International Ph.D. Program in Environmental Sciences and Engineering

Development of alternative methods for toxicological risk assessment based on human Mesenchymal Stem Cells in the framework of REACH

Advisors:
Prof. Giacomo Cao
Dr. Luisa Mancuso

Ph.D. Candidate:
Michela Scanu

(Academic Year 2010/2011)
## CONTENTS

<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of abbreviations</td>
<td>4</td>
</tr>
<tr>
<td>Thesis Summary</td>
<td>6</td>
</tr>
</tbody>
</table>

### Chapter 1 - Overview of the Toxicological Risk assessment and Toxicity Testing  

1.1. Introduction  
1.2. Animal methods of toxicity testing  
1.2.1. A brief historical overview  
1.2.2. Aims and objectives of animal studies  
1.3. Criticism to the use of animal methods  
1.3.1. Reasons for the criticism  
1.3.1.1. Scientific reasons  
1.3.1.2. Economic reasons  
1.3.1.3. Ethical reasons  
1.3.2. The 3RS principle  
1.4. Alternative methods  
1.4.1. In vitro tests for cytotoxicity  
1.4.1.1. General advantages and disadvantages  
1.4.1.2. Methods  
1.4.2. Target organ toxicity  
1.4.2.1. General advantages and disadvantages  
1.4.2.2. Methods  
1.4.3. Use of stem cells in toxicology  
1.5. Validation, regulatory acceptance and international organization  
1.5.1. Validation authorities  
1.5.1.1. ECVAM  
1.5.1.2. ICCVAM  
1.5.1.3. OECD  
1.5.1.4. Others
1.5.2. The Validation Process ..................................................................................... 41
  1.5.2.1. Definition ................................................................................................ 42
  1.5.2.2. Steps of the process ................................................................................ 43
1.5.3. Regulatory Acceptance ..................................................................................... 46
1.5.4. Validated and accepted alternative methods .................................................... 47
1.6. The legislative implementation of the 3RS principle ................................................. 54
1.7. Numbers of animals used ........................................................................................... 56

Chapter 2 - REACH, a new chemicals policy for the EU ................................................. 60
  2.1. Introduction ................................................................................................................ 60
  2.2. Scope: REACH and chemicals ................................................................................... 62
  2.3. How does reach work: REACH processes ................................................................. 63
    2.3.1. Pre-Registration ................................................................................................ 63
    2.3.2. Registration ....................................................................................................... 64
    2.3.3. Evaluation ......................................................................................................... 64
    2.3.4. Authorisation .................................................................................................... 65
    2.3.5. Restriction ......................................................................................................... 66
    2.3.6. Classification and labelling inventory .............................................................. 66
    2.3.7. Communication in the supply chain .................................................................... 67
  2.4. Timeline for REACH implementation ....................................................................... 68
  2.5. Impact assessment ...................................................................................................... 69
    2.5.1. Benefits ............................................................................................................. 69
    2.5.2. Costs ................................................................................................................. 71
  2.6. REACH and animal testing ........................................................................................ 75

Chapter 3 - Evaluation of the use of human Mesenchymal Stem Cells for acute toxicity tests ........................................................................................................................................ 78
  3.1. Introduction ................................................................................................................ 78
  3.2. Material and methods ................................................................................................. 80
    3.2.1. Isolation of Mesenchymal Stem Cells .............................................................. 80
    3.2.2. Phenotypic characterization of stem cell cultures ............................................. 80
    3.2.3. Determination of doubling time ....................................................................... 81
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>αMEM</td>
<td>Minimum Essential Medium</td>
</tr>
<tr>
<td>3T3</td>
<td>mouse fibroblast (BALB/c) 3T3 cells</td>
</tr>
<tr>
<td>ADME</td>
<td>Absorption, Distribution, Metabolism and Excretion</td>
</tr>
<tr>
<td>ALAT</td>
<td>Alanine aminotransferase</td>
</tr>
<tr>
<td>ASAT</td>
<td>Aspartate aminotransferase</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>CAAT</td>
<td>The Johns Hopkins Center for Alternatives to Animal Testing</td>
</tr>
<tr>
<td>CD</td>
<td>Cluster of Differentiation</td>
</tr>
<tr>
<td>CDI</td>
<td>Cellular Dynamics International</td>
</tr>
<tr>
<td>CMR</td>
<td>Carcinogenic, Mutagenic or Toxic for Reproduction substances</td>
</tr>
<tr>
<td>CSR</td>
<td>Chemical Safety Report</td>
</tr>
<tr>
<td>CTRL</td>
<td>Control sample</td>
</tr>
<tr>
<td>CYP</td>
<td>cytochrome P450</td>
</tr>
<tr>
<td>EC&lt;sub&gt;50&lt;/sub&gt;</td>
<td>Effective Concentration for 50% of the cells</td>
</tr>
<tr>
<td>ECHA</td>
<td>European Chemicals Agency</td>
</tr>
<tr>
<td>ECM</td>
<td>Extracellular matrix</td>
</tr>
<tr>
<td>ECOPA</td>
<td>European Consensus Platform on Alternatives</td>
</tr>
<tr>
<td>ECVAM</td>
<td>European Centre for the Validation of Alternatives to Animal Testing</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediamine tetraacetic acid</td>
</tr>
<tr>
<td>EINECS</td>
<td>European Inventory of Existing Commercial Chemical Substances</td>
</tr>
<tr>
<td>ELINCS</td>
<td>European List of Notified Chemical Substances</td>
</tr>
<tr>
<td>ESCs</td>
<td>Embryonic Stem Cells</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence-Activated Cell Sorting</td>
</tr>
<tr>
<td>FBS</td>
<td>Fetal Bovine Serum</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and Drug Administration</td>
</tr>
<tr>
<td>FRAME</td>
<td>Fund for the Replacement of Animals in Medical Experiments</td>
</tr>
<tr>
<td>GCCP</td>
<td>Good Cell Culture Practice</td>
</tr>
<tr>
<td>GHS</td>
<td>Globally Harmonized System of Classification and Labelling of Chemicals</td>
</tr>
<tr>
<td>GLP</td>
<td>Good Laboratory Practice</td>
</tr>
<tr>
<td>hMSCs</td>
<td>human Mesenchymal Stem Cells</td>
</tr>
<tr>
<td>IC&lt;sub&gt;50&lt;/sub&gt;</td>
<td>Inhibitory Concentration 50%</td>
</tr>
<tr>
<td>ICCVAM</td>
<td>Interagency Coordinating Committee on the Validation of Alternative Methods</td>
</tr>
<tr>
<td>Acronym</td>
<td>Term</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>IOM</td>
<td>Institute of Medicine</td>
</tr>
<tr>
<td>iPS</td>
<td>induced Pluripotent Stem cells</td>
</tr>
<tr>
<td>JaCVAM</td>
<td>Japanese Centre for the Validation of Alternative Methods</td>
</tr>
<tr>
<td>JRC</td>
<td>EU Joint Research</td>
</tr>
<tr>
<td>KoCVAM</td>
<td>Korean Center for the Validation of Alternative Methods</td>
</tr>
<tr>
<td>LD&lt;sub&gt;50&lt;/sub&gt;</td>
<td>Lethal Dose 50%</td>
</tr>
<tr>
<td>LDH</td>
<td>Lactate dehydrogenase</td>
</tr>
<tr>
<td>MDBK</td>
<td>Madin-Darby bovine kidney</td>
</tr>
<tr>
<td>mEST</td>
<td>mouse Embryonic Stem Cell Test</td>
</tr>
<tr>
<td>MSCs</td>
<td>Mesenchymal Stem Cells</td>
</tr>
<tr>
<td>MSDS</td>
<td>Material Safety Data Sheet</td>
</tr>
<tr>
<td>MTT</td>
<td>3-(4, 5-dimethylthiazol–2-yl)-2,5-DIphenyltetrazolium bromide</td>
</tr>
<tr>
<td>NCA</td>
<td>Netherlands Centre for Alternatives to Animal Use</td>
</tr>
<tr>
<td>NCI</td>
<td>U.S. National Cancer Institute</td>
</tr>
<tr>
<td>NHK</td>
<td>Normal Human Keratinocytes</td>
</tr>
<tr>
<td>NICEATM</td>
<td>NTP Interagency Center for the Evaluation of Alternative Toxicological Methods</td>
</tr>
<tr>
<td>NIFDS</td>
<td>National Institute of Food and Drug Safety</td>
</tr>
<tr>
<td>NIH</td>
<td>National Institute of Health</td>
</tr>
<tr>
<td>NIH-S</td>
<td>Japan's National Institute of Health Sciences</td>
</tr>
<tr>
<td>NK</td>
<td>Natural Killer</td>
</tr>
<tr>
<td>NOAEL</td>
<td>No Observable Adverse Effect Level</td>
</tr>
<tr>
<td>NR</td>
<td>Neutral Red</td>
</tr>
<tr>
<td>NRU</td>
<td>Neutral Red Uptake</td>
</tr>
<tr>
<td>NTP</td>
<td>National Toxicology Program</td>
</tr>
<tr>
<td>OECD</td>
<td>Organization for Economic Cooperation and Development</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate Buffered Saline</td>
</tr>
<tr>
<td>PSCs</td>
<td>Pluripotent Stem Cells</td>
</tr>
<tr>
<td>QSAR</td>
<td>Quantitative Structure-Activity Relationships</td>
</tr>
<tr>
<td>R&amp;D</td>
<td>Research and Development</td>
</tr>
<tr>
<td>RC</td>
<td>Registry of Cytotoxicity</td>
</tr>
<tr>
<td>REACH</td>
<td>Registration, Evaluation, and Authorization of Chemicals</td>
</tr>
<tr>
<td>SC4SM</td>
<td>consortium Stem Cells for Safer Medicine</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>ZEBET</td>
<td>German National Centre for the Documentation and Evaluation of Alternative Methods to Animal Experiments</td>
</tr>
</tbody>
</table>
Thesis Summary

Today’s modern society is becoming constantly more dependent on the use of chemicals (ranging from industrial or pharmaceutical products to pesticides and to personal care and household products) (Blaauboer, 2002a), which are a vital part of our daily life, and provide the society with a wide range of substantially high appreciated benefits, such as increased agricultural and industrial productivity. However, as a result of their manufacture, use or disposal, chemicals also have the potential to damage the environment. Moreover, humans are exposed to a multitude of chemicals on a daily basis, both through their interaction with the environment and their specific use (Holmes, et al., 2010). Poisoning due to chemicals is generally recognized as a severe health problem and a very important public health issue. According to the Institute of Medicine (IOM), indeed, more than four million poisoning episodes occur annually in the United States (ICCVAM, 2006a; IOM, 2004), and poisonings result to be the second leading cause of injury-related deaths, exceeded only by car accidents (Birnbaum and Stokes, 2010).

In order to ensure an adequate protection of human health and the environment, it is essential to evaluate if such products (chemicals, physical and biological agents) might cause adverse effects under normal and reasonably foreseeable use (Holmes, et al., 2010), and also to guarantee that reliable quality and adequate information on the adverse effects of exposure to such chemicals is provided (Blaauboer, 2002a). This process is named "risk assessment", and its principles are discussed in Chapter 1 in more details.

Toxicological testing, which provides such information, is therefore an important part of the regulatory safety assessment for chemicals worldwide (Holmes, et al., 2010). Currently, test methods used to inform risk assessments on chemicals rely largely on animal models developed over the last 50 - 60 years. The latest available statistics on the use of animals in scientific procedures (European Commission, 2010a) show that in 2008 the number of animals employed in experimental and other scientific purposes, in all the 27 Member States of the European Union (EU), was just above twelve millions. Although the majority of these procedures were related to pharmaceutical development and efficacy studies, a significant proportion were associated with other toxicological purposes, including the safety assessment
of chemicals (Holmes, *et al.*, 2010). However, the validity of the data obtained in these *in vivo* tests for the purpose of predicting the effects occurring in humans has been often questioned (Balls, 1991; Basketter, *et al.*, 1997; Blaauboer, 2002a, 2002b; BUAV, 2001; Creton, *et al.*, 2010; Frazier and Goldberg, 1990; Langley, 2005.; Worth and Balls, 2002; York, *et al.*, 1996; York and Steiling, 1998). The reason for these criticisms is referred to the fact that any use of animals as a surrogate for humans involves a number of assumptions and extrapolations. In some instances, test animals are often exposed to higher doses than would be expected for typical human exposures, requiring specific assumptions related to effects at lower doses. Moreover, a wide range of variation in responses due to species differences was observed, and therefore it is virtually impossible to know whether results from rodents, for example, will provide an accurate prediction in humans. These scientific limitations, together with economic and ethical concerns (all of them analyzed more in details in Chapter 1) and legislative changes (among the others the very recent and of primary importance REACH Regulation, discussed in Chapter 2), are driving toxicologists to explore the potential of non-animal alternative approaches to assess toxicity in the pharmaceutical and chemical industries (Holmes, *et al.*, 2010). Recently indeed, many *in vitro* methods for toxicity testing have been developed as alternatives to whole animal tests, according to the 3Rs approach – Refinement, Reduction, and Replacement – started in 1959 by Russell and Burch (1959) (discussed more deeply in Chapter 1). Clearly, especially replacement and/or reduction of unnecessary *in vivo* tests, that today have reached a high priority level and have been heavily promoted (Hartung, 2011), would have significant animal welfare benefits and would also result in lower testing costs and time (Creton, *et al.*, 2010; Ukelis, *et al.*, 2008).

