.114*t) in saltwater (5)

ACQ (ammoniacal copper quat):

- Flux = 265.14*exp(-0.924*t-0.239*pH) in freshwater (6)
- Flux = 4.25*exp(-0.0175*t) in saltwater (7)

CuN (copper naphthenate):

- Log10 (Flux) = 0.093 + 1.269*exp(-0.1375*t) in freshwater (8)

in which:

Flux	= loss rate per unit of surface area	[µg.cm^{-2} wood.d^{-1}]
t	= time after installation	[days]
S	= salinity	[ppt]
R	= preservative retention	[kg.m^{-3}]

Advantages:
- Equations fit data well.
- The equations are based on data sets that span longer time frames (typically a month to several months).
- The effect of the pH on the leaching rate is incorporated into some of the equations.

Disadvantages:
- It is an empirical model that does not give a conceptual description of the leaching process which is assumed to be a diffusion based process (see paragraph 4.2.2).
- Calculations assume stagnant water with no currents of flow, which is not representing the situation of treated wood in-service.
- Formulas are applicable/restricted to a limited amount of wood preservatives.
- Models are based on specific conditions in the original experiments (e.g. CCA and ACQ on treated yellow pine and ACZA and CuN on Douglas fir). Therefore, comparing leaching rates between different preservative treatments might be difficult and calculated leaching rates might not be valid for the specific conditions under which a wood preservative is meant to be put onto the market by the applicant.

3. Schoknecht et al. (2005) used an equation that was taken from Panelli (2001a, 2001b), who observed that all measurement points for copper, chromium and arsenic from leaching tests with CCA preservative were usually distributed regularly and that a polynomial regression of second order fitted the data well. This resulted in the following equation:

$$\log_{10} FLUX(t) = a + b \cdot \log_{10}(t) + c \cdot \log_{10}(t)^2 \qquad (9)$$

in which:

FLUX(t)	= Leaching rate	[$mg.m^{-2}$ $wood.d^{-1}$]
a	= parameter estimated by the model	[-]
b	= parameter estimated by the model	[-]
c	= parameter estimated by the model	[-]
t	= time	[d]

For this equation, three parameters have to be fitted by the model (a, b and c). As for all models, this model also has some advantages and disadvantages.

Advantages:
- The model fits the data well.
- FLUX can be determined for every time t (except t=0).

Disadvantages:
- FLUX cannot be determined for t=0 and subsequently the total quantity leached from t=0 cannot be calculated. Schoknecht et al. added the quantity that was measured during the first day of the experiment to the calculated cumulative losses to estimate the total quantity leached from t=0.
- It is an empirical model that does not give a conceptual description of the leaching process which is assumed to be a diffusion based process (see paragraph 4.2.2).
- Extrapolation to longer time periods is dangerous/difficult, because the cumulative leaching (in $mg.m^2$ wood) will approach infinity when the period (t) goes to infinity. This would result in a leached amount of wood preservative that is far greater than the amount that was actually put into/onto the wood in the first place.

The models that are described above are all used to fit a curve through calculated flux data. These fits subsequently result in an average flux (in mg.m^{-2} wood.d^{-1}) for a certain period. The fluxes that are calculated by the aforementioned models are based on experimental leaching data, which are mostly cumulative leaching data (in mg or mg.m^{-2} wood).
For the OECD Emission Scenario Documents (ESDs), the parameter of interest is the cumulative leaching (in mg or mg.m^{-2}) within a certain period. Presently, the cumulative leaching is calculated according to the OECD ESD as the product of the area of wood that is exposed (in m^2), the emission flux (in mg.m^{-2} wood.d^{-1}) and time (in days). If one can estimate the cumulative leaching within a certain period without using the transition of the leaching data to fluxes, this would make the emission estimations much simpler, mainly because there are several 'difficulties' that accompany the use of a flux as parameter:

1. In a leaching test, a piece of wood is submerged for a certain period in a vessel containing water. Next, the water is removed and the concentration of the leached wood preservative in the water is measured. This procedure is repeated several times (e.g. 4, 11, 18, 25, 32, 39, 46, 53, 60 and 64 days). This procedure results in a cumulative leaching (in mg or mg.m^{-2}) curve that can be considered as a continuous process. Fluxes (in mg.m^{-2} wood/day) are usually calculated per time step (e.g. the average flux for day 0-4 is calculated by dividing the amount of wood preservative leached from day 0-4 by the number of days). This method is applied to all time steps and subsequently a curve is fitted through all the calculated fluxes. The way the flux is determined, results in a discontinuous process, although leaching in itself is a continuous process.
2. It is an extra step in the risk assessment process. Since the experiments result in a cumulative leaching curve (in mg or mg.m^{-2}) and the parameter of interest for the emission scenarios is the cumulative leaching within a certain period, it would be preferable to extrapolate the cumulative leaching directly from the experiment to the field situation without the use of a flux.
3. It is unclear what part of the leaching rate curve is relevant for risk assessment. Is it the worst-case approach using the flux from the first couple of days or the equilibrium flux at the end of the leaching study?
4. It is unclear how the average flux over a certain period has to be calculated. Should one take the average of the measured fluxes, use the measured cumulative leaching (in mg.m^{-2} wood) from the leaching study divided by the number of days or calculate the flux by integration of the cumulative leaching?
5. It is unclear whether or not short term results can be extrapolated to long-term emissions.

In the following paragraph, several models that describe the cumulative leaching of wood preservatives from wood are summarized.

