Rijksinstituut voor Volksgezondheid
en Milieu
*Ministerie van Volksgezondheid,
Welzijn en Sport*

Integration of toxicokinetics and toxicodynamics testing essential for risk assessment

RIVM Letter report 055212001/2013
M.B. Heringa et al.

National Institute for Public Health
and the Environment
Ministry of Health, Welfare and Sport

Integration of toxicokinetics and toxicodynamics testing essential for risk assessment

RIVM Letter report 055212001/2013
M.B. Heringa et al.

Colophon

© RIVM 2013
Parts of this publication may be reproduced, provided acknowledgement is given to: National Institute for Public Health and the Environment, along with the title and year of publication.

Minne.B.Heringa, RIVM
Esther F.A. Brandon, RIVM
Jos G. Bessems, RIVM/JRC
Peter M.J. Bos, RIVM

Contact:
Minne Heringa
VSP
Minne.Heringa@rivm.nl

This investigation has been performed by order and for the account of Ministery of Economic Affairs, within the framework of project 5.5.2 'Implementatie van Alternatieven voor Dierproeven in de risicobeoordeling'

This is a publication of:

National Institute for Public Health and the Environment
P.O. Box 1│3720 BA Bilthoven
The Netherlands
www.rivm.nl/en

Publiekssamenvatting

Bij de beoordeling van schadelijke effecten van chemische stoffen is het van belang gegevens toe te voegen over de manier waarop stoffen zich in het lichaam gedragen (de kinetiek). Wordt een stof bijvoorbeeld door de darm opgenomen of direct uitgescheiden? Blijft het in vetcellen zitten? Met deze kennis kunnen de gezondheidsrisico's van een stof nauwkeuriger worden ingeschat. Ook zijn er minder proefdieren nodig, onder andere omdat bepaalde testen dan niet meer nodig zijn. Daarnaast kan worden voorkomen dat proeven opnieuw uitgevoerd moeten worden, wat in het huidige systeem aan de orde is als bijvoorbeeld achteraf de dosering van de te onderzoeken stof te hoog blijkt te zijn geweest of de gebruikte diersoort niet representatief blijkt voor de geteste stof.

Dit blijkt uit een evaluatie van het RIVM naar de voordelen van informatie over kinetiek voor de beoordeling van schadelijke eigenschappen van stoffen. Hierin is ook beschreven hoe de kinetiek meer geïntegreerd kan worden in beoordelingssystemen en op welke manier zo min mogelijk of zelfs geen proefdieren nodig zijn. Dit is van belang omdat er al enkele jaren wordt gewerkt aan een nieuwe strategie voor risicobeoordelingen van chemische stoffen waarvoor zo min mogelijk of geen proefdieren nodig zijn. Het idee daarvan is dat stoffen worden getest in cellen (*in vitro*) in plaats van in hele dieren (*in vivo*). Kennis van kinetiek is dan essentieel om de vertaalslag te kunnen maken van het gedrag van stoffen in cellen naar het hele lichaam.

Momenteel worden kinetische gegevens wel gebruikt bij risicobeoordelingen, maar nog niet (optimaal) voor alle beoordelingskaders (wel voor geneesmiddelen, niet voor industriële stoffen). Deze evaluatie beschrijft de beschikbare (*in vitro*) methoden om informatie over de kinetiek te verkrijgen en op welke wijze deze methoden zouden kunnen worden ingezet in een risicobeoordelingsstrategie. In vervolgonderzoek zullen deze methoden worden beoordeeld op hun toegevoegde waarde, beperkingen en kosten. Daarnaast zal worden uitgezocht welke kinetiek-informatie minimaal nodig is om de vertaalslag te kunnen maken van de werking van een stof in cellen naar een heel lichaam.

Abstract

In this report the importance is shown of data on the behaviour of chemical substances in the body (the kinetics) for determining their health risks. Not only do these help in a better estimation of the health risks, they can also aid in reducing the number of necessary test animals for determining the harmful effects of a substances. With knowledge on the kinetics, certain tests can turn out to be unnecessary, because, for example, the substance turns out not be taken up in the body. Furthermore, such knowledge can prevent that tests have to be repeated, because, for example, the applied amount of substance proves to be too high or the animal species proves not to be representative for humans in this special case. These benefits could be of more use in the current system of risk assessment when the collection of kinetic data would be more integrated into this system and would occur earlier than the studies on toxic properties. The methods with which this integration can be realized with as few test animals as possible are described in this report.
To further reduce or completely replace animal testing, research is currently performed on a new way to assess risks. The new approach is based on tests with cells instead of whole animals. Such tests do no incorporate the kinetics of a substance in a whole body, therefore these tests will always have to be accompanied with computer models describing the behaviour of the substance in a body. Kinetic data will be crucial for these models. The current progress of the research on what such a new risk assessment approach should look like is described in this report, as well as the progress on the available methods for this approach. In continued research, the available methods should be critically evaluated for their added value, limitations and costs. Additionally, it would be useful to select a base set of necessary kinetic data.

Contents

Summary

The importance of toxicokinetic data for the current system of chemical safety assessment is increasingly recognized and increasingly implemented in legislative frameworks and guidance documents. However, these data are usually collected in parallel to toxicity test, or even afterwards, which hampers the use of their benefits. This report highlights the benefits of toxicokinetic with regard to reduction of the use of test animals (due to better-designed studies) and a sounder safety assessment (with less assumptions and a better view on what may be expected in human beings as compared to test animals). An overview of possible methods to acquire toxicokinetic data without requiring extra test animals shows that microdosing in humans, *in vitro* tests, and collection of blood and excreta within *in vivo* toxicity tests offer potential.
The resistance against animal testing, the high costs and long duration of such tests, and also the desire to further improve the protection of the human population has led to efforts to increase the 3R application and change the classical, animal testing-based approach to an approach using more modern tools. This report argues how toxicokinetic data have an even greater importance for such a future system of chemical safety assessment.
Toxicokinetic data can be used for exposure-based waiving, but are also necessary for the extrapolation of the safe concentrations determined *in vitro* to safe external doses *in vivo* for humans. Physiology-based pharmacokinetic (PBPK) models are a centrepiece in such an extrapolation. These need input of toxicokinetic data determined, for example, by *in vitro* kinetic tests or *in silico* predictions of kinetic parameter values. Some of the necessary toxicokinetic methods and tools for the new risk assessment strategy are already available, but others still need major further development.
Further research will critically evaluate the different tools described in terms of added value, limitations and costs. Simultaneously, it is planned to develop strategies, both for integration of toxicokinetic data into the current system of risk assessment, and for a new risk assessment approach. Costs and uncertainty management will be important elements for consideration in the formation of these strategies. Together, this research is expected to lead to the formulation of a base set of toxicokinetic parameters and methods for their determination.

1 Introduction

Russell and Burch first introduced the 3R ethical framework for human experimental testing in 1959 in their book 'The principles of humane experimental technique'. The three Rs stand for replacement, reduction and refinement of animal testing, in order of decreasing preference. This principle is now widely applied in e.g. toxicology, following increasing public resistance against animal testing. For example, in the EU the relatively new REACH regulation for chemical substances, includes an obligation for registrants of chemical substances to follow the 3R principle and only perform animal testing when unavoidable and then, to use the least possible number of animals and cause the least possible discomfort (EU Regulation (EC) No 1907/2006 (REACH), preambule note 47).
However, a safety assessment is required when human exposure might occur during the production and use of a chemical or food component. Thus, the various EU legislative frameworks regulations for the risk and safety of chemicals and products containing chemicals (e.g. for pesticides, industrial chemicals, food additives) still place requirements and/or recommendations for equivalents to toxicological animal test data. In practice, this means animal data are required, as there are as yet insufficient equivalent animal-free alternatives. These regulations roughly require the same toxicity tests, but also show important differences. Especially the requirements for data on toxicokinetics are very diverse, as shown in Table 1, while it will be argued that these data can be very beneficial for both the 3R principle and the adequate protection of people. The toxicokinetics of a chemical can be defined as measure of its absorption, distribution (in the widest sense), metabolism, and excretion (abbreviated as ADME) when dosed to an organism. Altogether, these processes determine how much of the chemical swallowed, inhaled, injected, or applied on the skin (the "external dose") will reach the organs or tissues where it can exert its toxic effect (the "target dose"), and thus, at what external dose toxic effects can arise. A more detailed description of ADME is given in Chapter 2.
The requirements for toxicokinetic data in the different regulations range from an assessment of available data (e.g. REACH) to full determination of ADME (e.g. biocides regulation). Historically, most emphasis in toxicology has been placed on the determination of the so-called toxicodynamics: the type of effects a chemical can cause. Animals serve as models for humans, with the application of a safety factor of 10, by default, to correct for any difference between the average animal and an average human being. This simple, but therefore uncertain representation of reality was first refined by determining more toxicokinetic parameters in the pharmaceutical sector. Kinetic data are used on a standard basis in drug development to assist in candidate selection, appropriate species selection for toxicity testing, and dose selection for toxicity studies, as well as in safety assessment, by comparing experimental animal versus human systemic exposure (Barton *et al*., 2006). Thus even before clinical testing, a full assessment of the kinetics is common practice already for a long time. A special ICH guideline from 1995 (CPMP/ICH/384/95; EMEA) describes how toxicokinetic data can be obtained and how they are useful. The requirement for full ADME determination has also been included in the requirements for the safety evaluation of active substances for plant protection products and of biocides since long, at least since 1991 in the EU (directive 91/414/EC). Also in the requirements for safety evaluation of food additives, toxicokinetic data are presently included (though not further specified), as well

as in those for food contact materials showing a migration of more than 5 mg/kg food.
Thus, the importance of toxicokinetic data for the current system of chemical safety assessment is increasingly recognized and increasingly implemented in legislative frameworks. The general benefits of kinetic data are that use of test animals may be reduced due to better-designed studies and that the safety assessment is sounder, with less assumptions and a better view on what may be expected in human beings as compared to test animals (e.g. Barton *et al*., 2006; Bessems and Geraets, 2013).

For the proposed future approaches towards chemical safety assessment, toxicokinetic data have an even greater importance. The resistance against animal tests, the high costs of such tests and also the desire to further improve the protection of the human population has led to efforts to increase the 3R application and change the classical, animal testing-based approach to an approach using more modern tools that are more predictive for humans. In the USA, the National Academy of Sciences presented a long-term vision on toxicological testing without using experimental animals (Toxicity Testing in the 21st Century, TOX21, US-NAS, 2007). TOX21 will be based on expected "advances in toxicogenomics, bioinformatics, systems biology, epigenetics, and computational toxicology, transforming toxicity testing from a system based on whole animal testing to one founded primarily on *in vitro* methods that evaluate changes in biological processes using cell lines, or cellular components, preferably of human origin" (Vermeire *et al*., 2013). The International Life Sciences Institute - Health and Environmental Sciences Institute (ILSI-HESI) has launched the Risk 21 project, with a Risk 21 Technical committee, dedicated to design a new risk assessment framework. This framework is meant to incorporate the recent visions on improved risk assessment, such as using existing data before planning tests, applying an exposure-driven approach, and using tiers (http://www.hesiglobal.org/i4a/pages/index.cfm?pageid=3492). In the EU, intelligent/integrated testing strategies (ITS) are being developed to save animal tests and costs, as has been done in e.g. the EU 6th framework project OSIRIS (e.g. Vermeire *et al*., 2013; Rorije *et al*., 2013) . An ITS is largely based on the collection of already available data, the use of quantitative structure-activity relationship (QSARs) and read-across predictions and *in vitro* assays, with animal testing only as a last resort.