This research falls within the framework of the risk assessment of industrial/commercial chemicals, and specifically in the framework of the modern REACH Regulation, giving specific relevance to the 3Rs concept proposed by Russell and Burch (1959). The activity performed during the PhD, indeed, concerns the development of novel *in vitro* models in order to test the potential toxicity of drugs and chemicals, following two different lines of research, both of them based for the first time on the use of human Mesenchymal Stem Cells (MSCs) isolated from bone marrow. From our best knowledge indeed, both animal and human MSCs have never been adopted for developing *in vitro* model systems for acute toxicity tests, but their unique proprieties, such as unlimited proliferation ability, plasticity to generate others cell types and even being a more readily available sources of human cells,
clearly identify their potential benefits in toxicology (Davila, et al., 2004; Wobus and Loser, 2011). Properties of MSCs and their primary importance in toxicology are discussed in Chapter 1.

The first line of research, which resulted in a peer-review publication on *In vitro Toxicology* journal (Chapter 3), focused on the application of a selected *in vitro* toxicity method, the Neutral Red Uptake (NRU) assay, to detect chemical toxicity on human MSCs. Specifically, the aim of this study was to evaluate the applicability of human MSCs as cell line for *in vitro* cytotoxicity tests to correctly predict LD$_{50}$ and the hazard category of chemicals according to the Globally Harmonized System of Classification and Labelling of Chemicals – GHS (ONU, 2009). By comparing the behavior of human MSCs with the already validated 3T3 and NHK NRU test methods (validated by ICCVAM in 2006) (ICCVAM, 2006a, 2006b), the results clearly show that the tested cells can be confidently used to perform *in vitro* acute toxicity tests.

The second line of research, on the contrary, was mainly focused on the research of new *in vitro* methods in order to test, specifically, the vascular toxicity of chemicals. This research is based on the application of emerging tissue engineering concepts on the field of toxicology. For the moment, only preliminary studies were conducted, which regarded the application of decellularization techniques to blood vessels harvested from animals at the local abattoir, in order to obtain biological scaffolds. These scaffolds could be subsequently used for recellularization processes using human MSCs, and setting up of perfused *in vitro* human toxicity tests. Of course this goal might be reached only if the animal vessel to be used as scaffold is completely acellular. The study presented in this thesis focused, therefore, on the histological evaluation of the different decellularization protocols performed and the obtained results allow one to conclude that an optimized effective protocol has been achieved using a combination of two different chemicals (ionic and non ionic detergents) and a physical treatment.
Chapter 1
Overview of the Toxicological Risk assessment and Toxicity Testing

1.1. Introduction

As already introduced, modern society is increasingly depending on the use of chemicals produced for a wide variety of purposes, such as pharmaceuticals, personal care products, biocides, and so on. Moreover, their production and use may lead to many different adverse effects, both for humans and the environment. Recently, a systematic study of these risks has evolved in the direction of evaluating the balance between the benefits and the risk of adverse effects related to the use of chemicals, thus leading, in many cases, to the relatively safe use of chemicals in different sectors (Blaauboer, 2002a, 2002b). The process of assessment that aims to provide a scientific description of the risks of chemical exposures (Frazier and Goldberg, 1990), thus evaluating their potential impact on human health and welfare (Zurlo, et al., 1994), is named "risk assessment".

Such process consists of both qualitative information on the nature of the outcome and quantitative assessment of the chemical hazard, exposure and magnitude of risks, and it combines predictions of toxic hazard with evaluations of likely exposure under specified conditions (Balls and Fentem, 1992). Specifically, the procedures used for this purpose are deeply rooted in the so-called "risk-assessment paradigm", which consists of different well-defined elements: hazard identification, hazard characterisation, exposure assessment and risk characterisation (Blaauboer, 2002a, 2002b; Kroes and Feron, 1990). The hazard identification process involves the study of a chemical’s ability to cause adverse effects, and provide a quantitative description of the nature of these effects. The hazard characterisation, instead, provide a semi-quantitative evaluation of the nature of adverse effects (including, where possible, a dose-response assessment) (Blaauboer, 2002b; Holme and Dybing, 2002). Exposure assessment is again a semi-quantitative evaluation, but this time of the potential exposure to a chemical, both for human and the environment (Blaauboer, 2002b; Frazier and Goldberg, 1990). Information obtained from these three elements is then combined in a risk
characterization, which is a semi-quantitative estimation of the probability of occurrence and severity of adverse effects in a given population under defined exposure conditions (Blaauboer, 2002b). All these elements of the risk-assessment paradigm include a variety of experimental activities to be performed. While hazard identification and characterisation (in only one word hazard assessment) often depend on the use of animal experiments, exposure assessment is generally the result of chemical analysis, even if it might also depend on biomonitoring in animals or humans, and on computer-based estimations of exposure levels (Worth and Balls, 2002).

In general, it is possible to say that toxicological evaluations currently form the basis of the assessment of risks from existing and new chemicals. Such evaluations rely mainly on the basis of tests conducted on laboratory animals, but these methods has been heavily criticised for many years, both for ethical, scientific and financial reasons (Balls, 1991; Basketter, et al., 1997; Blaauboer, 2002a, 2002b; BUAV, 2001; Creton, et al., 2010; Frazier and Goldberg, 1990; Langley, 2005; Worth and Balls, 2002; York, et al., 1996; York and Steiling, 1998). The latter ones, together with legislative changes (REACH Regulation as last), are driving toxicologists to explore the potential of non-animal alternative approaches to assess toxicity in the pharmaceutical and chemical industries (Holmes, et al., 2010).

### 1.2. Animal methods of toxicity testing

Traditionally, as previously said, the toxic potential of industrial chemicals and household products was assessed by using animal models, and mammals (predominantly rodents) are often used as experimental animal species (Holmes, et al., 2010; Huggett, et al., 1996).

#### 1.2.1. A brief historical overview

The use of animals in scientific research has a surprisingly long history, dating back to the 4th century BC (Before Christ). But it was from the 17th century, and even more during the 18th and 19th ones, that the use of animals in experimentation slowly became more common and accepted. Since then, indeed, animals have been repeatedly used through the history of
biomedical research, and it has to be noticed that many advances in this field could not have been made without the use of animals in some way, leading to enormous benefit to humanity (Cohen and Loew, 1984). With specific regard to toxicology testing, they became important in the 20th century, following the birth of the synthetic chemical industry in the late 1800s, when the need to understand how all the new substances might affect the health of workers and consumers involved in their production and use became more important. The field of toxicology, therefore, grew in response to these factors, and the use of whole animals was the most logical choice, because very few alternatives were available at that time.

The use of living animals to study the potential adverse effects of new drugs, food additives, pesticides, and other substances effectively began during the 1920s, when a British pharmacologist, J.W. Trevan, proposed the so-called "lethal dose fifty percent" or "LD50" test (Trevan, 1927), which is a classic example of an exposure-response test and it is a measure of acute lethality (Eaton and Klaassen, 2001; Frazier, 1992). The principle of this test is to dose groups of animals (normally at least 5 animals) with a single dose of a test substance at concentrations expected to cause death in at least a fraction of the animals tested. Results of the test enable the calculation of the LD50 value, which is the dose that would kill 50% of the animals in the tested population within 14 days after a single exposure. The purpose was to standardize biological preparations, and then adapt them for testing the acute toxicity of chemical substances intended for human use (e.g. insulin) (Botham, 2002). This type of information has continuously provided an important basis for the safety assessment of chemical substances, and it is crucial essentially when considering accidental exposures at the workplace, at home (with specific regard to children) and following transport accidents as well (mainly for environmental assessments).

Two decades after the introduction of the LD50 test, in 1944, the U.S. Food and Drug Administration (FDA) scientist John Draize developed standardized tests for eye and skin irritation (simply known as "Draize tests") using albino rabbits (Draize, et al., 1944). This paper became one of the most cited publications in toxicology, also providing the author with the 10th medal from the American Society of Cosmetic Chemists in 1957 for his work on product safety (Draize, 1958). Moreover, a few years later the development of the Draize test, the U.S. National Cancer Institute (NCI) developed a standardized test for the identification of chemical carcinogens through the daily dosing of rats and mice for long term experiments (up to two years) (Ward, 2007; Weisburger, 1983). Then, in the early 1960s, as a reaction to
the Thalidomide tragedy (Kim and Scialli, 2011), when thousands of babies worldwide were born with debilitating birth defects caused by the drug, a number of new and more complex studies on animals were developed (i.e. reproductive and developmental toxicity studies), in which a test agent is given to a large numbers of animals before they mate, throughout their pregnancy, and after giving birth, in order to evaluate effects on reproductive performance and/or developing offspring (Collins, 2006; Hendrickx and Binkerd, 1990; Koren et al., 1998; Stern, 1981; Tyl, 2010).

Today, the use of living animals to study the potential adverse effects of new drugs, food additives, pesticides, and other substances cover a wide range of different endpoints that may be required by regulatory agencies/authorities. The main ones are: acute toxicity, skin irritation, eye irritation, corrosion, dermal sensitization, respiratory sensitization, chronic toxicity, mutagenicity, teratogenicity/embryotoxicity and carcinogenicity (Cote, et al., 1998).

### 1.2.2. Aims and objectives of animal studies

Animal-based studies have in general three major objectives:

1. The identification of the major toxic effects of the considered substances by examining a multitude of potential target tissues;
2. The identification of the toxic doses, both from single or repeated exposure;
3. The determination of the level of intake that does not result in adverse effects (NOAEL).

In order to achieve these aims, this kind of studies generally leads to observations of the clinical, histopathological and/or functional changes in the animals caused by the given dose of the considered chemical (Worth and Balls, 2002). The design, conduct and completeness of reporting of the experimental findings are of critical importance, in order to determine the validity and relevance of the results obtained from toxicological studies with mammalian species (IPCS, 1999).

However, even if in many cases the performance of this kind of tests has lead to a relatively safe use of chemicals, the use of in vivo tests has often been questioned. This matter will be discussed in details in the subsequent paragraph.
1.3. Criticism to the use of animal methods

The day-to-day practice of the procedures that rely on animal models in order to inform risk assessments on chemicals has, in many cases, led to the relatively safe use of chemicals in industry, or as agrochemicals, drugs, household chemicals or cosmetics. However, the use of *in vivo* tests for the purpose of predicting the effects occurring in humans has been often questioned, and there has been continuing intensive debate over thirty years (Balls, 1991; Baskettter, *et al.*, 1997.; Blaauboer, 2002a, 2002b; BUAV, 2001; Creton, *et al.*, 2010; Frazier and Goldberg, 1990; Langley, 2005.; Worth and Balls, 2002; York, *et al.*, 1996; York and Steiling, 1998). The reasons for these many criticisms are referred to ethical, scientific and also economic discussions and, as a consequence, many researches focus today on the way how these animal tests can be reduced, replaced or refined, accordingly to the 3Rs principle proposed by Russell and Burch (1959).

In what follows, a brief explanations of all these matters, which led to numerous legislative changes during the course of history.

1.3.1. Reasons for the criticism

1.3.1.1. Scientific reasons

The literature is full of unfortunate accidents where reliance on animal testing did not always correctly predict human toxicity (e.g. thalidomide) (Greenburg and Phillips, 2003; Kim and Scialli, 2011; Thorne, 2001), even if in most cases they led to a safe use of drugs and chemicals. As the famous NIH (National Institute of Health) pharmacologist Bernard Brodie said already in 1964: "… it is often a matter of pure luck that animal experiments lead to clinically useful drugs" (Hartung, 2008b).

Scientific reasons against the use of animals rely mainly on the fact that using animals as a surrogate for predicting the biological activities of compounds in humans is always prone to some degree of uncertainty, and involves a number of assumptions and extrapolations, mainly due to species differences (Blaauboer, 2002a; Hartung, 2008b; Hartung, 2009; Holmes, *et al.*, 2010; Langley, 2005). It is indeed recognized that different species may respond differently
to the same substance (Ekwall, et al., 1998; Gold and Slone, 1993; Hurtt, et al., 2003). Differences in reactions to chemicals even between closely related species, such as rats and mice, which are more closely genotypically and phenotypically related to each other than they are to humans, demonstrate why toxicity tests on rodents have a dubious predictivity for humans (Hartung, 2008b; Hoffmann and Hartung, 2006). Even monkeys can differ considerably from humans, as experienced tragically in 2006 when the TeGenero anti-CD28 antibody, after testing safe at 500-times higher concentrations in monkeys, led to multiple organ failure within hours in six human volunteers (Bhogal and Combes, 2006; Hartung, 2008b).