4.2.2 Models based on cumulative leaching

Schnute (1981):
Schnute proposed the following growth model, which includes numerous historical models as special cases. It has to be noted that this model is not specifically designed for leaching of wood preservatives, but was initially based on the growth of vertebrates and invertebrates:

$$Y(t_2) = Y(t_1) \times \left(\frac{\left(1 - e^{-at_2}\right)}{\left(1 - e^{-at_1}\right)} \right)^{\frac{1}{b}} \qquad (10)$$

Based on this model, Homan and Oosten (1999) proposed the following model to determine the (long term predictions of) cumulative leaching of wood preservatives from treated timber:

$$Q(t) = Q_{64} \times \left(\frac{\left(1 - e^{-a \cdot t}\right)}{\left(1 - e^{-a \cdot 64}\right)} \right)^{\frac{1}{b}} \qquad (11)$$

in which:

Q(t)	= cumulative amount of wood preservative leached after time *t*	[mg.m^{-2}]
Q_{64}	= experimentally determined amount of wood preservative leached after 64 days	[mg.m^{-2}]
e	= natural logarithm	[-]
a	= parameter estimated by the model	[1/days]
t	= time	[days]
b	= parameter estimated by the model	[-]

This model fits the experimental leaching data well and has a couple of advantages when compared to the Berbee model, but some critical remarks can also be made:

Advantages:

- The model itself 'chooses' the shape of the curve, depending on the value of parameters a and b, instead of the shape of the curve being based on the interpretation/view of the interpreter of the data. This is shown in Figure 4-1 and 4-2.
- According to SHR (1995), it is important to find out if the data make the presence of an asymptote likely or not, especially for the extrapolation of the experimental data to long term predictions. SHR proved that an asymptote in the cumulative leaching function is (almost) always present for fixating wood preservatives. However, this is only true when the results are plotted on a linear scale, not in reality.

Disadvantages:

- The model was developed as a general growth model for several species and is therefore not considered to provide an adequate description of the leaching process (diffusion, see this paragraph).
- It seems somewhat arbitrarily to use the experimental Q_{64}-value (usually the final experimental data point) to define the height of the model. If another value (e.g. 50 days) would be chosen, this might influence the height of the model. In a document by Paneli (2001), she compared equation 9 to the modified Schnute model as described above. Results from the modified Schnute model for two datasets containing three substances each, showed that the cumulative amount leached after 1 and even 10 years was in all cases similar to the amount leached after 64 days, which seems to be an unlikely scenario. It cannot be explained why using Q_{64} would be more valid than any other Q_t.

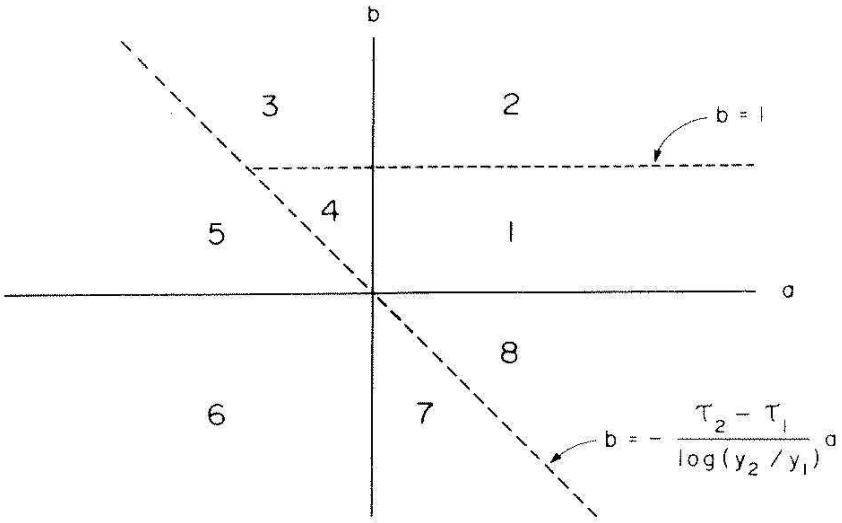


Figure 4-1: Set of eight regions in the a,b plane defined by four broken lines.
The solid lines correspond to the a-axis and b-axis. The broken line parallel to the a-axis is defined by b = 1. The diagonal broken line with negative slope is defined by the equation that is shown in the figure. Figure taken from Schnute (1981).