Clearly, *in vitro* assays will have an important role in future toxicity testing. The major limitation of *in vitro* assays, in comparison to whole animal tests, is the lack of toxicokinetics in the test system. An *in vitro* gene mutation test in mammalian cells, for example, does not include the absorption of a chemical from the gut, the distribution through the body and the rate of excretion. Metabolism is partly covered by the addition of a liver enzyme extract (S9) to the cells. This is also acknowledged by Tice *et al*. (2013) in their update on the Tox21 programme. They state that Tox21 faces "some very difficult issues", since for example, extrapolation from *in vitro* concentrations to *in vivo* doses or blood levels is by no means straightforward. Additionally, assessing the effects of chronic exposure conditions is not (yet) possible with *in vitro* techniques (Tice *et al*., 2013). Thus, the *in vitro* assays are suitable for assessing the toxicodynamics at the target site, but in order to derive quantitative safe external doses, especially in case of repeated doses, toxicokinetic data will need to be added. In a report of a EURL ECVAM workshop, Adler *et al*. (2011) also emphasized the crucial importance of toxicokinetics in risk assessment without animal testing, which was supported in the recent report of the European

scientific committees SCHER, SCENIHR, and SCCS (2013) on future risk assessment. A future toxicity testing approach therefore cannot do without integration of such data, much more than currently done.

A previous report presented an overview of the currently available *in vitro* methods to measure ADME parameters (Brandon *et al*., 2012). The general aim of the present report is to describe how such toxicokinetic data can aid in the current and future human risk assessment of chemicals, and how these data can and need to be integrated into the risk assessment. Therefore, Chapter 2 briefly provides a more detailed understanding of the kinetic processes. Chapter 3 describes the benefits of toxicokinetic data for current risk assessment, making a case for including the generation of such data in more regulations (e.g. REACH). Chapter 4 gives possibilities on how toxicokinetic data can already be used to gain these benefits, with the currently available methods. Chapter 5 describes into more detail why and how toxicokinetic data need to be fully integrated in a future integrated testing strategy and what methods will be necessary for that.

Table 1. Overview of required toxicokinetic data for human risk assessment in the different regulations for the use of chemicals in the EU

REACH (EU Regulation 1907/2006)				**Cosmetics (EU Regulation 1223/2009)**	**Pesticides**		**Biocides (EU Regulation 528/2012)**	**Food additives (Commission Regulation EC 1331/2008 + EFSA Scientific Opinion on data requirements for the evaluation of food additive applications)**	**Food contact materials (Regulation EC 1935/2004 + SCF Guideline**			**Pharmaceuticals (Directive 2001/83/EC + ICH Guidelines)**
1-10 t/y (Annex VII)	**10-100 t/y (Annex VIII)[1]**	**100-1000 t/y (Annex IX)[1]**	**>1000 t/y (Annex X)[1]**		**Active substances (Council Directive 283/2013)**	**Plant protection products (Council Directive 284/2013)**			**Migration <0.05 mg/kg food**	**Migration 0.05-5 mg/kg food**	**Migration > 5 mg/kg food**	
-	Assessment from available data	-	-	OECD 417[2]	ADME[3] after oral exposure (and other route)[4]	*In vitro* human skin absorption	ADME[3]	Data required	-	Accumulation potential study	ADME[3]	ADME[3]

[1] The shown required tests for this tonnage/migration level are additional to the tests required at lower tonnage/migration levels; the tests required at the lower tonnage/migration levels are also required here.

[2] No requirements stated in regulation, but the OECD 417 test guideline for toxicokinetics is mentioned in "The SCCS's Notes Of Guidance For The Testing Of Cosmetic Substances And Their Safety Evaluation, 8th Revision" of 11 December 2012.

[3] Absorption, distribution, metabolism and excretion data required.

[4] Only if appropriate.

2 Absorption, Distribution, Metabolism, Excretion, Transport: basic understanding

The toxicokinetics of a substance describes its concentration in time on and in the body. It is distinct from the toxicodynamics, which describes the toxic effect of a substance. In the pharmaceutical domain, the toxicokinetics is focused on the fate of a substance in the body with emphasis on determining the amount of the active substance (either parent compound or metabolite) that reaches the part of the body where it will exert its toxic effect (the target) within a certain time. Toxicokinetic data enable the determination of the relation between the observed effects and this target dose. In non-pharmaceutical domains (pesticides, biocides, REACH, food additives etcetera) it is in general more aimed at using ADME data to help substantiate various extrapolations needed during risk assessment. One example is to use route- and species-specific data on absorption (e.g. oral absorption in the animal toxicity study versus human dermal absorption) instead of relying on default factors.
The toxicokinetics are governed by four physiological processes: absorption, distribution, metabolism and excretion. Because of its importance, a fifth process, i.e. active transport, is described separately even though it is actually included absorption, distribution and excretion processes. These five processes are described in more detail in the subsequent paragraphs. Figure 1 gives an overview of how a substance can be transported into and out of organs and tissues due to these processes. Table 2 summarizes how the information on these five kinetic phases can be used.

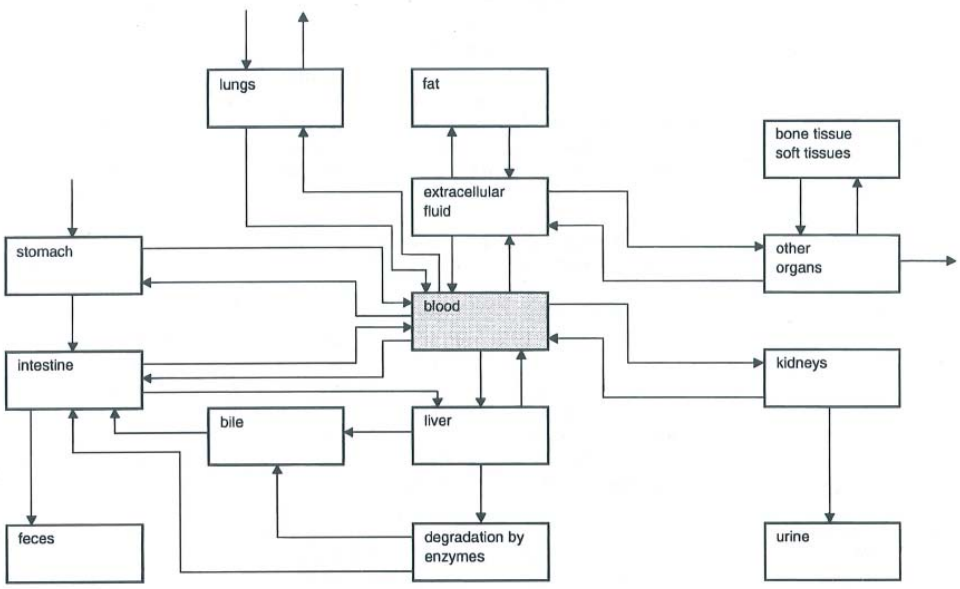

Figure 1. Schematic representation of the substance flow-pathways around different organs and tissues involved in the toxicokinetics of a substance, covering absorption, distribution, metabolism and excretion. Dermal absorption and excretion (through sweat) is not included here. (From: Niesink et al., 1996).

2.1 Absorption

When a chemical is applied on the skin (dermal exposure), inhaled into the lungs (inhalation exposure) or swallowed (oral exposure), it can exert local effects: irritation, corrosion , or an immuneresponse. For such effects, the kinetics are generally less relevant. To exert systemic effects (i.e. effects inside the body), however, a substance will first need to cross the skin, lung wall or gut wall to enter the inside of the body: absorption. Although absorption has also been defined as the process by which an unchanged drug (chemical) proceeds from the site of (external) administration to the site of measurement within the body (Bessems et al., 2013), it is in the present report defined as the fraction of a substance that passes the outer layer of the human body. The level at which a substance may be absorbed depends, for example, on its size and its hydrophobicity, as these determine how well the chemical can pass through a cell membrane. Also a number of external factors can influence the extent of absorption of a substance; for oral absorption this includes pH, presence of hydrolytic enzymes, influence of gut microflora, and the presence of food in the gut (Barton *et al*., 2006))
When discussing absorption, the term "bioavailability" is often used and sometimes even erroneously as synonymous to absorption. Bioavailability is defined as the fraction of a chemical in a certain matrix that reaches the systemic circulation unchanged within a defined time-frame. In specific cases, it may also refer to the bioavailability of a chemical to a specific organ. In that way, it is a complex parameter, combining several processes. Three processes (partially linear, partially in parallel) can be distinguished that determine bioavailability (Adler et al., 2011):
(1) release of the compound from its matrix, such as food (bioaccessibility),
(2) absorption of the released fraction and,
(3) metabolism before reaching the systemic circulation (in either the gut wall or in the liver).

Usually, bioavailability of a chemical refers to the parent compound, but it could refer to its metabolite in case this is the substance of interest for the toxic effects. It considers only one chemical form (parent or metabolite) and the entity it refers to (parent or a specific metabolite) has to be mentioned (bioavailability of a radiolabel without further information on the chemical speciation actually has very little, if any meaning). The difference between oral absorption (i.e., presence in gut tissue and portal circulation) and systemic bioavailability (i.e., presence in systemic blood and tissues) can arise from chemical degradation due to gut wall metabolism or efflux transport back to the intestinal lumen or pre-systemic metabolism in the liver, among other factors.
A very common method for establishing the extent of bioavailability is to compare the areas under curve (AUCs) of the concentration in blood or plasma over time, following intravenous (iv) administration and administration via the route of expected exposure (e.g. oral) based on the following relationship:

$$F = AUC_{\text{exposure route}}/AUC_{iv}$$

where *F* is the fractional bioavailability. It is also possible to use urinary excretion data to estimate bioavailability, based on the ratio of the total amount of unchanged chemical excreted in the urine after administration via the exposure route to that following iv dosing (Barton *et al*., 2006). However, this is not preferable, as one needs to be certain then, that the metabolism is the same for both administration routes. This information is often lacking.

2.2 Distribution

After entry into the body, the chemical will be distributed to different tissues and organs, depending on where it entered the body and on its chemical properties. For example, a substance entering through the skin, can be taken up by the lymphatic system and first reach lymph nodes before any other organs following the skin, while a substance entering through the lungs will directly come into the blood stream and reach the heart as the first organ following the lungs. Substances absorbed by the gut will directly reach the blood stream and pass the liver as the first organ following the gut. As another example, substances that are very hydrophobic will eventually accumulate in the body fat, in contrast to hydrophilic substances. Such substances, accumulated in fat tissues, can be released in relatively large amounts when the fat is digested, e.g. during a diet or during a breast-feeding period. Babies can then become exposed to lipophilic substances through their mothers' milk.
The distribution of a substance and its metabolites inside the body is governed by three main factors (Adler *et al*., 2011):
(1) the binding of the substance to plasma proteins,
(2) the partition between blood and specific tissues and,
(3) the permeability of the substance to cross specialized membranes, so-called barriers (e.g. blood–brain barrier/BBB, blood–placental barrier/BPB, blood–testis barrier/BTB).

2.3 Metabolism

In the liver, but also to some extent in other organs such as the gut and lungs, a chemical can be converted to other substances ("metabolism" or "biotransformation") by specific enzymes. These enzymes are meant to convert the "useful" nutrients to the necessary form to be used by the body and to convert "harmful" toxins to harmless substances that can be quickly removed from the body. However, these conversions can also lead to a more toxic substance than the original substances absorbed (i.e. bioactivation).
The general "aim" of the metabolic system is to make the substances more hydrophilic, so they can dissolve well in the blood and be distributed to the target organs (nutrients), or to make them more easily excreted via urine or bile (endogenous waste or xenobiotic chemicals). In general, the metabolic reactions can be clustered into phase I reactions (oxidations, reductions, hydrolyses) which introduce a polar group into the molecule, and phase II reactions, which attach (conjugate) an endogenous, hydrophilic substance to the molecule, e.g. glutathione (Niesink *et al*., 1996).

2.4 Excretion

Finally, a substance is removed ("eliminated", "cleared") from the blood via metabolism (see 2.3) or via excretion: removal through the faeces (including bile), urine, breath, and to a lesser extent via sweat, hair nails, milk, placenta or eggs. The route and extent of excretion of a substance depends on its physical and chemical properties. For example, volatile compounds will quickly leave the body via exhalation. Hydrophilic substances will remain dissolved in the urine and hardly be reabsorbed in the kidney tubules. Larger (typically >350 Da in rats and > 450-500 Da in humans (Barton *et al*., 2006)), substances are mainly excreted via the bile into the intestines and are subsequently excreted via faeces or re-absorbed into the body (enterohepatic cycle) (Niesink *et al*., 1996; Barton

et al., 2006). Transport of polar substances from the liver to the bile takes place through active transport (Niesink *et al*., 1996).