Important sources of uncertainty can be determined, for example, from differences in the mechanism of action of toxic compounds, or can be found in qualitative or quantitative differences between physiological and biochemical processes between species, and especially animals and humans. Parameters regarding the uptake, distribution, biotransformation and excretion may differ, resulting in some cases in dissimilarities in the concentration of a compound at the target tissue (Blaauboer, 2002a; Cohen, 2002, 2004; Haseman, et al., 1998; Robinson, et al., 2001; Schardein, 2000).

Other difficulties arise from the fact that extrapolations have to be made from a rather small, but homogeneous, group of laboratory animals to the very heterogeneous human general population (Blaauboer, 2002a), therefore not considering pre-existing pathological conditions. The tested animals, which are usually inbred, specific pathogen-free and genetically homogeneous, therefore, do not represent normal animals of their species or the human population of concern.

In concurring to such difficulties, there can be also problems in dosimetry (ACSH, 1997). Test animals are indeed often exposed to higher doses than would be expected for typical human exposures, requiring assumptions about effects at lower doses (Hartung, 2008b; Holmes, et al., 2010; Mehendale, 1995; Sumner and Stevens, 1994). That is because most small animals have short life spans and therefore can never be exposed to the chemicals as long as humans may be, and giving higher doses to them is an effort to maximize the sensitivity of the experiment, even if it has to be considered that, besides limiting costs and animal numbers (Hartung, 2009), many additional effects occur when extreme tissue concentrations are reached and defence systems are overwhelmed (Hartung, 2008b).
Moreover, traditionally, scaling up of doses has been used from small animals to larger humans, in order to adjust for metabolic differences in rates or sizes. However, available evidence shows that predicting the clearance of a chemical in humans on the basis of body weight ratios from animal tests does not give reliable results. Specifically, for drugs, that have several clinical trial results to confirm or contradict the animal test data, it is observed a more than 30% error rate; consequently, in the case of chemicals, for which there are seldom clinical trial results to be used, there is even less certainty (Hartung, 2008b; Langley, 2004; Mahmood and Balian, 1999).

As a consequence of these factors, despite efforts to standardize procedures, the results of some animal tests can be highly variable and difficult to reproduce, which means there is a poor reliability (Bremer, et al., 2007; Gottmann, et al., 2001; Griffith, 1964; Weil and Scala, 1971). Such a weak reproducibility can be due, indeed, to differences in strains and species, ages of animals used between the different laboratories, as well as the differences in weights and diet of the animal, for example. Other factors may include differences in technical ability of the investigators, as well as variations in ambient temperature, housing conditions of animals, humidity, noise and light/dark cycle as well (Balazs, et al., 1972; Dieke and Richter, 1945). All of these factors could significantly affect the results (Langley, 2005), and for these reasons animal-based models had been heavily criticized from the scientific point of view.

1.3.1.2. Economic reasons

The financial and economic feasibility of toxicity testing using conventional whole animals is another aspect that needed to be re-evaluated, being indeed very time-consuming and expensive (Blaauboer, 2002a; Zurlo, 1998). Indeed, in vivo toxicity tests generally take months or even years to be conducted and analyzed; for example, in the case of carcinogenicity studies, even 4 or 5 years are needed. On the other hand, what it is extremely costly is firstly the maintenance of an adequate animal number and also the employment of qualified staff that can amount even to a hundred thousand euros. As a consequence of course, the number of replicates and repetitions are typically limited, contributing to the poor reliability of the tests. Moreover, standard animal tests can be quite costly as well. A survey of testing costs in Europe (Fleischer, 2007) showed average costs of, for example, 1,200
euros for skin irritation, 1,350 euros for eye irritation, 50,000 euros for 28 days repeated-dose oral studies (even twice this for the inhalation route), 330,000 euros for a two-generation reproductive toxicity study or 780,000 euros for a rat carcinogenicity bioassay (Hartung, 2008b; USEPA, 2004). Therefore, considering the average cost for a project and that about 9,300 new projects involving animal tests are started annually in Europe, it is clear how animal testing represents an enormous market and business. A study performed by Prognos (Prognos, 2007) on behalf of the European Commission in the context of the revision of the Directive for the protection of laboratory animals (Directive 86/609/EEC) (Council Of The European Communities, 1986) has estimated that, in the EU, about 1,330 establishments (industry, contract research laboratories and universities) perform animal tests, with a turnover of about 3 billion euros per year (Hartung, 2008b).

1.3.1.3. Ethical reasons

Ethical concerns are to be considered as well while evaluating the use of animals for experimental purposes. Some conventional toxicity test methods consume indeed hundreds or even thousands of animals per substance examined (Cooper, et al., 2006; Doe, et al., 2006), and most of these tests cause many animals to experience serious pain and distress (Langley, 2005). This is particularly true for animals used in acute toxicity studies, especially those included in the high dose groups of repeated-dose experiments. Repeated injections, for example, often induce considerable local pain, and animals sometimes struggle desperately to avoid another injection, and also topical administration of irritant and corrosive substances to the skin and mucous membranes is a painful procedure. Moreover, considerable suffering must be assumed to affect tested animals, also in terms of, for example, perforated gastrointestinal ulcers, myocardial infarctions, liver necrosis and muscle wasting, besides functional disturbance (Balls, 1991). Moreover, it has to be considered also the stress and anxiety to which animals are subjected. There are many causes of stress in animals: transport, insufficient time for them to accustom, cages without enrichment, handling, procedures during daytime on night-active species, changes in groups for randomization, too many male animals caged together, just to make some important examples (Hartung, 2008b; Wolfer, et al., 2004). But there are also a long series of physical symptoms that cause such stressful condition, and they may include tremors, convulsions, loss of balance, unsteady gait, or even
paralysis, weight loss, breathlessness, excessive salivation, intestinal distension, diarrhoea, nasal or anal bleeding or discharge, lethargy, but also in the end coma and death (Balls, 1991; Langley, 2005).

For all these reasons, the number of animal protection groups throughout the world is considerably growing and have placed considerable opposition to the use of whole animals for product safety testing (Zurlo, 1998), which, as reported from some countries' statistics, accounts for up to 70% of the most painful procedures to which animals are subject for all experimental purposes.

1.3.2. The 3RS principle

As a consequence, all these scientific limitations, together with the economic and ethical concerns, are driving toxicologists to explore the potential of non-animal alternative approaches to assess toxicity in the pharmaceutical and chemical industries (Holmes, et al., 2010). Recently indeed, many in vitro methods for toxicity testing have been developed as alternatives to whole animal tests according to the 3Rs approach started in 1959 by Russell and Burch (1959).

The 3Rs principle refers to the concept of Refinement, Reduction, and Replacement, which were described by the authors as the major routes to achieve the principles of human experimental technique (Piersma, 2006), and are further defined as follows:

"Refinement" is defined as any method that reduces or eliminates pain and distress in animals during experiments. It is not enough to simply administer analgesics or anesthesia to animals in pain, but it also includes the substitution of species lower on the evolutionary scale for species that are higher in phylogeny. For example, using rodents instead of primates or using zebra fish instead of rodents in order to obtain the same or more scientific information (Goldberg and Locke, 2004).

"Reduction" is defined as any method that seeks to use fewer animals in an experimental protocol to obtain the same or similar information of scientific value, or use the same number of animals to obtain more scientifically valuable information (Goldberg and Locke, 2004; Robinson, 2005). Therefore, the number of animals used should be the minimum that is
consistent with the aims of the experiment (Zurlo, et al., 1996), which is extremely important. However, when thinking about ways to reduce the number of animals used, researchers also have to ensure that the design of their experiment is robust. If by reducing animal numbers, researchers end up with data that have no statistical significance, they have wasted animal lives, which would be unacceptable (Robinson, 2005).

Finally, "replacement" is defined by the use of techniques that do not use living animals and replace an animal model with alternative methods (Goldberg and Locke, 2004). The range of replacement alternative methods and approaches includes the improved storage, exchange, and use of information about previous animal experiments to avoid unnecessary repetition of animal procedures; use of physical and chemical techniques and predictions based upon the physical and chemical properties of molecules; use of mathematical and computer models; use of organisms with limited sentience such as invertebrates, plants and microorganisms; use of \textit{in vitro} methods including cell cultures, tissue slices, and perfused organs; and also the use of human studies, including use of human volunteers, postmarketing surveillance, and epidemiology (Robinson, 2005; Zurlo, et al., 1996). Replacement represents the ultimate goal of the principle, and although difficult, considerable progress on replacement has been made by scientists. In many areas of the biomedical sciences, indeed, \textit{in vitro} methods are increasingly used as the methods of choice in place of animal studies, not because they provide precisely the same information, but because they offer the best scientific approach (Zurlo, et al., 1996). Such \textit{in vitro} methods will be discussed in a subsequent paragraph.

Interestingly, the principles outlined by Russell and Burch in 1959 received little attention until the resurgence of the animal welfare movement in the mid-1970s, when the "Universal Declaration of the Rights of Animals" was promulgated (UNESCO, 1978), and today the concept of the 3Rs is very much in the forefront with regard to research, testing and education (Zurlo, 1998), and numerous legislative changes have been made as well in this sense. However, the greatest challenge for the 3Rs is perhaps the proposed new European Union legislation (referred to as REACH), which will require the safety testing of many already widely used household and industrial chemicals and new ones (Robinson, 2005). The entire Chapter 2 of this thesis work is dedicated to this legislation, to make clearer the matter from this point of view.
1.4. Alternative methods

For all the reasons expressed in the previous paragraphs, since its establishment and during the subsequent half-century, the issue of the 3Rs and their application, to biotechnology in general and to toxicological research and testing more specifically, has continued to influence research and development of new methodologies and testing strategies, and tens of millions of dollars have been invested by corporations, governments, and other stakeholders with the goal of advancing in this field. The research focused, therefore, on the development of alternatives to animal experiments, that are defined as "procedures which can completely replace the need for animal experiments, reduce the numbers of animals required, or diminish the amount of pain or distress suffered by the animals in meeting the essential needs of humans and other animals" (Balls, 1991, 1994). Specifically, the development of in vitro systems, thought to be easier, more reliable, reproducible and predictive of some potential hazards in humans, and also considered to be less time consuming, more humane and cost efficient (Gad, 2000) than animal systems, became indispensable.

Genetic toxicology was the first sub-discipline of toxicology in which in vitro test systems were used for toxicity testing to identify the mutagenic effects of compounds. An example of these tests could be the Ames Mutagenicity Assay with strains of Salmonella typhimurium (Ames, et al., 1973), which is dated back to the ‘70s. Anyway, the particular development of non-genotoxic in vitro toxicity studies for testing purposes did not begin to expand again until the mid 1980s, when in vitro tests have been proposed as a pre-screen or as an alternative method for endpoints such as prenatal toxicity, eye irritation, dermal irritation, tumour promotion and target organ toxicity (Atterwill, 1995; Frazier, 1992; Frazier and Goldberg, 1990; Purchase, et al., 1998). At the beginning, in vitro tests were based only on empirical cell assays and they were merely descriptive, while today they are used extensively in academia, government and industry, and it is highly recognized that they involve for example mechanistically-based assays, translational toxicology (i.e., the use of in vitro data to predict human health consequences), and that there is the need for batteries of in vitro assays to determine the safety/hazard potential of a compound and to do risk assessment (Goldberg, 2007).

This kind of systems range from relatively simple subcellular fractions, tissue slices or perfused organ preparations, through primary cultures and cell lines – grown both as
monolayer cultures or suspension cultures and as monocultures or co-cultures – to three-dimensional organotypic cultures, which include reconstructed tissue models. They can be used to evaluate both cytotoxicity and target organ toxicity, and each of these systems has its specific advantages and disadvantages, as reported by Bhogal et al. (2005) in Table 1.1 and briefly summarized in what follows.

Table 1.1. The advantages and disadvantages of the various types of tissue culture systems used in in vitro toxicology.