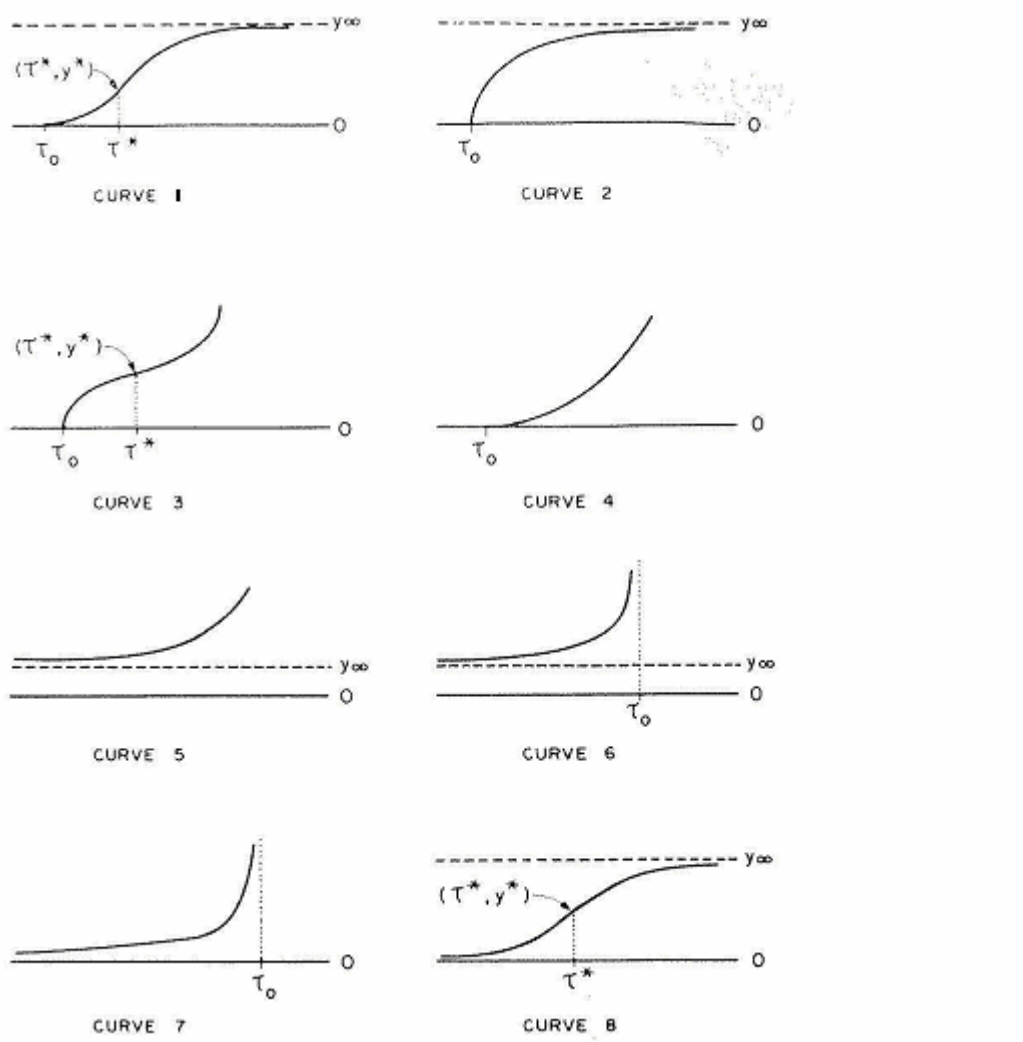


Figure 4-2. Set of eight characteristic growth curves appropriate to a parameter pair (a,b) in each of the eight regions shown in Figure 4-1. The curves all represent plots of size Y on the vertical scale against age t on the horizontal scale. Figure taken from Schnute (1981).

Based on the above, it seems preferable to find another (simplified) approach that fits the experimental data well and also provides some kind of logical description of the leaching process.
In terms of release behaviour, two types of construction products can be distinguished: monolithic and granular materials (Van der Sloot and Dijkstra, 2004). Monolithic products (e.g. concrete, blocks, bricks and wood) often show diffusion controlled release, whereas granular materials (e.g. sand, sinters, gravel, steel slag) usually show a percolation dominated release. For wood, the release process seems to be diffusion driven, which in turn can depend on environmental parameters (e.g. the (environmental) pH in case of speciation of metals). Since the emission of wood preservatives from the wood is a concentration dependent process with decreasing emission at decreasing concentrations in wood, this process can be compared to the degradation of substances in water or soil. Therefore, a model describing first order degradation might also be valid to describe leaching data:

$$C(t)_{wood} = C(0)_{wood} \times e^{-kt} + plateau \qquad (12)$$

in which:

$C(t)_{wood}$	= amount of wood preservative in wood on time *t*	[e.g. mg.m^{-3}]
$C(0)_{wood}$	= amount of wood preservative in wood on t=0	[e.g. mg.m^{-3}]
k	= emission constant, estimated by the model	[1/time]
t	= time	[e.g. days]
plateau	= remaining amount of wood preservative that will not leach	[e.g. mg.m^{-3}]

This model describes a continuous decrease of the wood preservative concentration in wood in time down to a certain plateau (concentration).
When looking from the cumulative leaching perspective, this can be re-written to:

$$C(t)_{leached} = C_{leached,\max} \times (1 - e^{-kt}) \qquad (13)$$

in which:

$C(t)_{leached}$	= cumulative amount leached on time *t*	[mg.m^{-2} wood]
$C_{leached, max}$	= maximum amount of wood preservative leached, estimated by the model	[mg.m^{-2} wood]
k_1	= emission constant, estimated by the model	[1/time]
t	= time	[e.g. days]

In Paneli (2001a), the values for parameter a and b from the modified Schnute model were given for six leaching curves. Five out of six leaching curves resulted in a curve 2 shape (see Figure 4-2) and one resulted in a curve 1 type leaching curve, because parameter b was determined to be 0.892 (<1). The shape of curve 2 appears to be similar to the shape of a curve as defined by equation 13.

When wood preservatives are applied to wood, they are assumed to be available in two pools:

1. One pool is present in the applied water phase and the boundaries of the wood;
2. Another pool is impregnated deeper into the wood.

Both pools will leach from the wood simultaneously, although the contribution of each pool to the total amount leached changes in time. At first, relatively fast diffusion from the applied water phase and the boundaries of the wood will be most important, which will eventually be dominated by a slower diffusion process of wood preservative that is impregnated deeper into the wood. Leaching in this phase can also be slower due to a lower wood preservative concentration in the treated timber and because of changing wood properties (e.g. weathering). The first phase will be relatively important for surfacial applications of wood preservatives, such as automated dipping or spraying, whereas the second phase can play an important role by vacuum pressure, because the wood preservative is ‘pressed’ deeper into the timber.
The description of this leaching process is more or less comparable to the bi-exponential model in the FOCUS degradation kinetics document (draft version 2004, p. 50), which is a double first order in parallel model. Taking a second phase into account, the cumulative leaching equation can be re-written to:

$$C(t)_{leached} = C_{leached,\max 1} \times (1 - e^{-k_1 t}) + C_{leached,\max 2} \times (1 - e^{-k_2 t}) \quad (14)$$

in which:

$C(t)_{leached}$	= cumulative amount leached on time *t*	[mg.m^{-2}]
$C_{leached, max1}$	= maximum amount of wood preservative leached in the first phase, estimated by the model	[mg.m^{-2}]
$C_{leached, max2}$	= maximum amount of wood preservative leached in the second phase, estimated by the model	[mg.m^{-2}]
k_1	= emission constant during the first phase, estimated by the model	[1/time]
k_2	= emission constant during the second phase, estimated by the model	[1/time]
t	= time	[e.g. days]

By using one of the equations above, an emission constant can be calculated and by putting this constant into the equation, the cumulative leaching within a certain period can be calculated. This way, the results from the leaching experiments can also be extrapolated to a long-term scale.

Another approach can be found in the Dutch Building Materials Decree, in which the following equation is used to calculate the emission of compounds from construction materials (e.g. concrete) during 100 years (from Aalbers et al., 1996):

$$E(t) = E_{avail} \times d_c \times \sqrt{\frac{D_e}{\pi}} \times \sqrt{t} \quad (15)$$

in which:

E(t)	= emission during t units of time	[mg.m^{-2}]
E_{avail}	= the availability of compounds for leaching	[mg.kg^{-1}]
d_c	= dry density of the construction material	[kg.m^{-3}]
D_e	= coefficient of effective diffusion	[m^2.s^{-1}]
t	= time	[s]

In this factsheet, D_e for wood is considered constant, based on the assumption that within the 64 days of the leaching experiment, the wood structure will not change. Therefore, equation 15 can be re-written as:

$$E(t) = c \times \sqrt{t} \quad (16)$$

in which:

E(t)	= emission during t units of time	[mg.m^{-2}]
c	= constant (comprised of $E_{avail} \times d_c \times \sqrt{\frac{D_e}{\pi}}$)	[mg.m^{-2}.t$^{-1/2}$]
t	= time	[s]

It has to be noted that by using this model for long term extrapolations, emissions can be bigger than the amount of wood preservative that was put into the wood in the first place, if *t* gets big enough.