2.5 Active Transport

The most common route by which substances pass through membranes is by passive diffusion (Niesink *et al*., 1996): the molecules diffuse through the (cell) membrane which consists of lipophilic fatty acids. Thus, the substance needs some lipophilicity to enable such diffusion through a fatty environment. Hydrophilic substances can cross membranes either through water-filled pores in the membrane or by active transport (Niesink *et al*., 1996). The passage through the pores ("filtration"), which are lined by proteins to create a hydrophilic boundary, is restricted to small molecules (generally <100 g/mol) and requires a pressure gradient.
Active transport, requiring energy, involves a carrier mechanism enabling passage against a concentration gradient (i.e. from low concentration to high concentration). As the molecules need to be "carried", the molecules need to bind to a transporter in the membrane. This implies that only certain molecular structures will fit (substrate-specific transport) and that the transporters can become saturated, limiting the passage across the membrane to a certain maximum (Niesink *et al*., 1996). Transporters can have a function in the uptake of substances into cells, but also in the direct removal of substances from cells after diffusion through the membrane.

Table 2. Use of various kinetic (sub)endpoints in testing strategies (qualitative) and in quantitative risk assessment (adapted from: Brandon et al., 2012)

Endpoint	**Testing strategy (qualitative use)**	**Risk assessment (quantitative use)**
Preabsorption/Liberation	Low liberation may indicate waiving of testing Choice of most relevant route for systemic toxicity testing	Low absorption lowers risk estimate Route-to-route extrapolation Interspecies extrapolation High to low dose extrapolation
Absorption	Low absorption may indicate waiving of testing Choice of most relevant route for systemic toxicity testing	Low absorption lowers risk estimate Route-to-route extrapolation Interspecies extrapolation High to low dose extrapolation
Distribution	Organ distribution may be indication for organs that will be affected in toxicity studies	Could be indication of persistence and/or bioaccumulation
Metabolism	Species selection, gender selection, age selection	Part of kinetic modelling to be used for interspecies, intraspecies, route-to-route and high to low dose extrapolation
Excretion	Very slow excretion may indicate human persistency and/or bioaccumulation	To be assessed in comparison with exposure rate
Bioavailability	Low bioavailability may indicate very low risk for systemic effects	Low bioavailability may indicate very low risk for systemic effects
Ratio: $\frac{absorption\_rate}{excretion\_rate}$		Indication of retention, bioaccumulation and persistence To be assessed in comparison with exposure rate.

3 Potential benefits of toxicokinetic data for current human risk assessment

The specific benefits of toxicokinetic data for the current frameworks for chemical safety assessment can be grouped as follows:

1. Kinetics data can aid to waive unnecessary animal tests (→ Reduction);
2. Kinetics data can improve study design (selection test species, exposure route, dose, and analyzed organs → Reduction and Refinement);
3. Kinetics data can improve extrapolation from animal to human (→ improved health risk estimation).

Each benefit will be described into more detail and illustrated in the subsequent paragraphs.

3.1 Avoidance of unnecessary animal experiments

If a chemical is not absorbed into the body through the skin, lung or gut, systemic toxicity can be ruled out, cancelling the need for additional animal tests on systemic toxicity in theory. In practice, hardly any substance will have absolutely no absorption at all, in most cases there will be some level of absorption. Absorptions below a certain limit, however, may induce such low systemic exposure that toxic effects can be expected to be negligible. This idea is an extension of the concept of the Threshold of Toxicological Concern (TTC; Kroes *et al*., 2004) that provides a threshold for the external oral dose below which no toxic effects are to be expected. An "internal TTC" could be determined by assessing the level of absorption of the chemicals that underlie the TTC.
If the measurement of the absorption of any new chemical proves the internal (absorbed) dose of a chemical to be below this internal TTC some tests may be waived, thereby reducing the number of animals tested. Only local effects would still need to be assessed.
If exposure to a chemical is only via the oral route and toxicokinetic data show that it is very quickly metabolized in the liver to metabolites that are already well characterized regarding their toxicity, the toxicity tests of this parent chemical might be limited to local and liver toxicity tests, saving on other systemic toxicity tests.
Chemicals that have shown to be genotoxic in somatic cells need to be tested for their genotoxic potential in germ cells (in practice, sperm cells in all stadia of development) unless it can be shown that the chemical does not reach the germ cells. Here also lies a potential for the use of toxicokinetic data, in this case on distribution, to save on an animal test. Considering that for the one test for germ cell genotoxicity (Comet assay), there is no accepted OECD guideline yet, and that the other test (transgenic rodent gene mutation assay) is only limitedly available, waiving such a test has an extra advantage.
As a last example, when chemicals are shown not to cross the placental barrier (and not to induce developmental effects indirectly, e.g. by damaging the placenta), this might support a waiving of developmental toxicity tests as the chemical will then not reach the embryo or foetus. Postnatal toxicity via lactation can be covered by a generational reproduction toxicity study.

3.2 Improvement of study designs

Toxicokinetic data can help to improve study designs through better selection of the test species, exposure route, dose and target organs to be analysed (Bessems and Geraets, 2013). Unnecessary tests can thus be avoided (Reduction) but also the discomfort of animals may be reduced (Refinement) as e.g. high doses inducing severe toxic reactions may be avoided or even turn out to be superfluous.

Test species selection

For the selection of the laboratory test species, important factors are the similarity of the anatomy and physiology to that of humans (especially for the endpoint of interest), the lifetime (e.g. it should not take too long to cover multiple generations), the ease with which organs can be dissected and analyzed (not too small), and the size of the animal (for costs of housing and feed). For these reasons, the species mostly used in toxicity tests for human risk assessment are the mouse and rat, followed by rabbits and dogs. Rats and mice, however, do have anatomical and physiological differences to humans that can affect the toxicokinetics of a substance and thus its toxicological effect. For example, rats do not have a gall bladder whereas humans do (possibly changing the digestion and thus absorption in the intestine and possibly affecting the excretion of substances through the bile). As another example, absorption of intact beta-carotene is very low if not zero in rodents as large quantities are converted in the gut to vitamin A, thus affecting its bioavailability. In contrast, in man 20–75% of the beta-carotene ingested is absorbed intact. This caused the European Food Safety Authority's (EFSA) Panel on Food Additives and Nutrient Sources Added to Food (ANS) to dismiss rodents as suitable models for evaluating the bioavailability and effects of beta-carotene in humans (Bessems and Geraets, 2013).

In addition, the dogs required for testing data for market acceptance of plant protection products are not always a relevant species. For example, in a comparison of the elimination data of multiple species (including humans) for phenoxyacetic acid herbicides, dogs appeared to have an exceptionally slow elimination leading to a higher sensitivity towards these substances compared to humans (Timchalk, 2004).

These examples illustrate the importance of insight into the toxicokinetics of a substance and of knowledge on how species differ in the mechanisms relevant for the toxicokinetics of that substance. Then, the species most resembling humans in these mechanisms can be selected.

In vitro studies could help to determine the best animal model by assessing e.g. the metabolites generated (Barton *et al*., 2006). Animal species with a different metabolic profile compared to humans could be considered inappropriate for toxicity studies (Barton *et al*., 2006). A major advantage of in vitro studies is that they can frequently be done for both the species in the toxicity study and in humans, thus facilitating interspecies comparisons (Barton *et al*., 2006). To know the relative importance of each of the different kinetic processes, however, an in vivo test would be necessary (Barton *et al*., 2006).

Dose selection

At the UK NC3R (National Centre for the Replacement, Refinement and Reduction of Animals in Research) workshop in 2008, dose selection for toxicity studies was the area where toxicokinetics were considered to have the greatest impact, both in improving risk assessment and in reducing and refining animal use (Creton *et al*., 2009). Testing protocols (e.g. OECD guidelines) require tests

to be performed at the maximum tolerated dose (MTD; determined in an earlier or range-finding test) and to two to three dose levels below that, usually spaced by a factor two to four. The driver for testing at the MTD is to obtain greater power with a limited number of animals (as the level of effect is as high as possible) but the MTD (and also the chosen lower doses) can also saturate or overload kinetic processes such as active transport and metabolism (Barton *et al*., 2006). High exposure levels, inducing saturation of detoxifying or eliminating systems, may result in toxic effects that would not arise at lower dose levels, more close to realistic exposure levels, that do not induce saturation. In such a case, extrapolation of the effects observed at the high dose levels to low, real-life dose levels will lead to an underestimation of the toxic effects at these lower doses.

For example, kinetic data for methylene chloride indicated that saturation of a high affinity but low capacity CYP-mediated detoxification pathway leads to a shift in metabolism via a glutathione transferase pathway and a subsequent production of disproportionately high levels of genotoxic metabolites that resulted in carcinogenicity. In hindsight, even the lowest concentration of methylene chloride administered in a US National Toxicology Program (NTP) bioassay was at a level where the CYP pathway was already saturated. This information, together with data on the enzyme kinetics for methylene chloride metabolism in different species and human tissue, was used to determine that low-dose environmental human exposures should present minimal risk of cancer (Creton *et al*., 2009).

As another example, in chronic inhalation studies with vinyl chloride, the incidence of tumours was approximately the same at exposure concentrations of 2500, 6000 and 10,000 ppm (ppm). Subsequent studies on the metabolism of this compound predicted that systemic exposure to the reactive metabolite associated with tumorigenic effects was essentially the same at all three exposure concentrations due to saturation of the enzyme responsible for producing the metabolite (Creton *et al*., 2009). Thus, prior knowledge on toxicokinetics (i.c. metabolism) could have shown that exposure to the two higher doses would be redundant which could have saved animals (i.e. a repeated test with lower doses would then not have been necessary). Therefore, preliminary *in vivo* and/or *in vitro* studies on metabolism can provide important information for dose selection. These studies should best be performed at different dose levels ranging from realistic human exposure levels to those planned for the toxicity experiments to identify saturable kinetics (Barton *et al*., 2006).

Guidance accompanying REACH indicates that toxicokinetic data could assist in dose selection for repeated dose toxicity studies by selecting the maximum dose level at the inflection point of dose-AUC relationships, as this can be regarded as the kinetically derived maximum dose (Creton *et al*., 2012). The selection of such a 'kinetically derived maximum dose' (KMD), has also been suggested by others as preferable to selection of an MTD or even to a dose causing toxicity, provided there is an adequate margin between test dose and predicted human exposures and the toxicokinetic processes in the test species are relevant to humans (Creton *et al*., 2012). In the determination of such a KMD, it must be considered that repeated daily exposures will increase the blood concentration over a period of 4–5 half-lives to establish a plateau if the half-life is longer than a few hours (Barton *et al*., 2006). In addition, repeated exposure can contribute to altered toxicokinetics and toxicological responses as compared to that seen following a single dose. This might primarily be due to induction of metabolizing enzymes but also due to inhibition of metabolizing enzymes or alterations in transporters (Barton *et al*., 2006). Thus, it is necessary to have kinetic data from single and repeated exposures. Finally, it must be taken into account that

protein binding differences between species can cause differences in the free fraction of the substance, which is generally the fraction available for causing toxic effects.
These data will aid to avoid unnecessary animal testing or avoid testing unnecessary high doses that are accompanied by animal discomfort. Further, insight in (differences in) kinetics at experimental dose levels and human exposure levels will lead to better prediction of human health risks.

Selection of organs to analyse
Prior knowledge on the absorption of a substance, on whether it is completely metabolized or not, on its distribution pattern and on excretion, can help to identify potential target organs to estimate the target organ doses. For example, assume a substance that is absorbed only by the oral route and that is completely metabolized in the liver with the metabolites completely excreted via the biliary route or known to be innocent. It can then be considered much less important to analyse organs other than the gut and liver for toxic effects (if indirect effects can be excluded). If *in vitro* genotoxicity tests for such a substance would indicate a genotoxic potential a follow-up *in vivo* erythrocyte micronucleus test would be superfluous (as the bone marrow is not reached) and an *in vivo* Comet assay would best be performed in gut and liver. This way, unnecessary animal experiments can be avoided leading to a reduction in the use of test animals.