<table>
<thead>
<tr>
<th>System</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary cultures</td>
<td>Obtainable from various target tissues; can retain in vivo tissue-specific characteristics</td>
<td>Short in vitro lifespan; progressively lose in vivo properties; prone to contamination</td>
</tr>
<tr>
<td>Monolayers and mono-cultures</td>
<td>Can be grown to confluence and subcultured; can be used as barrier models; used to quantify cell proliferation/growth; suitable for genetic manipulation</td>
<td>Limited interactions between cells; absence of other cell types, nervous, immune and endocrine systems</td>
</tr>
<tr>
<td>Co-cultures</td>
<td>Involve more than one cell type, so resemble in vivo situation more closely (e.g. blood-brain barrier)</td>
<td>Some cell combinations are incompatible with each other in culture; complicated/conflicting cell culture requirements</td>
</tr>
<tr>
<td>Continuous cell lines</td>
<td>Readily available and reproducible source of cells; avoids repeated cell isolation from animals or humans</td>
<td>Tend to lose in vivo differentiation and take on new properties induced by culture conditions; enter senescence and decline after a certain number of population doublings</td>
</tr>
<tr>
<td>Genetically engineered cell lines</td>
<td>Generated by transforming cells with foreign DNA; DNA might encode structural or functional proteins; used to create polymorphic cell line batteries</td>
<td>Techniques are specialized; methods do not always lead to permanent changes; limited potential for altering cellular features</td>
</tr>
<tr>
<td>Immortalized cell lines</td>
<td>Generated from human/animal cells by introducing oncogenes/telomere-controlling DNA; cells have cell line longevity but can retain tissue-type specific features</td>
<td>The immortalization techniques are specialized; there is not always permanent immortalization</td>
</tr>
<tr>
<td>Stem cells</td>
<td>Cells are able retain their stem cell capacity and to differentiate into many cell types</td>
<td>Limitations on cell types that can be generated; some animal species/strain limitations; ethical problems when using human embryonic stem cells</td>
</tr>
<tr>
<td>Tissue slices</td>
<td>Represents complexity of the organ; cellular contacts retained; useful for inter-species comparisons; many organs from same donor can be obtained; histological and biochemical tests possible; slices from different organs can be co-cultured; regional effects in same organ are particularly useful for metabolism studies</td>
<td>Difficult to produce reproducibly; exposure and activity of cells in slices can vary; limited in vitro lifespan</td>
</tr>
<tr>
<td>Organotypic cultures</td>
<td>Multilayered and spatially differentiated; exhibit cellular communication; good retention of in vivo physiology; can be generated from primary/immortalized cells; proprietary models available</td>
<td>Correct culture conditions can be difficult to define; batch variation of propriety models; limited in vitro lifespan</td>
</tr>
<tr>
<td>Perfused cultures</td>
<td>Applicable to a variety of the systems above; perfusion restores media and removes metabolites; allows cells to grow for extended periods; high cell densities possible; long-term repeat exposure/recovery possible; can be used for whole organs (e.g. kidney)</td>
<td>Technically complex; high risk of contamination; only a small number of samples can be set up; limited in vitro lifespan</td>
</tr>
<tr>
<td>Reconstructed tissue cultures</td>
<td>Components can be controlled and varied according to purpose</td>
<td>Technically complex</td>
</tr>
<tr>
<td>Whole organs</td>
<td>Organ functions modelled closely; different cell types with cellular interactions; particularly useful for embryo-toxicity studies</td>
<td>Can be difficult to culture; limited culture life; must be freshly isolated; tend to require complex perfusion systems</td>
</tr>
</tbody>
</table>
1.4.1. *In vitro* tests for cytotoxicity

The predictive value of *in vitro* cytotoxicity tests is based on the idea of "basal" cytotoxicity, a concept formalized by Ekwall (1983). Basal cell toxicity is defined as the adverse effect of such structures and functions that are essential for cell survival and proliferation (Seibert, *et al*., 1996). It is based on the fact that toxic chemicals affect the basic cell functions and structures, that are common to all cells (including integrity of membranes and cytoskeleton, cellular metabolism, cell division and so on), and that the toxicity can be measured by assessing the cellular damage. This interference with basal cell functions will in turn influence specific functions.

A large number of *in vitro* test methods have been developed to assess a chemical’s cytotoxic potential (Zurlo, 1998), employing a wide variety of cell lines that can be isolated from many tissues and species, such as for example a wide range of murine and bovine cells, but also some human cell lines and primary cells (e.g. HL60 cells, HeLa, erythrocytes, hepatocytes, and so on) (Clemedson *et al*., 2000; Clothier *et al*., 1999; Fry *et al*., 1990; Hulme *et al*., 1987; Mancuso *et al*., 2010; Reader *et al*., 1989; Riddell *et al*., 1986a, 1986b; Spielmann *et al*., 1999; Start *et al*., 1986; Worth and Balls, 2002). The introduction of the use of human cells, instead of animal ones, had the obvious advantage of avoiding the need for interspecies extrapolation (Combes, 2004) and therefore, today, the research is trying to focus mainly on them.

1.4.1.1. General advantages and disadvantages

Cell-based assays have several advantages, first of all the ease of performance and minimal or no animal use (Piersma, 2006; Zurlo, 1998), and the fact that a single organism can generate multiple cultures, which can be used also over extended periods of time and/or cryopreserved for future uses (Bhogal, *et al*., 2005; Brendler-Schwaab, *et al*., 1994). Moreover, these kinds of tests are more relevant and manageable than conventional laboratory animal models, and can provide several essential information about the potential effects of chemicals on specific cell properties, which can be studied both at molecular and cellular level (Bhogal, *et al*., 2005). That is a significant property even if, it has to be considered that projecting the effects of a chemical on a complex organism when the
observations are confined to a single type of cells in a dish is difficult, if not impossible. In addition, cell-based assays offer many advantages over in vivo studies, since they are generally less expensive and they also may be conducted under more controlled and automated conditions. These methods may be considered as both reduction alternatives (since many fewer animals are used compared to in vivo studies, despite the fact that small numbers of animals are still needed to obtain cells), and refinement alternatives (because they eliminate the need to subject the animals to the adverse toxic consequences imposed by in vivo experiments) (Zurlo, 1998).

However, their relative simplicity is at the same time the major disadvantage of cell-based assay, because it does not represent the complexity of the entire organism, as already mentioned (Piersma, 2006; Zucco, et al., 2004). Indeed, the entire toxicological process consists of events that begin with the organism’s exposure to a physical or chemical agent, progress through cellular and molecular interactions and ultimately manifest themselves in the response of the whole organism. On the contrary, in vitro tests are generally limited to the part of the toxicological process that takes place at the cellular and molecular level (Zurlo, 1998), and they allow the study of effects on single mechanisms only, e.g. cell proliferation [Pratt and Willis, 1985], cell adhesion (Braun, et al., 1979), metabolic cooperation (Trosko, et al., 1982), or cellular differentiation (Mummery, et al., 1984; Piersma, et al., 1993; Piersma, 2006; Spielmann, et al., 1997).

Furthermore, it is important to recognize that cells are finely balanced homeostatic machines that respond to external stimuli through complex pathways. Therefore, because toxicity could be the result of a multitude of cellular events (such as changes in cell morphology, differentiation, proliferation, function, excitability and/or communication), these in vitro systems - before being considered for risk assessment purposes - might require refinement (Bhogal, et al., 2005). For example, cell culture systems often lack essential systemic contributors to overall absorption, distribution, metabolism and excretion (ADME) (Bhogal, et al., 2005; Blaabjerg, 2002a, 2002b; Flint, 1990), and also the complex interactions and effects of the immune, endocrine and nervous systems. One way to circumvent problems with metabolic competence is by adding subcellular or cellular metabolizing systems and assessing the production of known metabolites (Coecke, et al., 1999). This is particularly important, for example, when considering the elimination of lipophilic compounds, which is
crucially reliant on metabolism, that can generate toxic intermediates from innocuous parent chemicals (Bhogal, et al., 2005).

Besides these limitations, it is important to be aware that also technical problems could also affect these tests. They include, among the others, the type of cells employed to study in vitro cytotoxicity, the solubility of the test chemicals studied, or the exposure period as well. Indeed, the contact time between cells and test chemicals, which is necessary to evaluate any delayed toxic potential of compounds, is not always sufficient (Blaauboer, 2002b).

As a consequence of all these factors, difficulties in the interpretation and extrapolation of the results may arise as well, and it is therefore clear that the full potential of in vitro methods in toxicological risk assessment will need some additional considerations (Blaauboer, 2002b). In order to interpret the results of in vitro toxicity tests, to determine their potential usefulness in assessing toxicity and also to relate them to the overall toxicological process in vivo, it is firstly necessary to understand which part of the toxicological process is being examined. In an ideal situation, the mechanism of toxicity of each chemical would be known, such that the information obtained from in vitro tests could be fully interpreted and related to the response of the whole organism. However, this is virtually impossible, since relatively few complete toxicological mechanisms have been elucidated. Thus, toxicologists face with a situation in which the results of an in vitro test cannot be used as an entirely accurate prediction of in vivo toxicity because the mechanism is unknown. However, frequently during the process of developing an in vitro test, components of the cellular and molecular mechanisms of toxicity are elucidated, and therefore these kinds of tests are of fundamental importance and significance (Zurlo, 1998).

### 1.4.1.2. Methods

In these kinds of systems, the cells are generally exposed to a concentration range of the test compound for certain periods of time, in order to determine by direct exposure a critical concentration at a specific target site and specific toxic endpoint. Many in vitro models, which are based on either a colorimetric or bioluminescence reaction for example, have been developed and assays are currently available to measure the total number of cells, the number of dead or living one, or even the mechanism of cell death (e.g. apoptosis or necrosis), as well
as specialized function effects, such as for example contractility in the case of muscle cells, motility and velocity in the case of spermatozoa, glycogen content for hepatocytes, and many different others (Worth and Balls, 2002). Table 1.2, provided by Worth and Balls (2002), summarized both the types of cells and of endpoints used, and a brief overview of some of the most common cytotoxicity tests is reported in what follows.

**Table 1.2. An overview of in vitro assays for cytotoxicity/acute lethal toxicity.**

*(LDH - lactate dehydrogenase; MDBK - Madin–Darby bovine kidney; NK - natural killer.)*

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Measurement</th>
<th>Cell line(s)</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell growth/ viability</td>
<td>ATP content or leakage</td>
<td>ELD cells (mouse); erythrocytes (mouse); LS-L929 cells (mouse); hepatocytes (rat); spermatozoa (bovine); HL-100/HeLa cells (human)</td>
<td>Clemadson et al., 2000</td>
</tr>
<tr>
<td>Cell morphology</td>
<td></td>
<td>C9 cells (rat); hepatocytes (rat); L2 cells (rat); MDBK cells (bovine); Chang liver cells (human); HeLa cells (human); McCoy cells (human); WI-1003/HeLa G2 cells (human)</td>
<td>Clemadson et al., 2000</td>
</tr>
<tr>
<td>Chromium release</td>
<td></td>
<td>LS-L929 cells (mouse)</td>
<td>Clemadson et al., 2000</td>
</tr>
<tr>
<td>Creatine kinase activity</td>
<td></td>
<td>Muscle cells (rat)</td>
<td>Clemadson et al., 2000</td>
</tr>
<tr>
<td>Haemolysis</td>
<td></td>
<td>Erythrocytes (human)</td>
<td>Clemadson et al., 2000</td>
</tr>
<tr>
<td>Killing index (sic)</td>
<td></td>
<td>SQ-5 cells (human)</td>
<td>Clemadson et al., 2000</td>
</tr>
<tr>
<td>LDH release</td>
<td></td>
<td>3T3 cells (mouse); hepatocytes (rat, human); Hep G2 cells (human); lymphocytes (human); SQ-5 cells (human); cornea cells (human)</td>
<td>Clemadson et al., 2000</td>
</tr>
<tr>
<td>Neutral red uptake</td>
<td></td>
<td>3T3 cells (mouse); L929 cells (mouse); V79 cells; H984A-43 cells (mouse); BHK cells (hamster); hepatocytes (rat, human); HeLa cells (human); Hep G2 cells (human); keratinocytes (human)</td>
<td>Clemadson et al., 2000</td>
</tr>
<tr>
<td>Neutral red release</td>
<td></td>
<td>Rabbit corneal fibroblasts or mouse embryonic fibroblasts or normal human epidermal keratinocytes</td>
<td>Reader et al., 1989</td>
</tr>
<tr>
<td>Plating efficiency</td>
<td></td>
<td>HeLa cells (human)</td>
<td>Clemadson et al., 2000</td>
</tr>
<tr>
<td>Viable cell count</td>
<td></td>
<td>LS-L929 cells (mouse); polymorphonuclear leukocytes (human)</td>
<td>Clemadson et al., 2000</td>
</tr>
<tr>
<td>Cell cycle distribution</td>
<td></td>
<td>Dauph (human); RERF-LC-A1 cells (human)</td>
<td>Clemadson et al., 2000</td>
</tr>
<tr>
<td>Glucose consumption</td>
<td></td>
<td>Muscle cells (rat)</td>
<td>Clemadson et al., 2000</td>
</tr>
<tr>
<td>Macromolecule content</td>
<td></td>
<td>HT1C cells (rat); Hep G2 cells (human)</td>
<td>Clemadson et al., 2000</td>
</tr>
</tbody>
</table>
One commonly used *in vitro* test for cytotoxicity is the Neutral Red Uptake (NRU) assay (Borenfreund and Puerner, 1985), which has been used also as method of choice in this thesis work (Scanu, *et al.*, 2011). This method is based on the ability of viable cells to incorporate and bind Neutral Red (NR), which is a supravital dye that readily diffuses through the plasma membrane and concentrates in lysosomes of viable cells. For most applications, the assay is performed in a 96-well plate format, where each well can be used for a single determination,
making this arrangement suitable for multiple concentrations of the test chemical, as well as positive and negative controls, and for a sufficient number of replicates for each concentration. Following the treatment of the cells with various concentrations of the chemical, the cells are rinsed and treated with NR, which may be added upon removal of the chemical to determine immediate effects, or also at various times after the removal, in order to determine cumulative or delayed effects. The intensity of the colour in each well, measured by a spectrophotometer, corresponds to the number of living cells in that specific well (Zurlo, 1998). What happens is that toxicants can alter the cell surface or the lysosomal membrane and cause fragility and other adverse changes that gradually become irreversible, causing cell death and/or inhibition of the cell growth. These effects cause a decrease of the amount of NR retained by the culture that is therefore directly proportional to the number of living cells. The NRU assay has an extended use among researchers, and it has been shown to be a reliable and sensitive assay for cell viability (Dierickx, 2000; Dierickx and Scheers, 2002; Eirheim, et al., 2004; Jirova, et al., 2003; Putnam, et al., 2002; Spielmann, et al., 1999). Recently, two proposed in vitro NRU assays using two different cell lines, mouse fibroblast (BALB/c) 3T3 cells and Normal Human Keratinocytes (NHK), have been validated by ICCVAM for predicting cytotoxicity in vitro (ICCVAM, 2006a, 2006b).