4.2.3 Comparison of cumulative leaching models against a flux model

In a document to the OECD Expert Group for the ESD for Wood preservatives (dated 7 August 2001), Paneli compared her equation (see equation 9) to the modified Schnute equation (see equation 11) and the Berbee equation (see equation 2).

The comparison was based on a dataset on CCA-salts from SHR (1995). Paneli concluded that her model often offered better accuracy in re-calculating the experimental results than the modified Schnute model and the Berbee model.
In this paragraph, the Paneli model is compared to the cumulative leaching models (equations 13 and 14). It has to be noted that the Paneli model was fitted on flux data and not on cumulative leaching data. Therefore, it is difficult to compare this model to the cumulative leaching models.
The correlation coefficients (r^2) of the Paneli model, the first order model and the bi-phasic model are shown in Table 4-1. The first order model was fitted using Graphpad 4.0, the bi-phasic model using Berkeley Madonna 8.0.1. The correlation coefficients from Paneli were taken from Panelli (2001a). Although a comparison of correlation coefficients might not be a waterproof method to determine if one model results in a better fit than the other, it can be considered to be an indication. It has to be noted that the presented results are currently only valid for the cases at hand and examining additional data sets is deemed necessary to find out if this can be considered to be a trend.

*Table 4-1: Comparison of correlation coefficients. If the r^2 for the cumulative leaching models is lower than for the Paneli model, the corresponding value is indicated in **bold**.*

Dataset/substance	Paneli	first order	bi-phasic
Data set 1 – substance 1	0.994	0.994	1.000
Data set 1 – substance 2	0.992	**0.991**	1.000
Data set 1 – substance 3	0.983	1.000	1.000
Data set 2 – substance 1	0.998	**0.992**	1.000
Data set 2 – substance 2	0.997	**0.984**	1.000
Data set 2 – substance 3	0.990	0.999	1.000
Data set 3 – substance 1	0.998	**0.987**	0.999
Data set 3 – substance 2	0.997	**0.986**	1.000
Data set 3 – substance 3	0.986	0.999	1.000
Data set 4 – substance 1	0.995	**0.989**	0.999
Data set 4 – substance 2	0.985	0.997	1.000
Data set 4 – substance 3	0.994	0.999	0.999
Data set 5 – substance 1	0.986	0.998	1.000
Data set 5 – substance 2	0.996	**0.995**	1.000
Data set 5 – substance 3	0.997	**0.995**	1.000
Data set 6 – substance 1	0.995	**0.994**	0.999
Data set 6 – substance 2	0.995	0.997	1.000
Data set 6– substance 3	0.991	0.997	1.000
Data set 7 – substance 1	0.991	0.996	1.000
Data set 7 – substance 2	0.992	0.994	1.000
Data set 7 – substance 3	0.980	0.996	1.000
Data set 8 – substance 1	0.983	0.998	0.999
Data set 8 – substance 2	0.978	0.999	1.000
Data set 8 – substance 3	0.992	0.997	0.999
Data set 9 – substance 1	0.996	**0.979**	1.000
Data set 9 – substance 2	0.997	**0.994**	1.000
Data set 9 – substance 3	0.982	0.999	1.000
Data set 10 – substance 1	0.991	0.998	1.000
Data set 10 – substance 2	0.999	0.999	0.999
Data set 10 – substance 3	0.992	0.997	1.000
Data set 11 – substance 1	0.991	0.997	0.999
Data set 11 – substance 2	0.996	**0.991**	0.999
Data set 11 – substance 3	0.984	0.993	1.000
Data set 12 – substance 1	0.994	**0.973**	0.999
Data set 12 – substance 2	0.979	0.993	1.000
Data set 12 – substance 3	0.992	0.997	1.000

From all the fits, 13 out of 36 fits showed a (slightly) higher correlation coefficient if the Paneli method was used, when compared to the first order model. When the bi-phasic model was used, all correlation coefficients were higher than the ones for the Paneli model, although it has to be noted that one additional parameter was introduced in the bi-phasic model (4 parameters) when compared to the Paneli model (3 parameters).

Based on these results, it can be assumed that the cumulative leaching models do not fit significantly worse or might even fit better than the flux model that was suggested by Paneli.

4.2.4 Comparison of the cumulative leaching models

In order to find out if one of the cumulative leaching models proves to be significantly better than the others, equations 13, 14 and 16 were applied to the same SHR dataset as in paragraph 1.1.3, which contains 36 cumulative leaching curves. The modified Schnute model was not taken into account, because this model lacks a description of the leaching process. The dataset had the following characteristics:

1. Substances: CCA-salt, individual curves for copper, chromium and arsenic;
2. Methods: according to NEN 7345 using different wood species (pinewood and firwood), pH (5.5, 7, 8.8), temperature (20, 8 ºC) and different fixation; methods (using steam and by putting the treated wood in a climate chamber);
3. Laboratory: SHR, the Netherlands.