3.3 Improvement of extrapolation of results from animal experiments

The benefit of well-designed studies is not only that unnecessary additional animal tests may be avoided but also that the data are better suited to use in the risk assessment for humans. As explained in the preceding section (under "Test species selection") there are physiological and anatomical differences between test animals amongst themselves but also between test animals and humans. To extrapolate a safe dose for a test animal to a safe dose for humans, these differences must be considered. It is preferred to derive a chemical specific assessment factor but this is seldom possible. Then, often a factor of 4 is taken for the kinetic differences between a rat and a human being by default. An example of a known difference is the following. Albumin, sex-hormone-binding globulin (SHBG), α1-acid glycoprotein, and α-fetoprotein are some of the major serum proteins involved in non-covalent binding of endogenous and exogenous compounds. SHBG is a primate protein with very specific binding affinity for the sex hormones estradiol and testosterone; it is not present in rodents. In contrast, rodent α-fetoprotein binds estradiol with high affinity, limiting its availability (Barton *et al*., 2006) while the human form does not bind estradiol. Thus, the availability and distribution of sex hormones and similar structures can be very different in rodents compared to humans. As another example, rat skin appears to be more permeable than human skin (Barton *et al*., 2006), overestimating the human absorption 1.5- to 14-fold for some pesticides (Barton *et al*., 2006). The ratio of human to rat in vitro dermal absorption can be used as a means to estimate in vivo human dermal absorption from in vivo rat dermal absorption data: $in\ vivo_{human} = (in\ vitro_{human}/in\ vitro_{rat}) \times (in\ vivo_{rat})$ (Barton *et al*., 2006).

Besides extrapolations from animal to human, route-to-route extrapolations are often necessary, for example, when toxicity data are available from one route only. In that case, route-specific kinetic data such as on absorption and on biotransformation are essential. Toxicokinetic data may also be important for

extrapolations over life stages as some kinetic parameters can change over the various life stages (Barton *et al*., 2006). For example, humans and rodents express *α*-fetoprotein at highest levels in utero with concentrations dropping dramatically after birth (Barton *et al*., 2006). This changes the protein binding and thus availability of substances over these life stages. Elderly show slower kinetics and excretion (decreased kidney function) but the available studies indicate that these are sufficiently covered by an uncertainty factor of 10 for intraspecies variation (Cooper *et al*., 2006).

Preferably, comparison of relevant exposures in animal experiments to humans should be made on internal dose metrics rather than external doses or concentrations. For example, the lowest liver concentration of a test substance that produces liver toxicity in the rat is to be compared to the liver concentrations in humans arising from their predicted (or measured) exposure. The differences in kinetics are then implicitly included in the parameters that are compared. To achieve this, current toxicity study designs should include blood and excreta analyses, enabling potentially more meaningful animal to human dose–exposure comparison if human blood levels are collected under conditions of real-world exposure (Saghir *et al*., 2006). Obviously, blood and excreta concentrations are easily measured, while e.g. liver concentrations would have to be modelled from kinetic data.

4 Integration of toxicokinetics in current risk assessment

The evident question following the description of the benefits of knowing the toxicokinetic characteristics of a substance is how kinetic measurements can be integrated with toxicity experiments. Important is that these measurements do not require that many additional animals themselves that in the end there is no net reduction in animal use. To be of maximum utility, it is desirable that toxicokinetic measurements be closely integrated with the toxicity testing protocol (Wilson *et al*., 1994). Currently however, most toxicokinetic data are obtained by separate animal studies performed simultaneously with or later than the required toxicity studies. A consistent set of recommendations is lacking that aids the registrant or registering authority to decide what ADME data are best applied in risk assessments (Barton *et al*., 2006). Thus, in non-pharmaceutical regulatory frameworks, toxicokinetic studies are often performed just before submitting the dossier to the regulators, mainly to fulfill mere regulatory requirements than to be used in an intelligent testing strategy. This way, extra animals are necessary for these measurements, which hampers waiving additional experiments or improving the design of these additional studies.

This chapter therefore describes how toxicokinetic data can currently be obtained *in vivo*, in humans and in animals, either by combining with other analyses or *in vitro*. The *in vitro* tests and the *in vivo* studies in humans are most preferable from an animal welfare point of view (of course where experiment s in humans would be ethically acceptable), while the *in vivo* analyses in animals are the least preferable, even though they are meant to reduce the number of animals and enhance their welfare. From the point of view of human health protection *in vivo* human data are most preferable as they prevent necessary extrapolations (e.g. interspecies). However, there are some drawback as well. Measuring ADME/TK under realistic exposure conditions may be unethical as long as little or no toxicity data are available.*In vitro* data with human cells or tissues and *in vivo* animal data have a roughly equal value as both have extrapolation uncertainties to be dealt with. Therefore, altogether, the order of preference is: first *in vivo* studies in humans, then *in vitro* tests, and if none of these are possible, *in vivo* animal studies.
Toxicokinetic data can also be obtained *in silico*, i.e. by computer models such as quantitative structure activity relationships (QSARs). These predict a specific property (e.g. binding constant to the protein albumin) based on the chemical structure of a substance. As no overview on the available QSARs for kinetic parameters was found, these models were not included in the description of methods below. It is expected that not many QSARs are available for kinetic parameters yet, as most focus has hitherto been on the toxicodynamic properties. Additionally, predictions from such QSARs always suffer from uncertainty inherent to model predictions, which makes them less preferable than *in vitro* determinations. They may serve well in obtaining first impressions of the kinetic parameter values, which is useful for prioritization of substances and of further tests. But finally, experimentally determined values will be desirable.

4.1 Collection of human toxicokinetic data *in vivo*

Human toxicokinetic data are of interest to determine internal human doses to enable the comparison of internal doses with those in experimental animals and

to detect interspecies differences. Such human data can since long be obtained by biomonitoring: the collection and analysis of human blood, urine, faeces, and exhaled air samples. Although this provides valuable information on internal dose and elimination it has its limitations. For example, biomonitoring is only possible for substances already in use and does not give specific information on distribution.
The recently developed technique of microdosing can provide more information on the internal absorption and distribution of a substance. New analytical techniques such as positron emission tomography (PET) and accelerator mass spectrometry (AMS) (Creton *et al*., 2009) have enabled the detection of minute amounts of radiolabeled substances in intact human bodies. Thus, a very low dose ("microdose") of radiolabeled test substance can be given to human volunteers, low enough not to cause any health risk, to determine its toxicokinetic profile. In order to ensure the microdose is indeed safe, normally short-term animal toxicity studies would need to be conducted, which would counter the 3R concept. However, a possible approach dismissing the need for prior animal testing could be to combine it with the TTC (Adler *et al*., 2011). In principle, human microdosing could possibly obtain ethical approval by keeping the total dose below the relevant threshold in TTC terms. Usually, an amount somewhere in between 1 and 100 µg is administered (Adler *et al*., 2011). If the chemical is not a genotoxic compound (sufficient *in vitro* methods are available to assess this) and not an organophosphate, the lowest threshold for exposure below which adverse effects are unlikely is 90 µg/day. Acknowledging that this threshold was based on lifelong exposure, it can be argued that health risks are not to be expected for the volunteers that receive microdosing. The exposure route should be considered then, because the current TTC concept is completely based on oral toxicity studies (Adler *et al*., 2011). Additionally, the detection of doses below the TTC in the blood and tissues after its distribution might be an issue (Bessems *et al*., 2013), just like the costs of the radiolabeled substance and the advanced analytical technique. Finally, it must be verified whether such a dose is not too different from the expected external exposure, or, if there is a considerable difference, if differences in kinetics are to be expected. Microdosing data may help to verify non-animal physiologically-based pharmaco-kinetic (PBPK) modelling predictions (see paragraph 5.1) before their widespread use for quantitative in vitro in vivo extrapolations (QIVIVE) and human risk predictions (Bessems *et al*., 2013).

4.2 Collection of human and animal toxicokinetic data *in vitro*

The available *in vitro* methods for determination of toxicokinetic data are described in a report of Brandon *et al*. (2012). This report makes clear that most of the relevant, available *in vitro* models are for the areas of oral and dermal absorption, protein binding and hepatic metabolism. To give a summarizing overview, the reported available methods for these endpoints (except protein binding) are given in tables 3-5. Protein binding and also any distribution between plasma and tissues can be determined by ultrafiltration, ultracentrifugation, equilibrium dialysis or nd-SPME (Brandon *et al*., 2012; Vaes *et al*., 1996; Artola-Garicano *et al*., 2000).

Absorption

For determination of oral absorption (table 3) the assay with the Caco-2 cell line seems to provide the best options in terms of availability of the biological material, throughput and validation. However, the validity of the Caco-2 model predictions for non-pharmaceuticals remains to be established, because this

model has primarily been applied for pharmaceutically active ingredients. While in the pharmaceutical R&D reliable prediction between 50 and 100% absorption is important, in the arena of industrial chemicals much lower absorption percentages can occur which are important to estimate correctly (Adler *et al*., 2011; Bessems *et al*., 2013). The possibility to translate this *in vitro* to an *in vivo* absorption percentage is easy for compounds which are passively absorbed, water soluble and do not require pre-systemic metabolism, but for other compounds the model is reported to have limitations, e.g. for active transport (Bessems et al., 2013).
In addition to the available *in vitro* assays, a quantitative property-property relationship (QPPR) relating percentage of absorption as a function of log K_{ow} has been published and re-formulated for derivation of *ka* (Bessems *et al*., 2013).

Regarding *in vitro* models for dermal absorption (table 4), it has been stated that currently still some *in vitro* skin models may be much more expensive than animal *in vivo* models. An example of this is the test with reconstructed human tissues while excised skin, either of human or porcine nature can be an economical alternative. However, it has also been reported that, the analytics or synthesis of radiolabelled compounds is at present rather a cost driver than the barrier itself (Bessems *et al*., 2013). As there is already an OECD Technical Guideline (TG 428; OECD 2004b) available for the method using animal and human skin, this *in vitro* method can already be used to produce acceptable data. The only issue to note is that it can be used as a worst estimate for the risk assessment as it determines the maximum absorption (risk assessment issue) but that improvement is needed with respect to the determination of the absorption rates. Therefore, it is not directly useful for PBPK modelling for QIVIVE purposes (Bessems *et al*., 2013).

In vitro models for absorption in the lung are in various stages of development and it has been estimated that several more years of intensive research will be needed to provide suitable models that can enter prevalidation (Adler *et al*., 2011).