Another relatively simple assay for cytotoxicity, based on a similar principle, is the MTT (3[4,5-dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide) test, developed by Mosmann (1983). Tetrazolium salts (Smith, 1951) have been used for many years to distinguish living cells from dead ones, and this assay is based on a tetrazolium dye (MTT) that is reduced by mitochondrial enzymes to formazan granules of a blue colour. Only cells with viable mitochondria will retain the ability to carry out this reaction; therefore, the colour intensity is directly related to the degree of mitochondrial integrity, and it is a direct measure of the viability of the culture, with the advantages of rapidity and precision. This test is useful in order to detect general cytotoxic compounds, as well as those agents that specifically target mitochondria (Zurlo, 1998).

Both of these methods are colorimetric ones and use the viability of the cells as major endpoint. Other similar tests, based on the same endpoint, could be, for example, the Lactate Dehydrogenase (LDH) test, that measure the integrity of the membrane by the activity of the cytoplasmic enzyme released by damaged cells (Decker and Lohmann-Matthes, 1988; Korzeniewski and Callewaert, 1983), or also the Adenosine Triphosphate (ATP) one, which
is a bioluminescence method that uses a specific enzyme (luciferase) in order to measure the concentration of ATP (present in all the metabolically active cells) (Crouch, 2000; Crouch, et al., 1993; Slater, 2001).

Besides these, there are many new methods being developed also to detect cellular damage, and more sophisticated methods which employ fluorescent probes to measure a variety of intracellular parameters, such as calcium release and changes in pH and membrane potential. In general, these probes are very sensitive and may detect more fine cellular changes, thus reducing the need to use cell death as an endpoint (Worth and Balls, 2002; Zurlo, 1998).

Once data have been collected on a series of chemicals using one of these tests, the relative toxicities may be determined and used to compare toxicities of different chemicals in vitro. This relative toxicity may be expressed as the concentration that exerts a 50% effect on the endpoint response of untreated cells, and can be referred to different values, such as the EC$_{50}$ (Effective Concentration for 50% of the cells) or the IC$_{50}$ (Inhibitory Concentration, which is the concentration of a chemical that causes a 50% inhibition of a cellular process, e.g. the ability to take up NR). These data are then generally used in order to predict the toxicity in vivo, even if it is not easy to assess whether they are comparable to their relative in vivo toxicities, since there are so many confusing factors in the in vivo system, as already analyzed above in this Section. However, despite the numerous complexities and difficulties in extrapolating from in vitro to in vivo, these in vitro tests are proving to be very valuable, mainly because they are simple and inexpensive to perform, and also generally rapid and precise. They may be used as screens to flag highly toxic drugs or chemicals at early stages of development (Zurlo, 1998).

1.4.2. Target organ toxicity

Several in vitro tests have been developed in order to assess specific target organ toxicity as well, using cells or tissues from the organs that are typically target of toxicity (e.g. liver, kidney, lung, cardiovascular system, and so on) (Spielmann, et al., 1998). The detection of specific target organ and target organ system toxicity is, indeed, another important aspect of toxicological testing (Worth and Balls, 2002). As defined by Klaassen (2001), “most chemicals that produce systemic toxicity do not cause a similar degree of toxicity in all
organs, but usually produce the major toxicity to one or two organs. These are referred to as target organs of toxicity for that chemical." Table 1.3, provided by Worth and Balls (2002), summarize some of the available models in relation to three of the most important target toxicity: hepatotoxicity, nephrotoxicity, and neurotoxicity.

**Table 1.3. An overview of models for chronic toxicity testing.** (ALAT - alanine aminotransferase; ASAT - aspartate aminotransferase; CYP - cytochrome P450; FACS - fluorescence-activated cell sorting).

<table>
<thead>
<tr>
<th>Models</th>
<th>Advantages</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hepatotoxicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isolated perfused liver</td>
<td>The <em>in vitro</em> model closest to the <em>in vivo</em> situation</td>
<td>Short life-time (2-3 hours)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complicated and demanding set-up</td>
</tr>
<tr>
<td>Liver slices</td>
<td>Retains the <em>in vivo</em> tissue organisation</td>
<td>Short life-time (7 days)</td>
</tr>
<tr>
<td>Isolated hepatocytes</td>
<td>The most frequently used <em>in vitro</em> model for long-term hepatotoxicity testing</td>
<td>Short life-time (24 hours)</td>
</tr>
<tr>
<td></td>
<td>Can produce a metabolite profile for a drug very similar to that found <em>in vivo</em></td>
<td>Loss of many liver-specific functions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Availability of human cells very limited</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-human models are not always predictive of human <em>in vivo</em> situation</td>
</tr>
<tr>
<td>Collagen sandwich cultures</td>
<td>Structural and functional integrity retained for up to 15 days</td>
<td>Loss of several differentiated functions over time</td>
</tr>
<tr>
<td></td>
<td>Normal cell shape is kept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intact structure, bile canaliculi are preserved</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ALAT and ASAT enzyme releases can be studied</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Successfully used in mimicking chronic treatment in the <em>in vivo</em> situation</td>
<td></td>
</tr>
<tr>
<td>Ito cells and Ito cell/hepatocyte co-cultures</td>
<td>Useful tools for studying liver fibrogenesis</td>
<td>Not a well-established model</td>
</tr>
<tr>
<td></td>
<td>Life-time 56 hours</td>
<td></td>
</tr>
<tr>
<td>Genetically engineered cells expressing single human or animal P450 enzymes</td>
<td>Used as a tool to assess the involvement of certain enzymes in metabolism, metabolite formation, and metabolism-dependent toxicity</td>
<td>Well-characterised, well-documented and easy to use</td>
</tr>
<tr>
<td>Cell lines derived from human hepatoma</td>
<td>Expression of CYP1A1</td>
<td>Low expression of other CYPs</td>
</tr>
<tr>
<td>HepG2 cells grown under continuous medium supply</td>
<td>Permanent cell lines, well characterised</td>
<td>Laborious systems for routine use</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Only used in a few pilot studies</td>
</tr>
<tr>
<td><strong>Nephrotoxicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Renal epithelial cell lines grown under continuous medium supply</td>
<td>Phenotypes with oxidative energy metabolism are available for tubular proximal cells</td>
<td>Laborious systems for routine use</td>
</tr>
<tr>
<td></td>
<td>Morphology very close to the <em>in vivo</em> parent cell type</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Culture periods of up to 6 weeks are possible</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Human-derived proximal tubular cell lines similar in function to parent cells are available</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Many functional parameters can be monitored, and a number of endpoints can easily be assessed</td>
<td></td>
</tr>
<tr>
<td>Models</td>
<td>Advantages</td>
<td>Limitations</td>
</tr>
<tr>
<td>------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>Neurotoxicity</td>
<td>Well characterised</td>
<td>Donor animals required</td>
</tr>
<tr>
<td>Primary neuronal cell cultures (rat) and their reaggregates</td>
<td>Identification of neurodegenerative compounds is possible</td>
<td>More intensive than cell lines</td>
</tr>
<tr>
<td></td>
<td>Long-term exposure possible (7-14 days for monolayer; 3 months for reaggregates)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Possibility to study oxidative stress and excitotoxicity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Used in industry for screening pharmaceuticals and agrochemicals</td>
<td></td>
</tr>
<tr>
<td>Permanent neuronal cell lines</td>
<td>Useful for detection of delayed neurotoxicity caused by organophosphates</td>
<td>Only one model available</td>
</tr>
<tr>
<td>Astrocytes</td>
<td>Culture system well developed</td>
<td>Require long-term culture for differentiation</td>
</tr>
<tr>
<td></td>
<td>Possibility to detect reactive astrocytes (cytokine production)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Specific surface markers can be detected by cell imaging and FACS analysis after a toxic insult</td>
<td></td>
</tr>
<tr>
<td>Oligodendrocytes</td>
<td><em>In vitro</em> models for studying demyelination and remyelination are established</td>
<td>Need long-term maintenance for maturation and/or testing of toxic compounds</td>
</tr>
<tr>
<td></td>
<td>Specific surface markers can be detected by cell imaging and FACS analysis after a toxic insult</td>
<td></td>
</tr>
<tr>
<td>Microglia</td>
<td>Rapid response to neuronal injury</td>
<td>Cultures require 1–2 weeks for the production of well-characterised microglia</td>
</tr>
<tr>
<td></td>
<td>Cytokine production, morphology, phagocytosis and proliferation are useful endpoints</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cell culture models and endpoints measurements are successfully developed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Primary glial cell cultures are widely used in academia and in industry</td>
<td></td>
</tr>
<tr>
<td>Brain slices from hippocampus</td>
<td>Useful for detecting excitotoxic and/or convulsive properties of drugs (hardly detectable in rodents)</td>
<td>Long-term culture conditions are required</td>
</tr>
<tr>
<td></td>
<td>Well accepted for studying learning and memory deficits (difficult to detect in rodents)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Used in agrochemical testing</td>
<td></td>
</tr>
</tbody>
</table>

### 1.4.2.1. General advantages and disadvantages

These systems have several different advantages, depending on the type of system used, ranging from the retention of tissue- and organ-specific functions, biotransformation capability, and in some cases also retention of cell-to-cell and cell-to-matrix interactions and organotopical three-dimensional structures.

However, there are also a number of difficulties associated with the design of such tests. The most notable one is the inability of maintaining many of the features of the organ *in vivo*. Frequently indeed, when cells are harvested from animals and placed into culture, they tend either to degenerate quickly and/or to dedifferentiate, which means they lose their organ-like functions and become more generic, with the consequence that within a short period of time,
usually a few days, the cultures are no longer useful for assessing organ-specific effects of a potentially toxic chemical. Moreover, even if it is possible to acquire much information about mechanisms of toxicity using one or more of these techniques, there is also in this case the difficulty of extrapolation from an in vitro system to the whole process occurring in vivo. Many of these problems are being overcome thanks to the recent advances in molecular and cellular biology, and all the new information obtained about the cellular environment in vivo may be utilized in modulating culture conditions in vitro. For example, many growth factors and cytokines are now commercially available and can be added to cells in culture, helping to preserve their integrity and to retain more differentiated functions for longer periods of time. Moreover, the knowledge of nutritional and hormonal requirements of cells in culture is increased, so that new media may be formulated (Zurlo, 1998).

Recent advances have also been made in tissue engineering and identification of both natural and artificial extracellular matrices, which is important since culture of cells on them can have profound effects on both their structure and function. Therefore, and despite all the difficulties, these systems can provide great insight into organ-specific mechanisms of toxicity (Zurlo, 1998; Lavik and Langer, 2004).

1.4.2.2. Methods

Many target organ toxicity studies are conducted in primary cells, which are defined as cells that are freshly isolated from an organ, and that usually exhibit a finite lifetime in culture. Besides this, such cultures are very useful for studying specific cellular targets of a chemical and provide many advantages for toxicity assessment (Zurlo, 1998). The major one includes tissue-specific functions and retention of capacity for biotransformation. However, cellular isolation can result in damage to cell membrane integrity, with damage to or loss of membrane receptors and cellular products, even if fortunately, when forming monolayers, such cell damage is often repaired (Broadhead and Bottrill, 1997; Knight, 2008).