The first order model was fitted using Graphpad Prism 4. Because Graphpad had difficulties with fitting the bi-phasic model to the data, the bi-phasic and √t models were fitted using Berkely-Madonna 8.0.1.
A chi-squared test was performed using a worksheet Version 1.0 (2004) for Microsoft Excel 2002 from the FOCUS degradation kinetics workgroup. The %-error for each curve is shown in Table 4-2:

Table 4-2: %-error from chi-squared tests for three different cumulative leaching models

Dataset/substance	**first order**	**bi-phasic**	**√t**
Data set 1 – substance 1	6.8	0.5	4.2
Data set 1 – substance 2	7.5	0.5	4.2
Data set 1 – substance 3	1.8	0.2	11.7
Data set 2 – substance 1	7.7	0.7	2.6
Data set 2 – substance 2	7.3	0.6	3.8
Data set 2 – substance 3	4.1	0.8	14.0
Data set 3 – substance 1	8.6	2.2	7.8
Data set 3 – substance 2	8.2	1.7	3.8
Data set 3 – substance 3	3.7	1.3	8.3
Data set 4 – substance 1	8.8	2.1	2.7
Data set 4 – substance 2	4.5	0.9	6.3
Data set 4 – substance 3	4.1	3.2	7.9
Data set 5 – substance 1	4.9	1.5	5.3
Data set 5 – substance 2	6.3	0.5	4.6
Data set 5 – substance 3	5.9	1.8	8.2
Data set 6 – substance 1	6.5	2.2	4.2
Data set 6 – substance 2	4.9	0.3	6.1
Data set 6– substance 3	4.8	1.7	12.2
Data set 7 – substance 1	6.3	1.7	6.3
Data set 7 – substance 2	6.3	0.8	4.7
Data set 7 – substance 3	5.9	2.0	11.5
Data set 8 – substance 1	4.9	2.0	5.5
Data set 8 – substance 2	2.5	0.9	9.0
Data set 8 – substance 3	5.2	2.1	8.8
Data set 9 – substance 1	9.4	1.0	16.7
Data set 9 – substance 2	4.9	0.9	25.4
Data set 9 – substance 3	3.0	1.6	18.6
Data set 10 – substance 1	11.9	1.1	17.8
Data set 10 – substance 2	6.1	1.7	23.3
Data set 10 – substance 3	5.1	1.8	20.0
Data set 11 – substance 1	4.9	2.2	17.6
Data set 11 – substance 2	6.5	2.1	20.3
Data set 11 – substance 3	6.9	0.5	9.9
Data set 12 – substance 1	10.7	1.4	12.0
Data set 12 – substance 2	6.4	1.2	15.0
Data set 12 – substance 3	6.6	0.7	17.0

Based on the results from the chi-squared tests, the models fitted the data best in the following (descending) order: bi-phasic model, first order model, √t model.
The bi-phasic model also appears to be the model with the best description of the leaching process of wood preservatives from treated timber (see paragraph 4.2.2).

In order to determine whether the bi-phasic model also fitted well to other data sets, another dataset was used. This dataset was taken from Schoknecht et al. (2005) and contained several substances and experimental methods (performed by several laboratories):

1. Substances: CCB-salt, propiconazole, Cu-triazole, individual curves for all individual components;
2. Methods: using OECD or EN113 wood sizes, with and without stirring, different pH (5 and 7), different emersion scenarios (3 immersions of 1 min per day, 2 immersions of 60 min per day or permanently immersed);
3. Laboratories: INIA Madrid, Spain; ICT Prague, Chec Republic; University Perugia, Italy; BAM Berlin, Germany; MPA Eberswalde; Germany; CTBA Bordeaux, France; DTI Taastrup, Denmark; ECN Petten, the Netherlands; HFA Wien, Austria.

This amounted to a total of 58 leaching curves. Because of the variety in experimental methods, laboratories and wood preservatives, this data set is considered to be suitable to test the applicability of the bi-phasic model. Results are shown in Table 4-3.

Table 4-3: correlation coefficient and %-error from chi-squared tests for the bi-phasic cumulative leaching model

wood preservative	**substance**	**r^2-values**	**%-error from chi-squared test**
CCB	boron 1	1.0000	0.2
	boron 2	0.9977	1.7
	boron 3	0.9982	1.5
CCB	chromium 1	0.9989	1.3
	chromium 2	0.9990	1.2
	chromium 3	0.9997	0.7
CCB	copper 1	0.9997	0.6
	copper 2	1.0000	0.1
	copper 3	1.0000	0.2
Propiconazole	propiconazole 1	0.9997	0.8
	propiconazole 2	0.9993	1.3
	propiconazole 3	0.9999	0.5
	propiconazole 4	0.9964	3.2
	propiconazole 5	0.9998	0.8
Cu-triazole	copper 1	0.9998	0.7
	copper 2	0.9986	1.8
Cu-triazole	triazole 1	0.9995	1.0
	triazole 2	0.9993	1.3
Cu-triazole	copper 1	0.9968	2.5
	copper 2	0.9979	2.0
	copper 3	0.9969	2.5
Cu-triazole	triazole 1	0.9960	3.5
	triazole 2	0.9957	3.3
	triazole 3	0.9957	3.5
Cu-triazole	copper 1	0.9987	1.2
	copper 2	0.9982	1.5
	copper 3	0.9973	1.8
	copper 4	1.0000	0.2
Cu-triazole	triazole 1	0.9980	2.1
	triazole 2	1.0000	0.2
	triazole 3	0.9966	2.7
	triazole 4	0.9999	0.4
Cu-triazole	copper 1	1.0000	0.2
	copper 2	0.9998	0.7
	copper 3	0.9998	0.7

	copper 4	0.9998	0.6
	copper 5	0.9999	0.4
	copper 6	0.9999	0.4
	copper 7	0.9998	0.6
	copper 8	1.0000	0.1
	copper 9	0.9996	1.0
	copper 10	0.9982	1.9
	copper 11	0.9996	0.9
	copper 12	1.0000	0.3
	copper 13	0.9995	1.0
Cu-triazole	triazole 1	0.9995	1.1
	triazole 2	0.9985	2.4
	triazole 3	0.9997	0.9
	triazole 4	0.9996	1.0
	triazole 5	0.9994	1.4
	triazole 6	0.9994	1.2
	triazole 7	0.9988	1.8
	triazole 8	0.9998	0.7
	triazole 9	0.9972	2.9
	triazole 10	0.9990	1.6
	triazole 11	0.9995	1.1
	triazole 12	0.9890	6.2
	triazole 13	0.9984	2.0

The bi-phasic model fitted the dataset really well and because this model also appears to provide a more realistic description of the leaching process of wood preservatives from treated timber when compared to the other diffusion based models, it can be concluded that the bi-phasic model can be used for PEC estimations.