Table 3. In vitro models to investigate the oral absorption and bioavailability (from: Brandon et al., 2012)

Test name	**Comments**	**Endpoints measured**
Oral bioaccessibility	Different *in vitro* digestion models have been developed that vary in simulated juices and static versus dynamic model.	Maximal bioavailiability
Cell monolayers, including transporter transfected	Provides information on the permeability of a compound, but also provides information on the involvement of transporters thus on active uptake or efflux. For example, Caco-2 cells transfected with P-glycoprotein. Current limitation: These models have been used and evaluated for 'pharmaceutical-like' physico-chemical properties and in the high output range, i.e. high absorption but not in the lower absorption range.	Absorption
Artificial intestinal membranes	Provides only information on passive diffusion. A draw-back of this method is the underestimation of the permeability of highly lipophilic drugs. For example PAMPA (parallel artificial membrane permeability assay)	Absorption
Intestinal segments	With an Ussing chamber, the transport across a section of the intestinal tract can be measured. The outcome of the experiment depends on the intestinal section used. To predict the human outcome, human intestine have been used.	Absorption + metabolism = bioavailability
Everted sac	The outcome of the experiment depends on the intestinal section used. Human and pig intestine have been used to predict human bioavailability.	Absorption + metabolism = bioavailability

Table 4. In vitro models to investigate the dermal absorption and metabolism (from: Brandon et al., 2012)

Test name	**Comments**	**Endpoints measured**
Stratum corneum	The model provides information on the maximal absorption across the skin. No metabolism is present in the stratum corneum. Human stratum corneum has to be used to predict human absorption.	Absorption
Immortalised keratinocyte cell line	The cells are cultured in monolayers to mimic the human skin. However, no stratum corneum is present and this could lead to an over-prediction of the absorption. Cell lines can have a smaller enzyme activity compared to primary cells. For example the human keratinocyte cell line HaCaT	Absorption + metabolism = bioavailability
Primary keratinocytes	Primary keratinocytes can be isolated form laboratory animals, pigs or humans. Human and pig will most closely predict the human outcome. However, the stratum corneum is not present and this could lead to an overprediction of the absorption.	Absorption + metabolism = bioavailability
Artificial skin	The artificial skin model consists of a collagen scaffold (the protein underlying the skin structure) and thus provides only information on the maximal absorption.	Absorption
Reconstituted epidermal models	The reconstituted human epidermal models are produced from human keratinocytes originating from a pool of different donors, *e.g.* from fore skin, and are cultured in 3D.	Absorption + metabolism = bioavailability
Animal and human skin	Skin from laboratory animals, pigs and humans can be used. However, to predict the human outcome human or pig skin has to be used. Pig skin most closely resembles the human skin compared to laboratory animals. Phototoxicity can be also be studied with this model.	Absorption + metabolism = bioavailability

Distribution
As previously explained, the distribution of a substance and its metabolites inside the body is governed by three main factors (Adler *et al*., 2011):
(1) the binding of the substance to plasma proteins,
(2) the partition between blood and specific tissues and,
(3) the permeability of the substance to cross specialized membranes, so-called barriers (e.g. blood–brain barrier/BBB, blood–placental barrier/BPB, blood–testis barrier/BTB).
Although there seem to be no OECD guidelines for methods to determine these parameters, there are well-accepted in vitro methods to measure protein binding (e.g. Oravcova *et al*., 1996; Vaes *et al*., 1996) and tissue-blood partitioning (e.g. Artola-Garicano *et al*., 2000). The BPB passage can be determined using discarded human placentas and for determination of the BBB passage in vitro methods are in development using transfected cell lines (Brandon *et al*., 2012).

Metabolism
In vitro methods for metabolism studies are available at different levels of biological complexity, ranging from methods with single metabolic enzymes (obtained from living animal or human donors or reconstituted in cell systems) to methods using liver slices (see table 5). The latter most resemble the *in vivo* situation but only provide a determination of which metabolites are formed and at what speed. Studies with single enzymes will need a battery of tests (more experimental work) and need a more elaborate extrapolation to a physiological *in vivo* environment, but they do provide qualitative and quantitative information on which enzymes are involved in the metabolism. This information can be used to predict potential interactions of chemicals (e.g., inhibition) to determine whether species differences may be expected. It can also be used to determine whether large individual differences can be expected in the human population (as some enzymes are subject to polymorphism) or with a distinct development (i.e., age-dependent) profile. Thus, studies of the metabolism of a chemical using single enzymes from different tissues, species, age groups or human populations present a valuable tool for obtaining qualitative and quantitative metabolic information for extrapolations (Barton *et al*., 2006).
In between these two described extremes are the popular systems consisting of subcellular fractions (e.g. microsomes) and liver cells either obtained fresh from whole livers (primary cells) or obtained from continuous cultures (cell lines). These provide the same type of information as the liver slices but the biological material is easier to acquire and handle enabling higher throughput. Microsomes are easy to obtain but do not contain all phase II enzymes (Bessems *et al*., 2013), thus lacking an important part of the metabolic system. An advantage of using microsomes is that they can be obtained from different tissues (e.g. gut and lung), to determine the extent of metabolism in these tissues and compare it to that in liver.

Table 5. In vitro models to investigate the metabolism of a compound (from: Brandon et al*., 2012)*

Test name	Comments	Endpoints measured
Recombinant enzymes	Microsomal or cytosolic subcellular fraction isolated from cells transfected with a specific human enzyme, *e.g.* CYP3A4 (major human detoxifying enzyme). Recombinant enzymes are available commercially.	Disappearance of parent compound, appearance of (major) metabolites and/or metabolite profile
Subcellular fractions	Microsomes, cytosol and S9 fraction isolated from organs from laboratory animals or humans. The subcellular fractions can be bought commercially.	Disappearance of parent compound, appearance of (major) metabolites and/or metabolite profile
Primary hepatocytes	Cells isolated from organs from either humans or laboratory animals, *e.g.* hepatocytes. Human cryopreserved primary hepatocytes are commercially available, even from different individual donors to study inter-individual variability.	Disappearance of parent compound, appearance of (major) metabolites and/or metabolite profile
Hepatocyte cell lines	Cell line isolated from hepatic tumours. Most cell lines have limited enzyme activity and are therefore less suitable to investigate the metabolism of a compound. For example, HepG2.	Disappearance of parent compound, appearance of (major) metabolites and/or metabolite profile
Liver slices	Liver slices can be used to study the metabolism, but are complex and not viable for a long period and therefore only suitable if the metabolism rate of the compound is high.	Disappearance of parent compound, appearance of (major) metabolites and/or metabolite profile
Induction of metabolism	Enzyme induction by a compound can be studied in primary human[1] hepatocytes.	CYP induction as activity
Inhibition of metabolism	The inhibition capacity of a compound can be studied in recombinant enzymes, subcellular fractions and primary hepatocytes.	CYP inhibition as activity

[1] As enzyme induction is very species-specific, it should only be determined in human material.

Primary liver cells (freshly isolated or frozen hepatocytes) have the advantage of resembling *in vivo* liver cells most but their preparation is cumbersome. They do not survive very long *in vitro* and also lose their initial metabolic activity very quickly, which makes their commercial purchase (which is possible) not very advantageous. Primary hepatocytes in sandwich culture (i.e. kept in between layers of collagen) show a better survival and viability but the collagen layers can form a diffusion barrier for the tested substance. Cell lines are easily maintained and give a stable metabolic activity. The most popular cell lines are the HepG2 and the HepaRG cells lines. First comparisons between these two cell lines indicate that HepaRG cells resemble the *in vivo* (and primary liver cell) metabolic activity better than HepG2 cells (Doktorova *et al*., 2012). It is worthwhile noting that a cell line such as HepaRG cannot provide information on human inter-individual variability while pooled human primary hepatocytes can provide a population average estimate (Bessems *et al*., 2013).
There is a general consensus that metabolically competent human primary hepatocytes or liver cell lines are the best enzyme source to perform the first primary screening of metabolism in the liver (Adler *et al*., 2011). The two most important endpoints measured are (1) intrinsic clearance (CL_{int}, i.e. disappearance of the test substance, mostly by metabolism) which can be extrapolated into hepatic metabolic clearance and (2) the identification of metabolites (stable, inactive, active or reactive metabolites of concern) (Adler *et al*., 2011).

Quantitative *in vitro - in vivo* extrapolation, i.e. scaling-up of the *in vitro* intrinsic clearance to the *in vivo* hepatic metabolic clearance can be performed using scaling factors (Barter *et al*., 2007) and physiological parameters (Brown *et al*., 1997). Wetmore *et al*. (2012) determined the intrinsic metabolic clearance (CL_{int}) for first-order conditions of metabolism in the liver at low concentrations *in vitro*. The *CLint* in the human hepatocyte culture was experimentally determined at 1 µM as was the slope of the disappearance of the chemical over time. Clearance was normalized to cell number, with the units µl/min/10^6 cells. *In vivo* intrinsic clearance was estimated by simply multiplying the *in vitro* clearance by the number of cells per gram of liver and the weight of the liver (see Basketter *et al*., 2012). (Bessems *et al*., 2013)
A more elaborate approach is to put the CL_{int} into a PBPK model. This has the advantage that instead of CL_{int}, also K_m and V_{max} can be used to simulate non-linear (saturating) dose ranges. The differences across species or life stages can also be incorporated into kinetic models. These models can serve to predict the final target dose in humans *in vivo*, to predict specific populations that may have a higher risk of toxicity, to make cross-species extrapolations (Barton *et al*., 2006) or life-stage extrapolations (Barton *et al*., 2006) either early in product development for purposes of toxicity study design, or for interpretation of toxicity studies and application in risk assessment (Barton *et al*., 2006). See chapter 5 for further description of such kinetic models.

Excretion
Whereas for biliary excretion some advances have been made with *in vitro* models (i.e. sandwich-cultured hepatocytes) no reports could be identified in the literature on *in vitro* models of renal excretion (Adler *et al*., 2011). This is therefore an area that deserves priority in further development of *in vitro* alternatives to animal tests.

4.3 Collection of animal toxicokinetic data *in vivo*

For an optimal integration of toxicokinetic data in the risk assessment process based on animal tests a tiered approach has been proposed by Wilson *et al*. (1994; see Table 6). An important aspect of this tiered approach is the generation of ADME data at an earlier stage during the toxicity testing of a chemical. For tier 1 a minimum experimental data set for a chemical was defined addressing three fundamental questions (Wilson *et al*., 1994):

1) Is the chemical absorbed?
2) Is the chemical metabolized?
3) Is the chemical persistent?

Table 6: Tiered approach in the collection of toxicokinetic data for risk assessment (from: Wilson et al*., 1994)*

Tier	Toxicokinetic data collected
Tier 1	Minimum data set: - absorption (analysis excreta and/or blood, iv + additional exposure) - metabolism (analysis excreta and/or blood for parent and metabolites) - elimination (analysis excreta and/or blood at different time points → $t_{1/2}$)
Tier 2	Dose response Repeat dosing Route of exposure Tissue distribution
Tier 3	Metabolite characterization Additional studies (e.g. binding studies; species, sex, strain, and age considerations) Toxicokinetic modeling

It was recognized that this minimum data set of tier 1 would not be sufficient to steer subsequent toxicity studies, it was meant more as a first set of data upon which it can be decided if more kinetic data are necessary (tier 2). A minimal experimental design was proposed for tier 1 studies, with the following characteristics (Wilson *et al*., 1994):

- 3 animals per dose level
- 1-3 dose levels
- 1 sex, male preferred
- Healthy young adults, matched for age and body weight
- Rat preferred, unless other species used in toxicity study
- 7 days or until 90% eliminated
- Relevant route(s), plus iv (absorption determination)

It is unclear whether this minimal design included a single dose or repeated dosing. In case of "significant findings" (not further defined) in tier 1 and a need for kinetic data to steer toxicity studies, tier 2 would be triggered. The studies in tier 2 are aimed to address, among others, the following questions:
Do the kinetics become nonlinear (i.e. saturated) at relevant doses (i.e between human exposure level and LD_{50})?
Do the kinetics change with repeated dosing?
The tier 3 studies are considered to be triggered either by the results of the kinetic studies in tier 2, or by the results of toxicity studies. Thus, it seems that Wilson *et al*. regard the tier 1 and tier 2 studies to be performed before the

toxicity studies (in order to steer those), and tier 3 before or after or in-between the toxicity studies, depending on when they are triggered (to provide more in-depth information).

For the *in* vivo determination of the extent of absorption and the time of elimination, the level of the substance in excreta, expired air and/or blood must be measured. The time for collection of excreta and blood samples should span a sufficient time period to allow accurate determination of the extent of absorption and the time course of elimination, which will differ per substance. To determine the metabolic conversion of the substance, blood or excreta should be analyzed for metabolites, for example using a radiolabeled substance. For tier 2 studies where the question of nonlinear dose effects are to be addressed, two or three dose levels should be used. If the study is designed to help in setting dose levels for a toxicity study a minimum of three dose levels should be used (Wilson *et al*., 1994).
In case a radiolabeled substance is used, this study would have to be a separate study, not integrated with a study for toxic effects, as the radiolabel can influence the toxic effects of a substance. In that case, animal reduction would be achieved if this study with only 3 to 9 animals would save on further animal studies (usually using much larger numbers of animals), as described in chapter 3. The radiolabel is mainly necessary to detect the metabolites of the substance among the thousands of substances present in blood and excreta. The radiolabel may be omitted in case the metabolites can be determined through a different method, e.g. *in vitro*, which should be feasible according to paragraph 4.2. In that case, the proposed minimal data set can be collected within a toxicity study with the limitation that the animals can only afford to lose a minimal amount of blood during such a study.