Thanks to molecular biological tools, the development of continuous cell lines, that can be useful for target organ toxicity testing thus eliminating the need for freshly isolated cells, has been possible. These cell lines are generated by transfecting DNA (usually from a virus, e.g. adenovirus) into primary cells. Such DNA contains a gene or genes that, when expressed,
allow the cells to become immortalized, which means able to live and grow for extended (often indefinite) periods of time in culture (Knight, 2008; Zurlo, 1998). That is a clear advantage compared to the short life time of primary cells. Immortalization of cell lines, however, may significantly alter their characteristics and function (Luttun and Verfaillie, 2006), therefore, before these cell lines may be used as a surrogate for a specific cell type, they must be carefully characterized to determine how "normal" they really are (Zurlo, 1998). The ability to generate immortalized cells, combined with the advances in cell culture technology, have greatly contributed to the creation of cell lines from many different species and organs, including rabbit kidney cells, mouse macrophages, rat hepatocytes, but also human lymphocytes, osteoblasts, and many others (Broadhead and Bottrill, 1997; Zurlo, 1998).

In recent years, tissue engineering has greatly facilitated the development of more-complex in vitro models than simple cell cultures (Lavik and Langer, 2004). In some instances, two or more cell types from an organ may be cultured together, and this provides an added advantage of being able to look at cell-to-cell interactions. It is possible to co-culture cells from the same, but also different organs, for example liver and kidney (Choucha-Snouber, et al., 2010), in order to assess the specific effects on kidney cells, of a chemical that must be bioactivated in the liver (Zurlo, 1998). It is also now possible to grow stratified layers, each of them exhibiting morphological and functional differentiation, of epidermal cells. This has given rise to several commercially-available organotypic and reconstructed in vitro culture models, including EpiSkin® (http://www.loreal.com) and EpiDerm™ (http://www.mattek.com), on which test protocols for skin corrosivity have been scientifically validated (Bhogal, et al., 2005; Botham, 2004; ICCVAM, 2002; OECD, 2004). Other organotypic models have also been developed, for example for ocular irritation (Zorn-Kruppa, et al., 2004), neurotoxicology (Atterwill and Purcell, 1999), and respiratory toxicity as well (Bhogal, et al., 2005; Gray, et al., 1996).

Of particular relevance to the development of many of these cell-based organotypic models is the use of substrates (Bhogal, et al., 2005), both natural and artificial extracellular matrices on which cells may be cultured, potentially having profound effects on both their structure and function, and thus providing great insight into organ-specific mechanisms of toxicity.
Other *in vitro* systems for studying target organ toxicity involve even more increasing complexity, and as they progress from single cells to whole organ cultures, they become more comparable to the *in vivo* situation, being very useful to assess interactions between multiple cell types. However, at the same time, they become much more difficult to control, because of an increased number of variables involved. Moreover, many of these systems can be used only in short-term studies, since they have limited *in vitro* longevities, which in turn limit their use (Zurlo, 1998).

These progresses involve, for example, the availability of new instruments that enable the researcher to cut uniform tissue slices in a sterile environment. Precision-cut tissue slices are being used extensively for toxicological studies, and offer some advantage over cell culture systems mainly because all of the cell types of the organ are present and they maintain their *in vivo* architecture and both cell-to-cell and cell-to-matrix interactions (Zurlo, 1998). Indeed, organ slices represent a multicellular three-dimensional *in vitro* model, that possesses all the biologically relevant structural and functional features of *in vivo* tissues. Their potential for evaluating mechanisms of organ toxicity is due to the fact that all the various cell types within an organ will contribute to the biotransformation of a chemical, as well as to the release of a multitude of mediators which regulate the function of other cells. In the case of some chemicals, the primary target could be a minor cell population of the tissue, such as for example endothelial cells. Hence, the various cell types and the cell-to-cell and cell-to-matrix interactions become an important factor in the overall assessment of drug-induced effects and will enhance our ability to make better predictions of *in vivo* outcomes (Vickers and Fisher, 2004). The main disadvantage of these systems, as already mentioned, is the difficult in the maintenance of viability. Slices, indeed, degenerate rapidly after the first 24 hours of culture, mainly due to poor diffusion of oxygen to the cells on the interior of the slices. However, different studies (Fisher, *et al.*, 1995) have indicated that, by gentle rotation, it is possible to achieve a more efficient aeration allowing, together with the use of a more complex medium, the slices to survive for up to 72 hours. Moreover, advances continue to be made in obtaining higher quality tissues, and today the organ slice methodology is readily adaptable to various organs and species (including human), facilitating experimentation of cross-species comparisons (Bussek, *et al.*, 2009; de Kanter, *et al.*, 2002; Farkas and Tannenbaum, 2005; Liberati, *et al.*, 2010; Parrish, *et al.*, 1995). The application of human tissue to slice studies of course increases the utility of this model, and provides an important
bridge between animal derived data and the extrapolation to human outcome (Vickers and Fisher, 2004).

Isolated perfused organs may also be used to assess target organ toxicity. Isolated organs represent three-dimensional biological system with a certain degree of retained physiological functions, native cellular architecture, and extracellular matrix that are superior to laboratory cell- and tissue-based bioassays. These systems are considered, for the above mentioned preservation of the three-dimensional organ structure, the closest model to the in vivo situation. They allow, indeed, for the maintenance of intra-organ interactions, and offer an advantage similar to tissue slices, since all cell types are present, but avoiding the stress to the tissue due to the manipulations involved in preparing slices. On the other hand, one of the major disadvantages is again their short-term viability, which limits their use for in vitro toxicity testing. Moreover, it has to be considered that, once extracted from the human body, an organ is instantly cut off from the original physiological environment in terms of blood supply, nervous modulation, immunoregulation, and thermo-homeostasis. Therefore, even if it is possible to set similar conditions for example in terms of body temperature and simulate the flow pattern due to normal heart beat and blood pressure, it is impossible to reproduce those dynamic regulatory changes due to nervous and immune regulation (Leung, 2009; Zurlo, 1998).

In terms of serving as an alternative, these systems may be considered a refinement method, since the animals do not experience the adverse consequences of in vivo treatment with toxicants; however, at the same time, their use does not significantly decrease the numbers of animals required (Zurlo, 1998), and therefore they cannot be considered neither as reduction, nor as replacement alternatives.

1.4.3. Use of stem cells in toxicology

Special attention regarding in vitro toxicology has to be posed to models based on human Pluripotent Stem Cells (PSCs), that today are becoming an attractive alternative. Their recent application in toxicology and drug research, indeed, provide new alternative to the standard routine tests performed by industry and offer new strategies and benefits for chemical safety assessment (Laustriat et al., 2010; Trosko and Chang, 2010). The unique properties of PSCs,
such as unlimited proliferation ability without needing immortalization, plasticity to generate other cell types and more readily available and unlimited sources of human cells of well-defined origin and homogeneous characteristics, clearly identified their benefits in toxicology (Davila et al., 2004; Wobus and Loser, 2011), showing significant advantages in comparison with somatic cells. To date, the attention of the scientific world focused mainly on Embryonic Stem Cells (ESCs), firstly established by Thomson et al., (1998), and on the more recently described induced Pluripotent Stem (iPS) cells, reprogrammed somatic cells firstly described by Yamanaka (Takahashi et al., 2007; Takahashi and Yamanaka, 2006). Their application studies, until now, regarded mainly embryo-, cardio-, and hepatotoxicity (Wobus and Loser, 2011), but there is a wide expectations of future development of tests based on human PSCs-derived keratinocytes (Guenou et al., 2009), fibroblasts (Cao et al., 2008), and also neural cells (Yla-Outinen et al., 2010; Zeng et al., 2006), that could significantly improve drug safety and drug development.

With regard to embryotoxicity, several in vitro assays based on ESCs have been developed (Fraichard et al., 1995; Laschinski et al., 1991; Lim et al., 2009; Meamar et al., 2010; Rohwedel et al., 2001; Schmidt et al., 2001; Strubing et al., 1995; Wiese et al., 2011; Wobus et al., 1988, 1991, 1994), but the one developed by Spielmann and colleagues in 1997 (Spielmann et al., 1997, 1998), the mouse Embryonic Stem Cell Test (mEST), is the only one that has been validated (Genschow et al., 2002). The mEST evaluate chemicals' cytotoxicity and its inhibitory effect on the spontaneous differentiation of murine ESCs into functional contractile cardiomyocytes in embryoid bodies. Importantly, data obtained from the study were used to generate a prediction model for the distinction of chemicals in three classes: non, weakly or strongly embryotoxic chemicals (Genschow et al., 2002, 2004; Seiler et al., 2006), predicting the embryotoxic potential of the tested compounds with a very high percentage of accuracy (about 78% for non and weakly embryotoxic compounds and 100% for strongly embryotoxic ones) (Genschow et al., 2002, 2004). However this test, even being the only well-established in vitro test for developmental toxicity based on mammalian cells (Bremer and Hartung, 2004), resulted to have several limitations (Piersma, 2004; Schmidt et al., 2001; Seiler et al., 2006). For example, it was laborious and time-consuming, and data on only the 20 chemicals used for the validation study were considered not sufficient to make final decisions on embryotoxicity. Moreover, there was the need for more reliable and reproducible differentiation procedures. Therefore, modifications and new endpoints were
recommended (Spielmann, 2009), such as the development of markers of neuronal, bone and cartilage in addition to cardiac ones, and also the inclusion of novel molecular endpoints in the future (Marx-Stoelting et al., 2009; Rohwedel et al., 2001; Spielmann, 2009). Such improvements would also be needed for the establishment of similar tests based on human ESCs, that would represent a significant progress, enhancing the predictivity of in vitro assays and avoiding problems associated with the interpretation of results from animal-based assays in a human context. Today, there are several recent studies (listed in Table 1.4 provided by Wobus and Loser (2011) that started to demonstrate that human ESCs could be used as a suitable model for analysing developmental toxicity, but it is clear that several problems (similar to those expressed for murine cells) still need to be solved.

The application of human iPS cells in embryotoxicity has been discussed (Heng et al., 2009) and considered as a promising tool as well, but future works are still needed to show if they can be applicable and if they may offer advantages over ESCs.

ESCs have been widely used also in the field of cardio- and hepatotoxicity, for which they could have significant advantages over the in vitro and in vivo systems currently used. Indeed, since the first successful derivation of human ESCs-derived cardiomyocytes (Kehat et al., 2001), and the early studies where mouse ESCs differentiated resulting in the formation of hepatocytic cells expressing hepatocyte-specific genes (Lavon and Benvenisty, 2005), a large number of studies has been published on the generation of cardiac myocytes or hepatocytes from ESCs and from iPS cells as well (Abe et al., 1996; Baxter et al., 2010; Beqqali et al., 2006; Burridge et al., 2007; Freund et al., 2010; Laflamme et al., 2007; Lavon and Benvenisty, 2005; Liu et al., 2010; Mohr et al., 2010; Moore et al., 2008; Moretti et al., 2010; Mummery et al., 2003; Pal and Khanna, 2007; Passier et al., 2005; Si-Tayeb et al., 2010a, 2010b; Song et al., 2009; Sullivan et al., 2010; Touboul et al., 2010; Yokoo et al., 2009; Zhang et al., 2009; Zwi et al., 2009). For a detailed review of the state of the art of these studies see Wobus and Loser (2011).
Table 1.4. Published developmental toxicity studies involving human ESCs and/or human ESC-derived cells.

<table>
<thead>
<tr>
<th>Study</th>
<th>hESC lines</th>
<th>Noxes tested</th>
<th>Endpoints used/identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adler et al. (2008a)</td>
<td>SA002; SA002.5</td>
<td>Known developmental toxicants (VPA, ATRA, 13-CRA); DMSO</td>
<td>hES cell viability (ATP content, resazurin reduction)</td>
</tr>
<tr>
<td>Adler et al. (2008b)</td>
<td>H1</td>
<td>Known developmental toxicant (RA)</td>
<td>hES cell viability (resazurin reduction); mRNA levels of genes coding for stemness markers (Oct4, hTert); mRNA levels of genes coding for mesodermal and cardiac differentiation markers (Brachury; GATA-4, Nkx 2.5; MyH6)</td>
</tr>
<tr>
<td>Flora et al. (2009)</td>
<td>ReliCell hES1</td>
<td>Arsenic (potential reversal of arsenic effects by MiADSA were also investigated)</td>
<td>mRNA levels of stemness markers; mRNA levels of genes representative of the three germ layers</td>
</tr>
<tr>
<td>Krishnamoorthy et al. (2010)</td>
<td>H9; Bg02; abBG02</td>
<td>Ethanol at low concentrations</td>
<td>CHRNA5 gene expression level</td>
</tr>
<tr>
<td>Lin et al. (2010)</td>
<td>H9</td>
<td>Smoke from conventional and harm-reduced cigarettes</td>
<td>hESC morphology; hESC apoptosis</td>
</tr>
<tr>
<td>Metha et al. (2008)</td>
<td>ReliCell hES1</td>
<td>Strong (busulfan; hydroxyurea), weak (caffeine, indomethacin) and non-embryotoxic (penicillin, saccharin) compounds</td>
<td>mRNA levels of genes coding for stemness markers; expression levels of genes representative of the three germ layers</td>
</tr>
<tr>
<td>Stummann et al. (2009)</td>
<td>H1</td>
<td>Methylmercury (MeHg)</td>
<td>mRNA levels of genes coding for neural lineage markers in differentiating cells (NCAM, NeuroD, MAP2)</td>
</tr>
<tr>
<td>West et al. (2010)</td>
<td>H9</td>
<td>Several drugs with known teratogenic effects in human</td>
<td>Abundance of certain small molecules (metabolites) as detected by mass spectrometry</td>
</tr>
<tr>
<td>Wilson et al. (2010)</td>
<td>H9</td>
<td>γ-radiation (0.4–4 Gy)</td>
<td>hESC viability; global gene expression changes in hESCs</td>
</tr>
<tr>
<td>Zdravkovic et al. (2008)</td>
<td>H7; H9</td>
<td>Nicotine exposure</td>
<td>Morphology and adhesion of hESCs; expression of stemness markers; hESC apoptosis</td>
</tr>
</tbody>
</table>