4.3 How to estimate a PEC based on the bi-phasic model?

The bi-phasic model was described as:

$$C(t)_{leached} = C_{leached,\max 1} \times (1 - e^{-k_1 t}) + C_{leached,\max 2} \times (1 - e^{-k_2 t})$$

By fitting the model to the data from an experimental leaching study, $C_{leached,max\ 1}$, $C_{leached,max\ 2}$, k_1 and k_2 are estimated by the model, as shown in Figure 4-3.

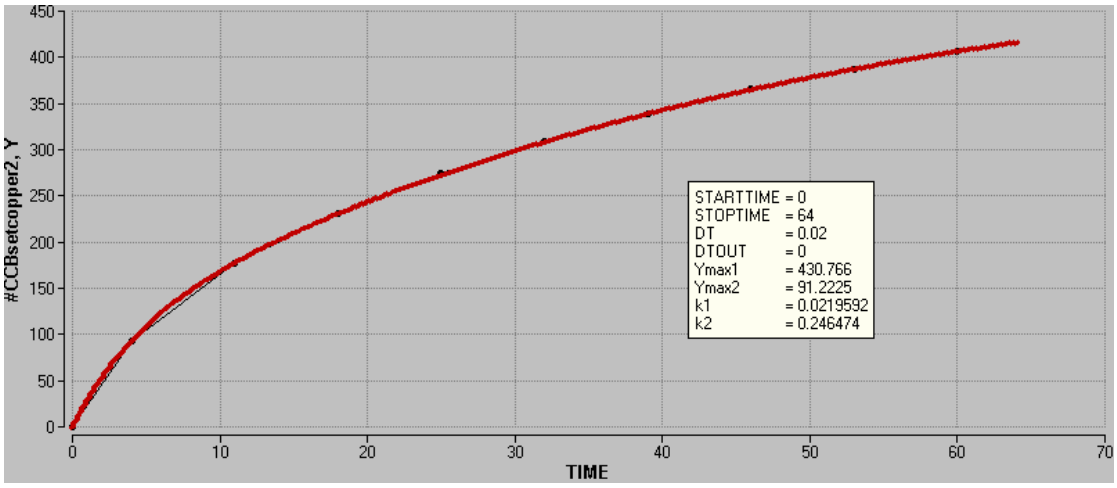


Figure 4-3: Example of a fit of the bi-phasic model to experimental data.

This example yields the following equation:

$$C(t)_{leached} = 430.766 \times (1 - e^{-0.0219592 \cdot t}) + 91.2225 \times (1 - e^{-0.246474 \cdot t})$$

By choosing a certain time period (t), the cumulative leaching for that period can easily be calculated.

4.4 Some critical remarks on the current test strategy

Even though the first order and bi-phasic equations fit the analysed data really well, one can wonder whether or not they also provide realistic leaching rates that can be used as input parameters for various emission scenarios. This (greatly) depends on the specific conditions and set-up of the leaching experiment that is used to generate the data and on the specific conditions (environment, wood species) the product will be used for. Several parameters that may differ from test to test can influence the leaching rate of wood preservatives, such as:

1. The chemical state of the impregnated substances which in turn can be influenced by the solid organic carbon content of the wood (serving as binding sites for the wood preservative), dissolved organic carbon (DOC) and the pH of the water;
2. The extent of impregnation, which influences the availability of the wood preservative. For instance, wood preservatives applied by vacuum pressure are expected to infiltrate the wood much deeper than wood preservative applied by brushing or dipping, which is likely to result in different leaching behaviour;
3. The porosity and pore structure of the wood differ between wood species and can influence the leaching behaviour of the wood preservative.

Transport of wood preservatives from treated wood is mainly diffusion dominated, but the chemical interaction between wood preservative and the wood and its environment are also important. Transport of free unbound wood preservative or transport of wood preservative adsorbed to DOC is not similar and complexation can also influence its activity.

Test methods that try to simulate reality (e.g. OECD Draft Proposal Guideline 1: immersion tests) might seem appropriate, but their results cannot be linked directly to reality, as was shown in the report of Schoknecht et al. (2005). In this report, guideline 1 (and 2) experiments were compared with field tests and results showed that the relation between outdoor and laboratory experiments was different for each substance and also depended on the duration of the dipping events (eg. 1 or 60 minutes). As stated by Schoknecht: ‘This means that there was no time schedule for immersion days that generally represents emissions that are caused by certain weather conditions. It is impossible to define the schedule of the immersion events only on the basis of water uptake. Additional environmental factors influence the emission rates under natural weathering. The emission is also determined by the active ingredients themselves, probably related to properties like water solubility and the formulation of the wood preservative and environmental factors that are not considered in the laboratory experiments’.

In order to gain insight in the most important parameters, it would be very useful to determine the influence of several parameters such as pH, solid organic carbon (i.e. wood with binding sites for metals and organic wood preservatives), dissolved organic carbon (DOC) and electric conductivity of the wood.
Methods such as the tank test (OECD Draft Proposal Guideline 2), in which the treated wood is immersed permanently, can relatively easily be adapted to provide information on these parameters, for instance by combining it with pH stat.

With the parameters estimated from a tank test like NEN 7345 or EN 1250 or OECD guideline 2 it is possible to use a more advanced diffusion model that takes the dimensions of the test specimen of the dimensions in the field application into account and thus avoid over-prediction of release by applying a one-dimensional diffusion model.
It should be noted that this type of modeling is only needed to properly describe a source term for deriving criteria. Once the criteria have been set more simple approaches are possible for the daily practice.
Such a more sophisticated approach is applied by van der Sloot et al. This approach implies using a chemical speciation model in combination with release and transport (LeachXS with embedded Orchestra). In this approach, the chemical processes controlling release of inorganic and organic compounds can be addressed and subsequently taken into account in describing release from a variety of situations. The advantage of that approach is that material intrinsic properties are derived that can subsequently be used in modelling different exposure scenarios (permanently wet, wet-dry cycles, etc) as well as transport of released substances into soil using a mechanistic approach rather than a Kd concept.
Steps are: speciation modeling using a pH dependence test on size reduced material Identifying release rates in terms of diffusion coefficients**,** initial surface flush and/or solubility controlled release from the surface, by using a tank test approach like NEN 7345 or EN 1250 or OECD guideline 2). Figure 4-4 shows an example of speciation/partitioning modeling of Cu from CCA treated wood.