To this end, Burtin *et al*. (1996) have described the use of a single blood sample from each animal (with different animals sampled at different time points) to calculate the area under the blood concentration – time curve (AUC) and the maximal blood concentration (C_{max}). The AUC provides an indication of the total systemic exposure over time (expressed as quantity * time, e.g. µg * h), while C_{max} indicates to which maximal concentration the animal is exposed internally. Both are important kinetic parameters. This approach with one single blood sample represents a minimally invasive approach, with the variability between the individually sampled animals as a downside.
A more accurate measurement of AUC may be obtained with additional sampling in the same animal. Jochemsen *et al*. (1993) determined that three blood samples taken at selected times from each animal during a 24 h period would accurately estimate AUC and C_{max} with short, intermediate, and long half-life pharmaceuticals. With rats as nocturnal animals, eating the spiked food mostly at night, the best times for these three samplings were found to be at early morning when the lights were turned on (timepoint of C_{max}), at mid-morning (3 h after lights are turned on), and in late afternoon (2-3 h before lights were turned off (timepoint of C_{min}) for rats. The AUCs determined from these three time points were found to be 103 ± 10% of the original AUC-24 h, with 13 out of 17 values ranging between 96 and 105% of the original (Saghir *et al*., 2006). In contrast to rats, timing of C_{max} and C_{min} in mice varied and did not follow any pattern with reference to the light and dark cycle as feeding/drinking patterns of mice are not as nocturnal compared to rats. However, in the case of exposure through diet, the described schedule for collecting 3 blood samples for AUC-24 h in mice still afforded 98–99% conformity despite the variable nocturnal feeding behavior (Saghir *et al*., 2006).

The collection and analysis of 1–3 blood samples (≤ 100 µL each) per dose on one day may provide insight into possible saturation of absorption or elimination or some other phenomenon warranting further investigation. In addition, collection of the terminal blood samples from rats, which is usually conducted after 18 h of fasting, will be helpful in rough estimation of blood/plasma half-life of the compound (Saghir *et al*., 2006). This information can be obtained from study animals without compromising the toxicity results, at least in rats. In mice, with their low blood volumes, separate satellite animals would be necessary as blood donors.
However, the use of dried blood spot technology combined with high performance liquid chromatography–mass spectrometry (HPLC–MS) seems to offer a solution for these small animals. This technology permits high quality toxicokinetic information to be obtained using significantly smaller volumes of blood (typically 10-20 µL) than are traditionally required (200-250 µL) avoiding or reducing the need for satellite animals in the case of rats and mice, respectively (Creton *et al*., 2012). The small sample volume also provides a refinement because warming prior to sampling is not needed for rats, and warming time for mice can be reduced from around 10 min to 5 min. The technology also offers considerable financial advantages through the reduction in animal numbers and simpler sample collection, storage and shipping (Creton *et al*., 2009).

4.4 Conclusion

In conclusion, this overview that techniques and concepts are available to acquire relevant toxicokinetic data in a tiered way, without net costs of extra test animals. These techniques and concepts also allow integration of the toxicokinetic data into the risk assessment procedure, more than is currently done. A more in-depth analysis of these techniques and concepts is still needed, to gather their exact possibilities and limitations and form a strategy on how to integrate these techniques and concepts in the current risk assessment procedure in a cost-effective manner.

5 Future human risk assessment strategy including toxicokinetics

As explained in the introduction, the resistance against animal testing, the high costs of such tests and also the desire to further refine the protection of the human population has led to efforts to increase the 3R application and change the classical, animal testing-based approach to an approach using more modern, non-animal tools, such as QSARs and *in vitro* tests. For cosmetics, non-animal-based ("1R") risk assessment is even already regulatory reality in the EU. Adler *et al.* (2011) have concluded that toxicokinetic data are essential in such as new risk assessment approach, a conclusion repeated in the report of the SCHER, the SCENIHR and the SCCS (2013).

5.1 New risk assessment strategy

Various strategies for a new way of risk assessment have been suggested (e.g. Adler *et al*., 2011; SCHER/SCENIHR/SCCS, 2013; Thomas *et al*., 2013), at different levels of detail. Together, they show the need for and the way to incorporate the following desired ingredients:

- A shift from a hazard-based to an exposure-based hazard testing and risk assessment;
- Extension of external exposure to internal exposure-based testing;
- Animal testing only as a last resort or not at all;
- A tiered approach
- Integration of toxicokinetic data, mostly through physiologically-based toxicokinetic (PBTK) models;
- Mode of action (MOA) determinations.

A synthesis of these strategies needs to be made, taking into account pragmatic limitations of each method and costs. Some elements for consideration into such a new strategy are further discussed here.

For an exposure-based risk assessment, the assessment will most likely start with an exposure assessment. This serves to determine the routes through which humans may be expected to be exposed and estimates of external doses for these routes. This is theoretically simple, but in practice, this may turn out be almost impossible and far too expensive. Many manufacturers of chemical substances do not know the complete downstream use of their substances, and obtaining this information and to process these in complicated exposure models can become very demanding, if not impossible in case of a new substance for which not all uses are foreseen yet. Therefore, this step may turn out to remain a very, possibly too uncertain assessment.
For exposure-based waiving, the estimated exposures will need to be compared to the external TTC. To apply the right TTC, it is useful to know whether the substance is genotoxic, as a different TTC is applied for genotoxic substances (Kroes *et al*., 2004). Therefore, it would be logical to perform the common battery of *in vitro* genotoxicity tests at the start of a risk assessment, too, in parallel to the exposure assessment. Such early genotoxicity tests match with common practice in industry, as genotoxic substances are usually not developed further and therefore need to be identified as soon as possible. Likewise, an intelligent testing strategy (ITS) for sensitization could also be performed at the beginning of the risk assessment.
Exposure-based waiving could be extended for internal exposures, by determining the absorption and comparing the internal dose to an internal TTC

(e.g. Adler *et al*., 2011). Only if the internal exposure would be above this internal TTC, there would be a need for systemic kinetic and toxicity tests. Such an internal TTC still needs to be developed, but this is found to be impossible at the moment as too few absorption data are available.

In parallel to the absorption assessment for internal exposure-based waiving, the ability of the substance to cause local effects at the relevant exposure sites can be determined, as these effects occur independent of absorption.
An important consideration is to perform a metabolism study before performing the toxicity tests, so that the relevant metabolites can be included in the toxicity tests. This is especially important when the successive steps are performed *in vitro*, where no/not all metabolites will be included. The U.S. EPA OPPTS 870.7485 Guidelines for Metabolism and Pharmacokinetics and the OECD 417 Guideline indicate that metabolites in excreta accounting for 5% or more of the administered dose should be structurally characterized. (Barton *et al*., 2006; OECD, 2010), thus this can be taken as limit for including a metabolite for further risk assessment or not.
It has been pointed out that the design of the *in vitro* toxicity assays for systemic effects would benefit from prior PBPK modelling, because the determined target dose can help in the selection of the optimal *in vitro* test concentrations (Adler *et al*., 2011). The predictions of a PBPK model could even help in deciding which systemic effects are necessary to study, as effects in organs that are not reached by the substance, would not be expected to occur. Furthermore, a PBPK model of the *in vitro* system could optimize the experimental design, just like it can optimize the design of an *in vivo* test (see chapter 3). As discussed in paragraph 5.1, the PBPK model would need some minimal input of parameters. These would thus have to be determined prior to the systemics tests, too.

The obtained no observed effect concentrations (NOECs) or benchmark doses (BMDs) from *in vitro* toxicity tests can theoretically be used in two ways: for *in vitro – in vivo* comparison (IVIVC) or for *in vitro – in vivo* extrapolation (IVIVE) (Adler *et al*., 2011). For IVIVC, the target (or intrinsic) NOEC is compared to the target dose estimated by the PBTK model earlier. If the target NOEC is higher for all targets, it may concluded there no human health risk is expected from the estimated exposure. If the target NOEC is lower for one or more target doses, it may be concluded there is possible human health risk for these effects from the estimated exposure. For IVIVE, the target NOECs are extrapolated to external no observed adverse effect levels (NOAELs) using the PBTK model, which are then used to derive human limit values (HLVs) and to classify the substance.
A point to consider in these comparisons, which is not discussed in literature, is the use of assessment factors, to correct for variations and uncertainties. If all data have been determined using human cells and tissues, no uncertainty factor for interspecies differences would need to be included in this comparison and extrapolation. An uncertainty factor for intraspecies differences will be needed unless the variability in the human population is included in both the determinations of the toxicokinetics and the toxicodynamics, by e.g. the use of cells from different donors. Furthermore, it may be worthwhile to introduce a new type of uncertainty factor, namely for the use of *in vitro* models, as they are only parts of a whole body and also show experimental variability and uncertainty. Just like *in vivo* animal models have always been regarded to differ from the human body, leading to the use of the default uncertainty factor of 10, the battery of *in vitro* tests in a new strategy may be regarded to differ from the human body, too. Instead of dismissing them for this reason, a new uncertainty factor could be introduced to make their limitations acceptable.

5.2 Role toxicokinetics

It may be clear from the previous section how important toxicokinetic data are in the new risk assessment strategy: they are often included in a first or second tier of presented strategies (e.g. Adler *et al*., 2011; SCHER/SCENIHR/SCCS 2013; Thomas *et al*., 2013), are determining in internal exposure-based waiving, and are a component in the final step of assessing whether there is a risk or when deriving HLVs (by the use of the target doses or PKPK model).

Information on toxicokinetics in the body is essential under 1R to address three major issues, according to Adler *et al*., (2011):
(1) Development and design of more efficient testing strategies: As a key starting point for any toxicological testing, it is essential to know whether a substance will be bioavailable by one of the relevant uptake routes: only in cases where a substance is bioavailable following dermal, oral or inhalation exposure, further tests on systemic and not just local toxicity will be necessary. Furthermore, knowledge on the distribution of the substance and final target concentrations (e.g. through PBPK models) permit the selection of the relevant *in vitro* toxicodynamic tests and of the dose range that should be used in these tests.
(2) *In vitro–in vivo* extrapolation: To relate toxicodynamic information from non-animal-testing (1R) to real-life situation relevant for humans, i.e. to transform *in vitro* concentration–effect relationship into an *in vivo* dose–effect relationship. The most sophisticated challenge under 1R is to make *in vitro* data (from any type of toxicological endpoint) usable for risk assessment, i.e. to properly relate toxicodynamic information from *in vitro* studies to the *in vivo* situation, because test results under 1R will be presented as an *in vitro* concentration–effect relationship instead of an *in vivo* dose–effect relationship. This can be achieved using PBPK models, which need toxicokinetic data as input.
(3) Identification of clearance rates and the role of metabolites: For the *in vitro* dynamics experiments, it is essential to know whether the cell or tissues are exposed to the parent compound and/or its metabolites. To be able to use this information to optimize the toxicity test designs, this information has to be known upfront. It can be based on toxicokinetic alternative methods identifying the main metabolites and the clearance rates of the parent compound and/or its metabolites.

Additionally, knowledge of the kinetics in the *in vitro* system is also crucial to translate *in vitro* results to the (human) *in vivo* situation. Nominal (i.e. added) concentrations in *in vitro* media may greatly differ from the actual intracellular concentration due to altered bioavailability (interactions with the medium, the plate, the cell itself) or to physiological cellular processes (mechanism of transport across the membranes, biotransformation, bioaccumulation) (e.g. Heringa *et al*., 2004). Especially in case of repeated treatments for prolonged times of exposure, to mimic exposure to chemical substances, these processes create uncertainty about the actual level of exposure of cells *in vitro*. For this reason, *in vitro* kinetics should also be considered in the experimental design for and the treatment of the results from *in vitro* dynamics experiments in order to correlate *in vitro* results to *in vivo* actual situations (Adler *et al*., 2011). Measurement of the free medium concentrations (available for cell uptake) or free cellular concentrations provide a means to derive *in vitro* concentration-response relationships that are devoid of the kinetic effects mentioned above, and that can thus be directly used for extrapolation to *in vivo* dose-response relationships (e.g. Heringa *et al*., 2004; Kramer *et al*., 2012). Whether such

measurements are necessary, and which dose-parameter should then be measured depends on a.o. the physical-chemical properties of the substance (e.g. volatility) as has been reviewed by Groothuis *et al*. (2013). The latter paper also provides a decision tree for the selection of the right dose-parameter to measure.