With regard to cardiac safety pharmacology, current preclinical models, both in vivo and in vitro, do not accurately predict clinical outcomes and present some limitations, mainly related to costs and sensitivity. Cultures of human cardiomyocytes are an excellent in vitro model system for safety evaluation (Bird et al., 2003), however the availability of primary cells from healthy donors is limited. As a consequence, the establishment of novel test systems to screen new chemicals and identify their safe use in humans is highly desirable (Davila et al., 2004; Wobus and Loser, 2011). The potential of ESCs in safety toxicology and pharmacology has already been demonstrated by studies on murine cell lines (Boheler et al., 2002; Wobus and Boheler, 2005), however only human cells could overcome inter-species limitations. The potential usefulness of human ESCs-derived cardiomyocytes was shown in several recent studies (Braam et al. 2010; Caspi et al. 2009; Jonsson et al. 2010; Liang et al. 2010; Otsuji et al. 2010; Pekkanen-Mattila et al. 2010; Peng et al. 2010), even if information on their applicability to safety pharmacology is still restricted. The advantages of these cells are related mainly to the human origin and the fact that express characteristic cardiac-specific
genes, therefore avoiding species-specific differences inherent to animal-based test systems; moreover, they are able to respond to external stimuli and show the anticipated response to a broad range of pharmaceutically active substances. Furthermore, these cells are genetically identical and can be kept in culture for at least several weeks. For all these reasons, human PSC-derived cardiomyocytes represent a promising option and may offer the basis for the development of novel, human-specific platforms for assessing the safety of novel compounds, even if several problems still have to be solved in order to make them widely applicable to screening procedures and to gain the acceptance of regulatory authorities, such as, among the others, the establishment of more efficient and reproducible differentiation protocols for the production of mature cells at large numbers (Wobus and Loser, 2011).

The knowledge obtained from progresses made on the differentiation of human ESCs into cardiomyocytes is now being applied to iPS cells as well (Freund et al., 2010; Moretti et al., 2010; Yokoo et al., 2009; Zhang et al., 2009; Zwi et al., 2009), leading to consider these cells as a promising tool for toxicological studies and also attracting the interest of pharmaceutical industry. For example, the Madison-based company Cellular Dynamics International (CDI, www.cellulardynamics.com) has launched human iPS cells-derived iCell® Cardiomyocytes designed to aid drug discovery and improve the predictability of compound efficacy and toxicity (Cellular Dynamics, 2009).

With regard to hepatotoxicity, the availability of primary cells from healthy donors is limited, because of the use of large numbers of human hepatocytes in the basic and clinical studies (Davila et al., 2004). Initial progresses have been made in the establishment of protocols for the generation of human PSCs-derived hepatocyte-like cells. However, until now, a validated standardized protocol for the differentiation is not available and there are only limited data available on long term cultivation of these cells; this is a critical factor for the implementation of novel human cell-based in vitro assays, since detection of drug effects might require the application over an extended period of time or repeated exposure. Results are anyway encouraging (Basma et al., 2009; Guguen-Guillouzo et al., 2010; Hay et al., 2008) and several advantages could rise from this kind of tests, for example they may have the potential to predict hepatotoxicity already in early phases of drug development, leading consequently to a reduction of costs and to improved safety profiles of drugs and chemicals. However, before such systems can be implemented into routine toxicity testing, also in this case some
problems still have to be solved, including again the development of robust and reproducible differentiation protocols (Wobus and Loser, 2011).

We can conclude, therefore, that human PSCs offer enormous opportunities for the development of promising and innovative in vitro toxicological systems, providing a new tool for better understanding the mechanisms involved in drug-induced adverse reactions and to potentially predict and avoid toxicity in humans. However, the controversy surrounding the use of human ESCs (Vojnits and Bremer, 2010) has limited the development of these cells, and the research is now focusing especially on the use of different sources of cells, such as the iPS cells. At the same time, also iPS cells present some important limitations, such as the serious problem of presenting a potentially tumorigenic status. Therefore, further studies are highly required. The high interests in using stem cells in safety toxicology and pharmacology is illustrated by collaborative programmes between industries and research institutions (Baker, 2010; Caspi et al., 2009; Ebert and Svendsen, 2010; Sartipy et al., 2006; Trosko and Chang, 2010), such as the consortium Stem Cells for Safer Medicine (SC4SM) (Jha, 2007).

1.5. Validation, regulatory acceptance and international organisation

In order to facilitate the replacement of old in vivo toxicity tests with new in vitro alternatives, which is of extreme importance especially with regard to the issues addressed in this Thesis Chapter, national and international authorities have developed processes and criteria for evaluating and determining if a new toxicity test methods can replace an existing one. The major difficulty in this sense is represented by test method validation and regulatory acceptance, which will be briefly discussed in what follows.

1.5.1. Validation authorities

The several organisations that begun to pursue, form, and validate in vitro toxicity systems, are distributed all around the three main continents, and are represented by groups of companies, industry associations, animal welfare organisation, and national government and
international agencies (Balls and Fentem, 1999). Three of them are the main organisations that developed the validation criteria for new toxicological test methods in use today, working together in order to harmonize and avoiding major differences between them. They are the Organisation for Economic Cooperation and Development (OECD), the European Centre for the Validation of Alternative Methods (ECVAM), and the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM).

1.5.1.1. ECVAM

The EU was the first to establish a formal validation process by creating, in 1991, the European Centre for the Validation of Alternative Methods (ECVAM, http://ecvam.jrc.it/) (Commission of the European Communities, 1991b), which became operational in 1993 and was established, pursuant a requirement in Directive 86/609/EEC on the protection of laboratory animals (Council of the European Communities, 1986), in order to conduct research, develop non-animal methods, and implement validation studies. ECVAM, as a service of the EU Joint Research Centre (JRC), has pioneered the validation process (Zuang and Hartung, 2005), participating in early validation studies and international efforts to define validation principles for alternative test methods, and publishing recommendations concerning practical and logistical aspects in a workshop report of 1995 (Balls, et al., 1995). Current principles and procedures for validation have been published by Worth and Balls (2002) and also ECVAM Guidelines have been provided for the submission of a test method (http://ecvam.jrc.it/upload_docs/m_5/Guidelines.pdf). ECVAM has continuously worked with other organizations, such as ICCVAM or OECD, with the aim of having an early exchange of information on the validation of new methods, as to facilitate internationally harmonized validation and acceptance of alternative methods.

1.5.1.2. ICCVAM

The U.S. Government established in 1994 the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM, http://iccvam.niehs.nih.gov/), with the aim of establishing a formal process by which in vitro methods would be validated and
recommended for regulatory acceptance. ICCVAM is an interagency committee composed of representatives from 15 U.S. Federal agencies, who in turn collaborate with the NTP (National Toxicology Program) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), that provides administrative and technical support for the activities of ICCVAM. NICEATM-ICCVAM conduct rigorous reviews of both new or revised test methods, and coordinate issues on validation, acceptance, and national and international harmonization of toxicological test methods (Schechtman and Stokes, 2002). It places priority attention on those methods that may improve prediction of adverse human, animal or ecological effects, and those that may reduce, refine or replace animal use. ICCVAM published its original validation and regulatory acceptance criteria document in 1997 (ICCVAM, 1997), and then a submissions guidance document has been developed (ICCVAM, 2003) to be used in conjunction with information in the 1997 document.

1.5.1.3. OECD

The Organisation for Economic Cooperation and Development (OECD, http://www.oecd.org/home/0,2987,en_2649_201185_1_1_1_1_1_1_1,00.html) is an intergovernmental organization, established in 1961 and composed today of 30 different countries, whose general mission is to promote policies that will improve the economic and social well of people all around the world. The OECD started working on guidance for test method validation around 1994, when the international debate on test method validation was considerable, and with the aim of "internationally harmonize the various published and advocated concepts for the validation of alternative test methods" (OECD, 2002). The OECD validation and acceptance criteria have been refined over time, and the current guidance is the "OECD Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment" (OECD, 2005). According to the OECD, a test method can be validated prior to being considered for an OECD Test Guideline (that is published only when consensus is reached between all members), but validation of a test is not a prerequisite for initiating the development of a Test Guideline.
1.5.1.4. OTHERS

Other organizations are distributed all around the world. For example, in Europe, there are also several other organizations besides ECVAM, such as the Fund for the Replacement of Animals in Medical Experiments (FRAME, www.frame.org.uk), established in 1982, with the aim to develop and validate alternative *in vitro* test systems and with the ultimate goal of eliminating the need to use laboratory animals in any kind of medical or scientific procedure. In the U.S., it is important also the John Hopkins Centre for Alternatives to Animal Testing, (CAAT, www.caat.jhsp.edu), established to find new methods to replace the use of laboratory animals in experiments, reduce the number of animals tested, and refine necessary tests to eliminate pain and distress. There are also, among the others, the Netherlands Centre for Alternatives to Animal Use (NCA), established as a national information centre on alternatives in 1994; the European Consensus Platform on Alternatives (ECOPA), founded in 2002 in Brussels by national platforms from 10 European states (i.e. Austria, Belgium, Czech Republic, Finland, Germany, Italy, Netherlands, Spain, Switzerland, and UK); and also the centre for the Documentation and Evaluation of Alternative Methods to Animal Experiments (ZEBET), established in Germany in 1989. Notable are also the Japanese Centre for the Validation of Alternative Methods (JaCVAM), as a part of the Japan's National Institute of Health Sciences (NIHS) and the recently established Korean Center for the Validation of Alternative Methods (KoCVAM), as a part of the National Institute of Food and Drug Safety (NIFDS) in the Korean Food and Drug Administration. Both JaCVAM and KoCVAM conducted validation studies and participated in international validation and harmonization activities.

1.5.2. The validation process

The implementation of these new *in vitro* systems in toxicological risk assessment, together with their regulatory acceptance, must be preceded by a detailed evaluation of their relevance and reliability, which means that it has to be previously performed a process, known as validation (Balls and Fentem, 1999; Balls and Karcher, 1995).
1.5.2.1. Definition

The concept of validation was recognized as necessary for acceptance of an alternative method in the ‘80s, and since then the need to create a process that would codify validation arose. The concept was simple: once a test was validated, everyone would use it (Goldberg, 2007). Validation was defined by the ECVAM as "the process by which the reliability and the relevance of a procedure are established for a specific purpose" (Balls, et al., 1995), definition that has been agreed at a workshop on the principles of the validation of toxicity test procedures, held in Switzerland in 1990 (Balls, et al., 1990a, 1990b), following a report on validation produced by Frazier (1990) for the OECD and the discussions at the 1987 CAAT symposium on validation (Goldberg, 1988). The current criteria and principles have evolved from those established there, and very similar definitions were then proposed by OECD and ICCVAM (ICCVAM, 2003; OECD, 2005).

From a scientific point of view the validation of any test procedure has a number of interesting aspects, which are also reflected in the above-cited definition. The "relevance" of a procedure refers to the scientific value and the practical usefulness of the results it provides (Balls and Fentem, 1999), for assessing whether the compound is a hazard or not (Goldberg and Locke, 2004). Its meaning was clarified in OECD Guidance Document No 34 (OECD, 2005) as follows: "Description of relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method."

During the process of validation, "reliability" needs to be established as well, which means that all the procedures should be reproducible and meaningful, with an adequate predictive ability (Balls and Fentem, 1992). Such predictivity should be defined by comparison of in vitro results with existing in vivo data (Piersma, 2006), while reproducibility concerns with results within and between laboratories, and over time. In the official definition, reliability was used instead of reproducibility, because it implies that results are good (Balls and Fentem, 1999). The term reliability was also clarified in OECD Guidance Document No 34 (OECD, 2005) as follows: "Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility and intra-
laboratory repeatability." Relevance and reliability are off course strictly related, since a procedure that provide reliable but irrelevant results, or that provide relevant but unreliable results, is definitely of any utility (Balls and Fentem, 1999).

Another crucial point is that the "purpose" needs to be clearly identified, since it covers the scope of the procedure and what can be expected of it. Essentially, it is the basis on which the validity of the method should be judged (Balls and Fentem, 1999; Barile, et al., 1994).