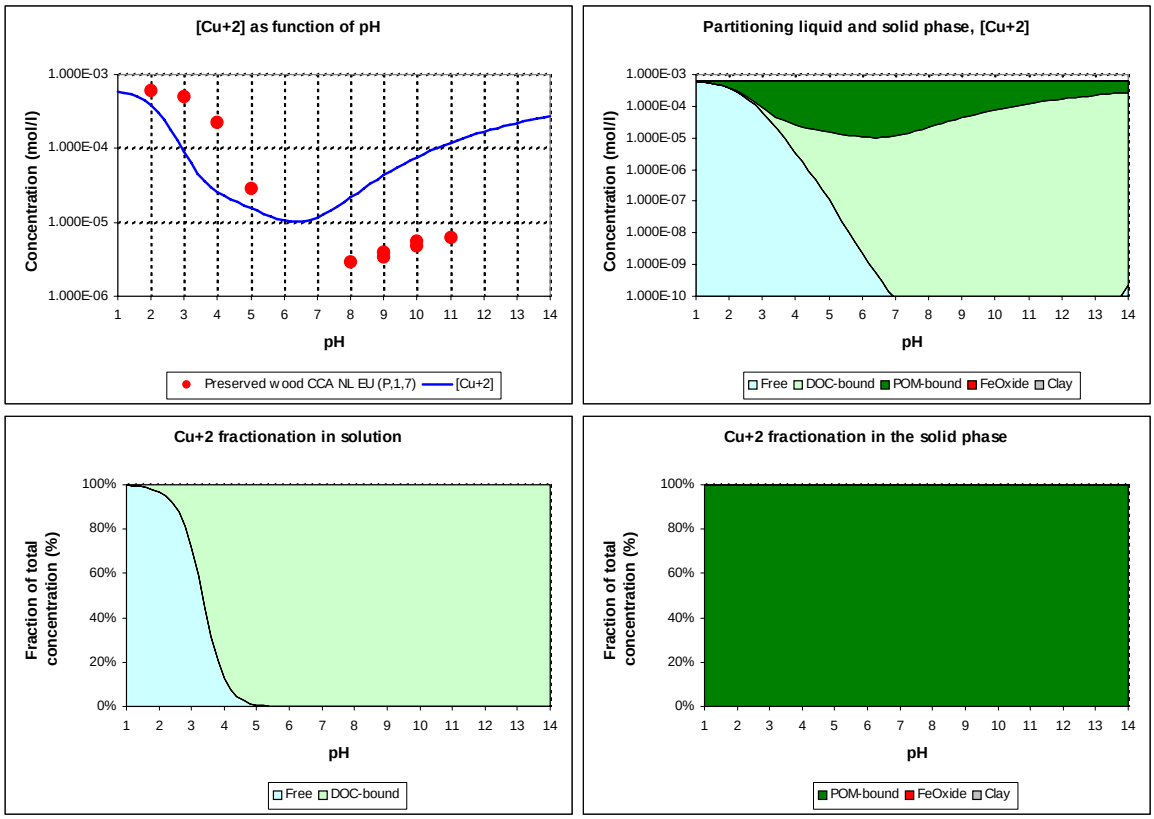


Figure 4-4: Partitioning of Cu from CCA treated wood using LeachXS-Orchestra (provided by H. van der Sloot).

Very preliminary studies have revealed a number of interesting aspects of release of biocides from treated wood (presentation Berlin 2004 and EU report BAM). Using a scenario developed for other materials has shown already a quite useful match between field observations and

model predictions as shown in Figure 4-5. Release of Cu (and possibly other elements) in DOC associated form is of relevance. This may also be the case for some organic biocides (not studied in detail).

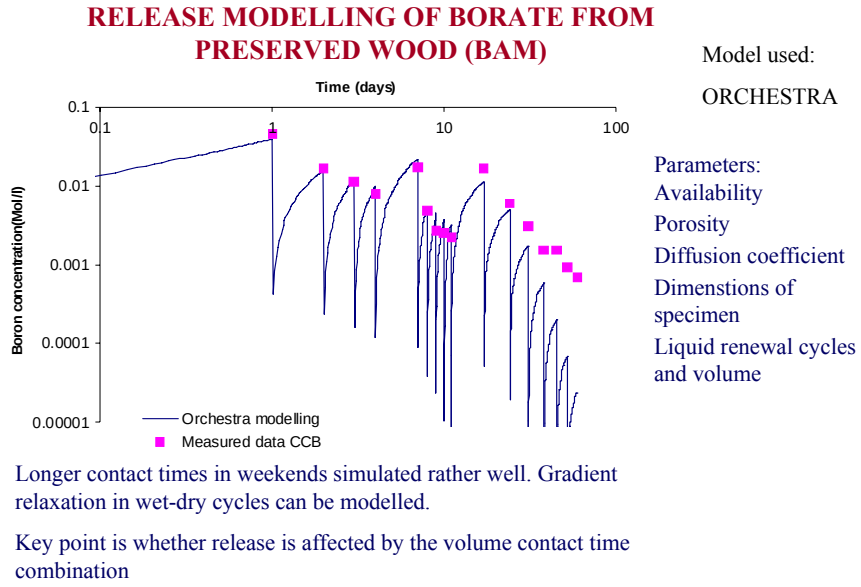


Figure 4-5: Prediction of release for field data obtained by BAM using diffusion coefficients derived from tank test data (provided by H. van der Sloot).