5.3 PBPK models

Clearly, PBPK models are seen to play a crucial role in a new risk assessment strategy. They are seen as "the most adequate approach to simulate the fate of compounds in the human body" (Adler *et al*., 2011) and "likely to be an important tool by which the *in vitro* to *in vivo* interface can be improved" (SCHER/SCENIHR/SCCS, 2013). Therefore, these models are described more in detail in this paragraph.
PBPK models are mathematical models of an organism, containing its important physiological features to describe the concentration of a substance in time for any physiological location in that organism. They can therefore provide a link between the external doses (e.g. the concentration of the substance in the inhaled air) and the target dose (e.g. the concentration in the liver) at a certain time point. Currently, PBTK models are mostly used to estimate the target doses from the known external doses or to convert safe animal external doses into safe human external doses. In a new risk assessment strategy, they would not only need to predict the target doses from a known or expected exposure, but also to convert safe target doses, determined with *in vitro* assays, into external safe doses (*in vitro – in vivo* extrapolation, IVEVE). The latter has already been quite successfully performed by a handful of researchers, as reviewed by Punt *et al*. (2011): predicted LOAELs or BMD_{10}s were within factor 1-10 of the *in vivo* experimental values.

Input data for PBPK models
A challenge in the use of PBPK models is that they are very data-demanding. As input data for the parameters in a PBPK model, both organism-specific or "system" data (e.g. blood flows, organ weights, blood volume) and substance-specific data (e.g. volatility, metabolism rate, tissue partitioning coefficients) are necessary (Adler *et al*., 2011). The organism-specific data are generally well known, as the physiology of mammals has been studied extensively, especially for humans. The success of a PBPK model is therefore mostly dependent on the availability and good quality of the substance-specific data ("trash in, trash out"). In the absence of human *in vivo* ADME data, *in vivo* animal ADME data are often used in PBPK models for the specific test animal species. These are then converted to models for humans by so-called allometric scaling: a scale-up between animals and humans of such parameters as tissue volume and blood flow and by their relation to body weight. However, with the desire to replace these animal tests, the substance-specific data will have to come from *in vitro* tests or QSARs. In this line, there is commercial PBPK modelling software that is more directed to the use of human *in vitro* data, e.g. from Simcyp.
Of course, the more parameter values available, the more precise the model will become and the lower the uncertainty of its predictions. However, as explained in paragraph 4.2, not all ADME parameters can already be determined by *in vitro* or *in silico* methods. To know if and which of these methodology gaps have priority to fill, efforts are made to define the minimum data set for PBPK models. Thomas *et al*. (2013) suggest that only metabolism and protein binding are needed for a PBPK model in their first tier of risk assessment, and the volume of distribution (estimated with QSARs) and oral bioavailability (determined with and *in vitro* assay using Caco-2 cells) are added in the second tier. Bessems *et*

al. (2013) explain that a first, simple ("tier 1") PBPK model only needs to include passive processes and hepatic metabolism, as *in vitro* or *in silico* methods are not readily available for active transport determination yet, and this process also complicates the model greatly. In fact, there are *in vitro* models for active transport available already, but their extrapolation is indeed problematic. The metabolism in other tissues is also left for later stages of model development, as this is also less readily assessed and less important than the metabolism in the liver. The parameters necessary for such a first, simple PBPK model were considered to be:

- Oral, dermal and inhalation absorption
- Tissue:blood partitioning coefficients (not specified for which tissues)
- Metabolic rate
- Blood:air partitioning coefficient (for excretion through exhaled air)
- Renal excretion rate
- Protein binding

Further analysis will have to show whether these parameters are indeed sufficient and necessary, also considering the costs of their determinations.

PBPK model types

To date, PBPK models have mostly been built for each single substance separately, tailored to the specific properties of the substance. As this is too labour-intensive to perform for the large number of substances needing a risk assessment, generic PBPK models will be necessary. These are PBPK models containing all physiological features, for which the chemical-specific parameters need to be filled in to tailor the model to the substance in question. For pharmaceutical drug development, a number of generic PBPK models have been developed to predict human *in vivo* kinetics, such as GastroPlus and Simcyp (Creton *et al*., 2009). These models use *in vitro* and *in silico* (QSAR) data, on factors such as metabolism, plasma protein binding and lipophilicity (Creton *et al*., 2009). *In silico* models to estimate oral bioavailability (Adler *et al*., 2011) have also been developed for the use in conjunction with PBPK models. However, relevance and reliability of these *in silico* models outside the pharmaceutical R&D, where they were developed, will need quite some years of extra investigations before prevalidation would be reachable (Adler *et al*., 2011). Additionally, the pharmaceutical origin of these models makes them primarily targeted to the oral route of exposure. In these package of, for example, Simcyp, PK-Sim and Cloe, which are somewhat „black box" in that model structure and code is inaccessible to the user, the inhalation and dermal route, are not described in sufficient detail. It is noted that recently a mechanistic dermal absorption model has been added to Simcyp (Bessems *et al*, 2013)
Attempts to facilitate PBPK modelling in the public domain are IndusChemFate and MEGen. MEGen is a web application for the rapid construction and documentation of custom-built deterministic PBPK model code. MEGen comprises a parameter database and a model codegenerator that produces code for use in several commercial software packages and one that is freely available. IndusChemFate is a generic PBPK model, freely available as an MS Excel spreadsheet-file for the estimation of biomonitoring equivalent guidance values for chemical agents related to health based exposure rates for inhalation, oral intake and/or skin exposure. (Bessems *et al*, 2013)
A number of factors, that would help improve the accuracy of PBPK models and promote their greater application, have been identified. These include development of better statistical models and methods for characterising variability and uncertainty; improved databases on physiological parameters and their intra- and interindividual variation; and development of principles and guidance on good modelling practice, to support both modellers during model

design and risk assessors in evaluating their utility in the risk assessment (Creton *et al*., 2009). To enable probabilistic risk assessment, Monte Carlo sampling can be incorporated into the IVIVE model to account for interindividual variability and derive an oral equivalent dose that represents the 95th percentile of the population. (Thomas *et al*., 2013).

More pragmatically, the generic models will need to be critically analysed for their practical value and limitations in a risk assessment strategy (additional to the scientific evaluation in Bessems *et al.*, 2013). Ultimately, it is desirable to obtain a generic PBPK model for all exposure routes, validated for the whole domain of chemical substances. Furthermore, it needs to be further determined which parameters are necessary to determine for such a model. The costs of determining these parameters and the use of the PBPK model will need to be considered in order to promote their application.

5.4 Further needs

Tools that are still lacking, are incomplete, or still need major further development, as e.g. discussed in paragraph 4.2,are reported to be:
- *In vitro* absorption tests: mostly for absorption in the lung;
- An internal TTC: the internal dose below which no chemical is expected to cause a health risk
- *In vitro* barrier tests, mostly for the blood-brain barrier and the blood-testes barrier
- *In vitro* excretion tests

Final validation and standardization into OECD guidelines is also still necessary for all toxicokinetic methods presented, except for *in vitro* dermal absorption (Guideline no. 428).

The derivation of an internal TTC is found to be impossible at the moment, as the currently available data are too sparse (Adler et al., 2011). This shows the need for a database with absorption and toxicodynamic data. Similarly, Bessems *et al*. (2013) conclude that a database with animal and human chemical-specific kinetic data is necessary to enhance the use of non-animal PBPK models.
The report of SCHER, SCENIHR and SCCS (2013) even concludes that five categories of databases are needed for the new risk assessment approach:
1. Human exposure data;
2. Human effects data from exposure to individual chemicals;
3. Data on the adverse effects of chemicals in animal models;
4. *In vitro* findings;
5. Metabolism and other kinetic data.

Bessems *et al*. (2013) further recommend the generation of freely available, easy to use software to create PBPK models and their standardization through guidelines. Now, software programs are either commercially available or only applicable to a limited set of substances. For example, IndusChemFate is freely available from the Cefic-LRI website, but is only applicable to volatile and semi-volatile chemicals.

A synthesis needs to be made of the already presented suggestions for a new risk assessment strategy, with a critical evaluation of the value and limitations of each approach. A consideration of the costs will need to be included, as well as a way to handle uncertainties. This exercise will also make clear what will be the minimum set of necessary toxicokinetic parameters. The strategies (paragraph 5.2) and the list of toxicokinetic parameters considered necessary for a first

PBPK model (paragraph 5.1) already show selections towards a minimum set of parameters necessary to perform a risk assessment according to the new approach. However, not all details are clear yet, such as for which tissues the partitioning coefficients need to be determined. Also, it is not defined yet with which test the desired parameters should be determined exactly, from the range of different *in vitro* and also *in silico* methods sometimes available, and the costs are not considered. Therefore, this base set of parameters and corresponding tests need to be further defined.

6 Conclusions and outlook

In 2012, an overview of available *in vitro* methods for determination of kinetic parameters was prepared (Brandon *et al*., 2012). The current report builds further upon that overview by including *in silico* methods and describing how toxicokinetic data can be beneficial or even crucial in terms of 3R and improved human safety assessment.
In the current system of chemical safety assessment such data can be beneficial in the following ways:
Kinetics data can avoid unnecessary animal tests (→ Reduction);
Kinetics data can improve study design (selection test species, exposure route, dose, and analyzed organs → Reduction and Refinement);
Kinetics data can improve extrapolation from animal to human (→ improved human safety).
Methods by which these data can be obtained without losing the benefit of animal reduction, are human microdosing, *in vitro* tests, inclusion of blood and excreta sampling in the first *in vivo* toxicodynamic test performed, or a separate *in vivo* toxicokinetic test with a minimal number of animals (3-9). Important is that these toxicokinetic data are obtained in an early phase, in order to enable to use of the data in the design of further toxicodynamic tests, i.e. to integrate toxicokinetics in the further risk assessment.
In a new risk assessment strategy, completely or largely replacing animal tests, the toxicokinetics clearly play a crucial role, e.g. to be able to extrapolate the safe concentrations determined *in vitro* to safe external doses *in vivo* for humans. PBTK models are a centrepiece in such an extrapolation, which should be fed the necessary toxicokinetic data by *in vitro* kinetic tests or *in silico* predictions of kinetic parameter values. Some of the necessary toxicokinetic methods and tools for the new risk assessment strategy are already available, but others still need major further development.
In 2014, it is planned to follow up this work by critically evaluating the different tools described in terms of added value, limitations and costs. Simultaneously, it is planned to develop strategies, both for integration of toxicokinetic data into the current system of risk assessment, and for a new risk assessment system. Costs and uncertainty analysis will be important elements for consideration in the formation of these strategies. Together, these exercises are expected to lead to the formulation of a base set of necessary toxicokinetic parameters and the methods to determine these.

7 References

Adler, S., Basketter, D., Creton, S., Pelkonen, O., van Benthem, J., Zuang, V., Andersen, K. E., Angers-Loustau, A., Aptula, A., Bal-Price, A., Benfenati, E., Bernauer, U., Bessems, J., Bois, F. Y., Boobis, A., Brandon, E., Bremer, S., Broschard, T., Casati, S., Coecke, S., Corvi, R., Cronin, M., Daston, G., Dekant, W., Felter, S., Grignard, E., Gundert-Remy, U., Heinonen, T., Kimber, I., Kleinjans, J., Komulainen, H., Kreiling, R., Kreysa, J., Leite, S. B., Loizou, G., Maxwell, G., Mazzatorta, P., Munn, S., Pfuhler, S., Phrakonkham, P., Piersma, A., Poth, A., Prieto, P., Repetto, G., Rogiers, V., Schoeters, G., Schwarz, M., Serafimova, R., Tahti, H., Testai, E., van Delft, J., van Loveren, H., Vinken, M., Worth, A., and Zaldivar, J. M. (2011) Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. *Archives of toxicology* **85**(5), 367-485.

Artola-Garicano, E., Vaes, W. H. J., and Hermens, J. L. M. (2000) Validation of Negligible Depletion Solid-Phase Microextraction as a Tool to Determine Tissue/Blood Partition Coefficients for Semivolatile and Nonvolatile Organic Chemicals. *Toxicology and Applied Pharmacology* **166**(2), 138-14.

Barter, Z. E., Bayliss, M. K., Beaune, P. H., Boobis, A. R., Carlile, D. J., Edwards, R. J., Houston, J. B., Lake, B. G., Lipscomb, J. C., Pelkonen, O. R., Tucker, G. T., and Rostami-Hodjegan, A. (2007) Scaling factors for the extrapolation of in vivo metabolic drug clearance from in vitro data: Reaching a consensus on values of human microsomal protein and hepatocellularity per gram of liver. *Current Drug Metabolism* **8**(1), 33-45.

Barton, H. A., Pastoor, T. P., Baetcke, K., Chambers, J. E., Diliberto, J., Doerrer, N. G., Driver, J. H., Hastings, C. E., Iyengar, S., Krieger, R., Stahl, B., and Timchalk, C. (2006) The acquisition and application of absorption, distribution, metabolism, and excretion (ADME) data in agricultural chemical safety assessments. *Critical reviews in toxicology* **36**(1), 9-35.

Basketter, D. A., Clewell, H., Kimber, I., Rossi, A., Blaauboer, B., Burrier, R., Daneshian, M., Eskes, C., Goldberg, A., Hasiwa, N., Hoffmann, S., Jaworska, J., Knudsen, T. B., Landsiedel, R., Leist, M., Locke, P., Maxwell, G., McKim, J., McVey, E. A., Ouédraogo, G., Patlewicz, G., Pelkonen, O., Roggen, E., Rovida, C., Ruhdel, I., Schwarz, M., Schepky, A., Schoeters, G., Skinner, N., Trentz, K., Turner, M., Vanparys, P., Yager, J., Zurlo, J., and Hartung, T. (2012). A roadmap for the development of alternative (non-animal) methods for systemic toxicity testing - t4 report. *ALTEX* **29**(1), 3-91.

Bessems, J.G.M. and Geraets, L. (2013) Proper knowledge on toxicokinetics improves human hazard testing and subsequent health risk characterisation. A case study approach. *Regulatory Toxicology and Pharmacology* **67,** 325–334

Bessems, J.G. Loizou, G., Krishnan, K., Clewel, H.J., Bernasconi, C., Bois, F., Coecke, S., Collnot, E.-M., Diembeck, W., Farcal, L.R., Geraets, L., Gundert-Remy, U., Kramer, N., a,⇑ Küsters, G., Leite, S.B., Pelkonen, O.R., Schröder, K., Testai, E., Wilk-Zasadna, I. and Zaldívar-Comenges, J.-M. (2013) PBTK modelling platforms and parameter estimation tools to enable animal-free risk assessment – Recommendations from a joint EPAA – EURL ECVAM ADME

workshop. *Regulatory Toxicology and Pharmacology*, doi: 10.1016/j.yrtph.2013.11.008.

Brandon, E.F.A., Bessems,J.G.M., and Bos, P.J.M. (2012) Overview on *in vitro* models to measure absorption, distribution, metabolism and excretion for use in testing strategies and human risk assessment. Letter report RIVM.

Brown, R.P., Delp, M.D., Lindstedt, S.L., et al. (1997) Physiological parameter values for physiologically based pharmacokinetic models. *Toxicology and Industrial Health* **13**, 407–484.

Burtin, P., Mentre, F., van Bree, J., and Steimer, J. L. (1996) Sparse sampling for assessment of drug exposure in toxicological studies. *European Journal of Drug Metabolism and Pharmacokinetics* **21**(2), 105-11.

Carmichael, N. G., Barton, H. A., Boobis, A. R., Cooper, R. L., Dellarco, V. L., Doerrer, N. G., Fenner-Crisp, P. A., Doe, J. E., Lamb, J. C. t., and Pastoor, T. P. (2006) Agricultural chemical safety assessment: A multisector approach to the modernization of human safety requirements. *Critical reviews in toxicology* **36**(1), 1-7.

Cooney, D.O. (1976) Biomedical Engineering Principles. Marcel Dekker.

Cooper, R. L., Lamb, J. C., Barlow, S. M., Bentley, K., Brady, A. M., Doerrer, N. G., Eisenbrandt, D. L., Fenner-Crisp, P. A., Hines, R. N., Irvine, L. F., Kimmel, C. A., Koeter, H., Li, A. A., Makris, S. L., Sheets, L. P., Speijers, G., and Whitby, K. E. (2006) A tiered approach to life stages testing for agricultural chemical safety assessment. *Critical Reviews in Toxicology* **36**(1), 69-98.

Creton, S., Billington, R., Davies, W., Dent, M. P., Hawksworth, G. M., Parry, S., and Travis, K. Z. (2009) Application of toxicokinetics to improve chemical risk assessment: implications for the use of animals. *Regulatory Toxicology and Pharmacology* **55**(3), 291-9.

Creton, S., Saghir, S. A., Bartels, M. J., Billington, R., Bus, J. S., Davies, W., Dent, M. P., Hawksworth, G. M., Parry, S., and Travis, K. Z. (2012) Use of toxicokinetics to support chemical evaluation: Informing high dose selection and study interpretation. *Regulatory Toxicology and Pharmacology* **62**(2), 241-7.

Doktorova, T. Y., Pauwels, M., Vinken, M., Vanhaecke, T., and Rogiers, V. (2012) Opportunities for an alternative integrating testing strategy for carcinogen hazard assessment? *Critical Reviews in Toxicology* **42**(2), 91-10.

EMEA (European Medicines Agency) (1995) ICH Topic S3A: Toxicokinetics: A guidance for assessing systemic exposure in toxicology studies. Step 5. PMP/ICH/384/95.

Groothuis, F. A., Heringa, M. B., Nicol, B., Hermens, J. L., Blaauboer, B. J., and Kramer, N. I. (2013) Dose metric considerations in in vitro assays to improve quantitative in vitro-in vivo dose extrapolations. *Toxicology*, doi: 10.1016/j.tox.2013.08.012.

Heringa, M. B., Schreurs, R. H. M. M., Busser, F., Van Der Saag, P. T., Van Der Burg, B., and Hermens, J. L. M. (2004) Toward more useful in vitro toxicity data

with measured free concentrations. *Environmental Science and Technology* **38**(23), 6263-6270.

Jochemsen, R., Bazot, D., Brillanceau, M. H., and Lupart, M. (1993) Assessment of drug exposure in rat dietary studies. *Xenobiotica* **23**(10), 1145-1154.

Kramer, N. I., Krismartina, M., Rico-Rico, Á., Blaauboer, B. J., and Hermens, J. L. M. (2012) Quantifying processes determining the free concentration of phenanthrene in basal cytotoxicity assays. *Chemical Research in Toxicology* **25**(2), 436-445.

Kroes, R., Renwick, A. G., Cheeseman, M., Kleiner, J., Mangelsdorf, I., Piersma, A., Schilter, B., Schlatter, J., van Schothorst, F., Vos, J. G., and Wurtzen, G. (2004) Structure-based thresholds of toxicological concern (TTC): guidance for application to substances present at low levels in the diet. *Food and Chemical Toxicology* **42**(1), 65-83.

Niesink, R.J.M., de Vries, J., Hollinger, M.A. (Eds.) (1996) Toxicology. Principles and applications. CRC Press, Boca Raton, U.S.A.

OECD (Organisation for Economic Co-operation and Development) (2010) OECD Guidelines for the Testing of Chemicals, Section 4: Health Effects. Test No. 417: Toxicokinetics.

Oravcova, J., Bohs, B., and Lindner, W. (1996) Drug-protein binding sites. New trends in analytical and experimental methodology. *Journal of chromatography. B, Biomedical applications* **677**(1), 1-28.

Punt, A., Schiffelers, M. J., Jean Horbach, G., van de Sandt, J. J., Groothuis, G. M., Rietjens, I. M., and Blaauboer, B. J. (2011) Evaluation of research activities and research needs to increase the impact and applicability of alternative testing strategies in risk assessment practice. *Regulatory Toxicology and Pharmacology* **61**(1), 105-14.

Ramsey, J.C., Andersen, M.E. (1984) A physiologically based description of the inhalation pharmacokinetics of styrene in rats and humans. *Toxicology and Applied Pharmacology*
73 (1), 159-175.

Rorije, E., Aldenberg, T., Buist, H., Kroese, D., and Schuurmann, G. (2013) The OSIRIS Weight of Evidence approach: ITS for skin sensitisation. *Regulatory Toxicology and Pharmacology*, doi: 10.1016/j.yrtph.2013.06.003.

Saghir, S. A., Mendrala, A. L., Bartels, M. J., Day, S. J., Hansen, S. C., Sushynski, J. M., and Bus, J. S. (2006) Strategies to assess systemic exposure of chemicals in subchronic/chronic diet and drinking water studies. *Toxicology and Applied Pharmacoogyl* **211**(3), 245-60.

SCHER (Scientific Committee on Health and Environmental Risks), SCENIHR (Scientific Committee on Emerging and Newly Identified Health Risks), SCCS (Scientific Committee on Consumer Safety) (2013) Addressing the new challenges for risk assessment. Report adopted March 2013.

Thomas, R. S., Philbert, M. A., Auerbach, S. S., Wetmore, B. A., Devito, M. J., Cote, I., Rowlands, J. C., Whelan, M. P., Hays, S. M., Andersen, M. E., Meek, M.

E., Reiter, L. W., Lambert, J. C., Clewell, H. J., 3rd, Stephens, M. L., Zhao, Q. J., Wesselkamper, S. C., Flowers, L., Carney, E. W., Pastoor, T. P., Petersen, D. D., Yauk, C. L., and Nong, A. (2013) Incorporating new technologies into toxicity testing and risk assessment: moving from 21st century vision to a data-driven framework. *Toxicological Sciences* **136**(1), 4-18.

Tice, R. R., Austin, C. P., Kavlock, R. J., and Bucher, J. R. (2013). Improving the human hazard characterization of chemicals: a Tox21 update. *Environmental Health Perspectives* **121**(7), 756-65.

Timchalk, C. (2004) Comparative inter-species pharmacokinetics of phenoxyacetic acid herbicides and related organic acids. evidence that the dog is not a relevant species for evaluation of human health risk. *Toxicology* **200**(1), 1-19.

US-NAS (2007) Toxicity testing in the 21st century. A vision and a strategy. Committee on Toxicity Testing and Assessment of Environmental Agents, Board on Environmental Studies and Toxicology, Institute for Laboratory Animal Research, Division on Earth and Life Studies. National Research Council, The National Academic Press, Washington D.C., USA. <http://www.nap.edu/catalog/11970.html>.

Vaes, W. H. J., Urrestarazu Ramos, E., Verhaar, H. J. M., Seinen, W., and Hermens, J. L. M. (1996) Measurement of the free concentration using solid-phase microextraction: binding to protein. *Analytical Chemistry* **68**(24), 4463-4467.

Vermeire, T., Aldenberg, T., Buist, H., Escher, S., Mangelsdorf, I., Pauné, E., Rorije, E., and Kroese, D. (2013) OSIRIS, a quest for proof of principle for integrated testing strategies of chemicals for four human health endpoints. *Regulatory Toxicology and Pharmacology* **67**(0), 136-145, doi: 10.1016/j.yrtph.2013.01.007.

Wetmore, B. A., Wambaugh, J. F., Ferguson, S. S., Sochaski, M. A., Rotroff, D. M., Freeman, K., Clewell, H. J., Dix, D. J., Andersen, M. E., Houck, K. A., Allen, B., Judson, R. S., Singh, R., Kavlock, R. J., Richard, A. M., and Thomas, R. S. (2012) Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. *Toxicological Sciences* **125**(1), 157-174.

Wilson, A. G., Frantz, S. W., and Keifer, L. C. (1994) A tiered approach to pharmacokinetic studies. *Environmental Health Perspectives* **102 Suppl 11**, 5-11.

RIVM
De zorg voor morgen begint vandaag