4.5 Conclusions

In this factsheet, several models are described which can be used to fit to experimental wood preservative leaching data. Each model has its own advantages and disadvantages, but based on the two datasets that are described in this factsheet, the use of a bi-phasic model seems preferable because:

1. the bi-phasic model shows the best fit of the models that were tested;
2. the bi-phasic model is believed to adequately describe the leaching process of wood preservatives from treated timber (diffusion);
3. the bi-phasic model results in leaching rate constants that are constant in time, which is in contrast to a flux, which decreases in time;
4. the results of the bi-phasic model can easily be used to calculate the cumulative leaching for any given time period or time interval. This is in contrast to a flux-based model, for which it is unclear which flux has to be taken to calculate the cumulative leaching within a certain time span;
5. the bi-phasic model has an asymptote which assures that the estimated cumulative leaching will never exceed the amount of wood preservative that was put into/onto the timber in the first place;
6. extrapolations to long-term periods that exceed the experimental duration (mostly 64 days) can easily be made by filling in the time of interest into the equation;
7. Programs like PEARL contain a module which can use a DT_{50} to estimate the release of a substance in time. Normally, this is done for the ‘release’ of a metabolite from the parent compound. Because the leaching process is described by a first order model, the release from the compound from the wood, can be estimated based on the DT_{50} (leaching half time based on the first order leaching *k*). This way, one can simulate a

continuous emission to the receiving environmental compartment which decreases in time and one can combine this with processes in soil or surface water (such as degradation and sorption) which can take place simultaneously.

If the cumulative leaching is expressed as mg wood preservative.m^{-2} wood, the OECD Excel ESD worksheets can easily be adapted to use cumulative leaching as input parameter.

A downside of the bi-phasic model is that a relatively simple program as Graphpad Prism 4 has difficulties fitting the model to the data and a more sophisticated, less simple program such as Berkeley Madonna 8.0.1. had to be used. If a simple approach would be preferable, then the first order diffusion model can be suggested. This model generally fits the experimental leaching data well, although less accurately than the bi-phasic model (see Table 4-1 and 4-2).

4.6 Recommendations for further research

The following suggestions can be used to investigate whether the bi-phasic model can be applied to all leaching curves:

- Perform a study using the modified Schnute model if all leaching curves have a curve 2 shape, i.e. if parameter a>0 and b>1 (see Figure 4-2).
- Perform leaching studies over a longer period of time (or search for existing data) and determine whether or not long term estimations from the bi-phasic model are in agreement with the experimental data.
- Address issues related to the chemical speciation of inorganic and organic biocides, such as release of Cu as DOC bound species (other impact on organisms than anticipated), conversion of Cr(VI) to Cr(III) and release of Cr III as DOC bound species, release of organic biocides in DOC bound form?
- Evaluate the potential of LeachXS Orchestra for biocide release modeling.
- Also take notice of developments in other EU frameworks, such as the Building Materials framework in which a lot of similar work is taking place.

References

Aalbers ThG, De Wilde PGM, Rood GA, Vermij PHM, Saft RJ, Van de Beek AIM, Broekman MH, Masereeuw P, Kamphuis Ch, Dekker PM, Valentijn EA. 1996. Environmental quality of primary and secondary construction materials in relation to re-use and protection of soil and surface water. RIVM-report no. 771402007. National Institute for Public Health and the Environment, Bilthoven, the Netherlands.

Berbee RPM. 1989. Onderzoek naar uitloging in oppervlaktewater van PAK en koper, chroom, arseen uit geimpregneerd hout. D.B.W./RIZA notanr. 89.049, Lelystad, the Netherlands.

Brooks KM. 1995. Literature review, computer model and assessment of the potential environmental risks associated with creosote treated wood products used in aquatic environments. Prepared for Western Wood Preservers Institute. Vancouver, WA, United States of America, 1995 (revised 1997).

Brooks KM. 1997a. Literature review and assessment of the environmental risks associated with the use of ACZA-treated wood in aquatic environments. Prepared for Western Wood Preservers Institute, Vancouver, WA, United States of America.

Brooks KM. 1995. Literature review and assessment of the environmental risks associated with the use of CCA-treated wood in aquatic environments. Prepared for Western Wood Preservers Institute. Vancouver, WA, United States of America.

Brooks KM. 1998. Literature review, computer model and assessment of the potential environmental risks associated with pentachlorophenol treated wood products used in aquatic environments. Prepared for Western Wood Preservers Institute. Vancouver, WA, United States of America.

Brooks KM. 2003. Literature review, computer model and assessment of the potential environmental risks associated with Copper Naphthenate treated wood products used in aquatic environments. Prepared for Merichem Chemicals and Refinery Services LLC, Tuscaloosa, Alabama, United States of America.

Dickey P. 2003. Guidelines for selecting wood preservatives. For the San Fransisco Department of the Environment, September 2003. Washington Toxics Coalition.

FOCUS (2004). Guidance document on estimating persistence and degradation kinetics from environmental fate studies in EU registration. Report of the FOCUS Work Group on Degradation Kinetics. EC Document Reference Sanco/???/2004 version 1, Draft 29/02/2004.

Homan WJ, Oosten AL. 1999. Statistically stable models for determination of PEC, 30st Annual IRG meeting, Rosenheim.

Paneli M. 2001a. Suggestions for calculation of the long term emissions of preservative component(s) from treated wood, based on experimental leaching data. Working document EXGP (2001)11 of the Expert Group for the development of the OECD ESD for Wood preservatives.

Paneli M. 2001b. Comparison of the performance of Model I and Model II for calculation of the emissions of preservative component(s) from treated wood, based on experimental leaching data. Working document EXGP(2001)15 of the Expert Group for the development of the OECD ESD for Wood preservatives.

Schnute J. 1981. A versatile growth model with statistically stable parameters. Can. J. Fish Aquat. SCi. 38: 1128-1140.

Schoknecht U, Mathies H, Morsing N, Lindegaard B, Van der Sloot HA, Van Zomeren A, Deroubaix G, Legay S, Tadeo JL, Garcia-Valcárcel AI, Gigliotti G, Zadra C, Hajšlová J, Tomaniová M, Wegner R, Bornkessel C, Fürhapper C. 2005. Inter-laboratory evaluation of laboratory test methods to estimate the leaching from treated wood, European Comission Grant Agreement Number 04/375757/C4, April 2005.

SHR. 1995. Conceptrapport uitloging van gewolmaniseerd hout tijdens het gebruik in de waterbouw. Rapport code 94.005. Stichting Hout Research, Wageningen, the Netherlands.

Van der Sloot HA, DIjkstra JJ. 2004. Development of horizontally standardized leaching tests for construction materials: A material base dor release based approach? Identical leaching mechanisms for different materials. ECN-C—04-060, ECN Petten, the Netherlands.