Validation is therefore a process by which the credibility of a candidate test is established for a specific purpose (Balls and Fentem, 1999). More importantly, it is not a process to develop new approaches, optimize approaches, or compare one approach to another, but it is essentially a process which verifies that the method or procedure under investigation performs as intended (Wilcox and Goldberg, 2011).

1.5.2.2. Steps of the process

A formal validation study requires substantial organization and planning, involving an evaluation of the performance of the test with an optimized protocol and an optimized prediction model, as well as an independent analysis of data and outcome in relation to the predefined goal of the study (Balls and Fentem, 1999). Five main stages have been identified in the evolution of new test methods: test development, prevalidation, proper validation, independent assessment, and then progression toward regulatory acceptance (Balls, et al., 1995; Wilcox and Goldberg, 2011).

Before undertaking a validation study, it is necessary to properly develop the test satisfying a set of important criteria. That means it is necessary to identify the purpose and its practical applications, explain the basis of the test, define the protocol, operating procedures and also an adequate set of chemicals and adequate controls to be used, together with a specification of the endpoints and methods for calculating and expressing results. Moreover, as proposed by Bruner et al. (1996), the development of an appropriate prediction model is also necessary (Balls and Fentem, 1999; Worth and Balls, 2004). An explanation of the need for the method in relation to existing in vivo ones and other non-animal methods is also required in this phase of the process, as well as a statement about its limitations (Worth and Balls, 2004).
As suggested by Curren et al. (1995), soon after the development of the test, a prevalidation study is now formally required in order to improve the efficiency and speed of the validation process and maximize the successful results (Balls and Fentem, 1999; Goldberg and Locke, 2004). It consists in a small scale inter-laboratory study, that provides the opportunity to optimize the test method protocol, therefore ensuring that an optimized protocol is available for further assessment in a formal validation study, and to obtain a preliminary assessment of its performance and reproducibility (Curren, et al., 1995). Moreover, it is performed to verify that the protocol is transferable to other laboratories, which is of primary importance for the concept of reliability (Balls and Fentem, 1999; Goldberg and Locke, 2004; Worth and Balls, 2004). The prevalidation process is divided into three consecutive phases (Curren, et al., 1995). During the so-called "method refinement" phase (phase I), the protocol and prediction model of a test method are refined in a single laboratory (that have previous experience in the use of the test), whereas during the "method transfer" phase (phase II), an assessment is made of the transfer ability of the method to a second laboratory, making any necessary refinements to the protocol and prediction model. During the "method performance" phase (phase III), finally, the relevance and reliability of the test are assessed under blind conditions in three or more laboratories (generally including the first two ones) (Worth and Balls, 2004). Preferably, some flexibility in methodology should be allowed, since it generally facilitates the introduction of a test system in a laboratory and when possible is a sign of robustness (Balls and Fentem, 1999; Piersma, 2006).

Then, the formal larger inter-laboratory validation study, designed to obtain a more definitive assessment of relevance and reliability, can be performed. It can be considered as an extended version of the last phase of prevalidation, in which additional chemicals are tested under blind conditions (Worth and Balls, 2004). The key to a successful validation study is clarity of definition of goals, excellence of design and management, high quality of laboratory work and statistical analysis, and strength of partnership at many levels (Balls and Fentem, 1999). The validation studies are therefore obviously complex, expensive and time consuming. They should be performed in a series of reference laboratories, since both intra-laboratory repeatability and inter-laboratory reproducibility of the candidate test must be demonstrated, using calibrated sets of chemicals from a chemical bank and cells or tissues from a single source (Zurlo, 1998), together with the application of Good Laboratory Practice (GLP) (Cooper-Hannan, et al., 1999) and Good Cell Culture Practice (GCCP) (Coecke, et al., 2005;
Hartung, et al., 2002). The results must then be subjected to appropriate statistical analysis (Zurlo, 1998). A schematic representation of the ECVAM validation process is given in Figure 1.1, taken from Worth and Balls (2004).

Figure 1.1. A schematic representation of the ECVAM validation process.

In the ECVAM process, a scientifically validated method is one that has been endorsed by the ECVAM Scientific Advisory Committee (ESAC). If the method is appropriate for chemicals testing, a draft Annex V guideline, incorporating the method, will be submitted to the EU Competent Authorities for Directive 67/548/EEC for consideration for regulatory acceptance and application.

After a formal validation has been completed, the next step is an independent evaluation of the goals, design, management, performance and outcome of the study, by one or more appropriate agencies, together with the publishing of a report in the peer-reviewed literature (Balls and Fentem, 1999; Goldberg and Locke, 2004).

Efforts to define and coordinate the validation process have been made both in the United States and in Europe (Zurlo, 1998), and its scientific characteristics have been discussed on several occasions, being at the centre of some controversies as well (Balls, 1992; Flint, 1992; Green, 1992; Walum, 1992). However, today there is a wide literature on the principles of validation (Balls, et al., 1990a; Balls, et al., 1995; ICCVAM, 1997; OECD, 1996, 2005;
Walum, et al., 1994), even if the current reality recognizes that validation is only the first step. It can lead to scientific and industrial acceptance of the test and to the drafting of a proposed regulatory guideline, marking the beginning of the progression of the test towards regulatory acceptance and application (Balls and Fentem, 1999), but the path from validation to acceptance and implementation remains a long and arduous process (Goldberg, 2007).

### 1.5.3. Regulatory acceptance

A formal validation study is an important step on the way of regulatory application and acceptance of a new methodology (Balls and Fentem, 1999). The formal adoption of a validated test method by a regulatory agency/authority, which has to be considered also during the design and conduct of the validation study, is defined as "Regulatory Acceptance" of new or revised toxicological test methods, whose principles and criteria were developed simultaneously with validation criteria. These regulatory acceptance criteria (OECD, 2005) involve for example a transparent and independent peer review process of test method and validation study data; the fact that the test method should generate data useful for hazard/risk assessment purposes; that applicability and limitations of the test method should be clearly described; and also that the test method should be time and cost effective and likely to be used in a regulatory context. What it has to be considered, however, is that, in some cases, such as in the U.S., validation does not guarantee regulatory acceptance. Regulators do not necessarily require that agencies adopt formal validated alternatives; each agency can decide for itself whether a validated alternative test will be acceptable under its regulatory programs. Authorities simply need to be convinced that the proposed method performs to its intended use in measuring the endpoints in question for regulatory approval (Goldberg and Locke, 2004; Wilcox and Goldberg, 2011). Anyway besides this, as it was concluded at a workshop held in Vouliagmeni, Greece, in 1990 "All conceivable and practicable steps should be taken to make the formal acceptance and incorporation on non-animal toxicity test procedures into regulatory practice as smooth and rapid a process as possible" (Balls and Fentem, 1999; Balls, et al., 1990b).
### 1.5.4. Validated and accepted alternative methods

In **Table 1.5** taken from AltTox.org (http://alttox.org/ttcr/validation-ra/validated-ra-methods.html), a very useful website dedicated to advancing non-animal methods of toxicity testing, all the already validated and accepted alternative methods, updated at 27 September 2011 are listed.

**Table 1.5. Validation and Regulatory Acceptance Status of Alternative Test Methods and Testing Strategies.**

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Method Name</th>
<th>Test Type</th>
<th>Endorsement of Scientific Validity</th>
<th>Regulatory Acceptance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute mammalian toxicity (inhalation)</td>
<td>Acute toxic class method</td>
<td><em>In vivo</em></td>
<td></td>
<td>OECD TG 436 (2009)</td>
</tr>
<tr>
<td>Fixed concentration procedure</td>
<td></td>
<td><em>In vivo</em></td>
<td></td>
<td>Draft TG OECD 433</td>
</tr>
<tr>
<td></td>
<td>Toxin binding inhibition test for</td>
<td><em>In vitro</em></td>
<td>ESAC 10</td>
<td>European</td>
</tr>
<tr>
<td>Test Type</td>
<td>Endpoint Description</td>
<td>Species</td>
<td>Reference</td>
<td>Notes</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>---------------------------------------------------------------------------------------</td>
<td>---------</td>
<td>-----------</td>
<td>-------</td>
</tr>
<tr>
<td>Endocrine active substances(^3)</td>
<td>Androgen receptor binding assay (rat prostate cytosol)</td>
<td>Ex vivo</td>
<td></td>
<td>OPPTS TG 890.1150 (EPA, 2009)</td>
</tr>
<tr>
<td></td>
<td>Aromatase inhibition assay (human recombinant)</td>
<td>In vitro</td>
<td></td>
<td>OPPTS TG 890.1200 (EPA, 2009)</td>
</tr>
<tr>
<td></td>
<td>Estrogen receptor (ER)-alpha transcriptional activation assay for estrogen agonists (STTA)</td>
<td>In vitro</td>
<td>OECD/EPA</td>
<td>OPPTS TG 890.1300</td>
</tr>
<tr>
<td></td>
<td>Estrogen receptor binding assay rat uterine cytosol (ER-RUC)</td>
<td>Ex vivo</td>
<td></td>
<td>OPPTS TG 890.1250 (EPA, 2009)</td>
</tr>
<tr>
<td></td>
<td>H295R steroidogenesis assay</td>
<td>In vitro</td>
<td>OECD/EPA</td>
<td>OPPTS TG 890.1550 (EPA, 2009)</td>
</tr>
<tr>
<td></td>
<td>US EPA Tier 1 Screening Battery</td>
<td>In vivo/in vivo</td>
<td></td>
<td>US EPA (2009)</td>
</tr>
<tr>
<td></td>
<td>BIG1Luc ER TA test method for estrogen agonists and antagonists</td>
<td>In vitro</td>
<td>ICCVAM (expected 2011)</td>
<td></td>
</tr>
<tr>
<td>Eye corrosion</td>
<td>Bovine corneal opacity permeability (BCOP) test</td>
<td>In vitro</td>
<td>ICCVAM (2007)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ESAC 27 (2007)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>INVITOX Protocol 127</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>JacVAM (2009)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ICCVAM Regulatory Acceptance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OECD TG 437 (2009)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OECD Pesticide Proficiency Standards</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Draft OECD GD on Supplement to TG 437 and</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Name</td>
<td>Type</td>
<td>Source Code(s)</td>
<td>Location</td>
<td>Notes</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------</td>
<td>----------</td>
<td>----------------------------------------------------</td>
<td>--------------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>INVITTOX Protocol 192 modified</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluorescein Leakage (cytotoxicity/cell-based assay)</td>
<td>In vitro</td>
<td>ESAC 40 (2009)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hen’s egg test-chorioallantoic membrane (HET-CAM)</td>
<td>In vitro</td>
<td></td>
<td></td>
<td>EU Competent Authorities for Dangerous</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Substances Directive</td>
</tr>
<tr>
<td></td>
<td>or Ex</td>
<td>INVITTOX Protocol 80</td>
<td></td>
<td>OECD Proficiency Standards</td>
</tr>
<tr>
<td></td>
<td>vivo</td>
<td>ICCVAM (2007)</td>
<td></td>
<td>Draft OECD GD on Supplement to TG 437 and</td>
</tr>
<tr>
<td>Isolated rabbit eye test</td>
<td>In vitro</td>
<td>ICCVAM (2009)</td>
<td></td>
<td>EU Competent Authorities for Dangerous</td>
</tr>
<tr>
<td></td>
<td>or Ex</td>
<td></td>
<td></td>
<td>Substances Directive</td>
</tr>
<tr>
<td></td>
<td>vivo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>endpoints</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cytosensor Microphysiometer modified (cytotoxicity/cell-function based</td>
<td></td>
<td>INVITTOX Protocol 192 modified</td>
<td></td>
<td></td>
</tr>
<tr>
<td>in vitro assay</td>
<td></td>
<td>ICCVAM (2010)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotype</td>
<td>Test Name</td>
<td>Species</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>---------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>Genotoxicity</td>
<td>Bacterial reverse mutation (Ames) test</td>
<td>in vitro</td>
<td>OECD TG 471 (1997)</td>
<td></td>
</tr>
<tr>
<td>Genotoxicity</td>
<td>In vitro cell gene mutation test</td>
<td>in vitro</td>
<td>OECD TG 476 (1997)</td>
<td></td>
</tr>
<tr>
<td>Genotoxicity</td>
<td>In vitro chromosomal aberration test</td>
<td>in vitro</td>
<td>OECD TG 473 (1997)</td>
<td></td>
</tr>
<tr>
<td>Genotoxicity</td>
<td>In vitro mammalian cell micronucleus test</td>
<td>in vitro</td>
<td>ICCVAM comments to OECD</td>
<td></td>
</tr>
<tr>
<td>Genotoxicity</td>
<td>In vitro sister chromatid exchange test</td>
<td>in vitro</td>
<td>OECD TG 479 (1986)</td>
<td></td>
</tr>
<tr>
<td>Genotoxicity</td>
<td>In vitro unscheduled DNA synthesis test</td>
<td>in vitro</td>
<td>OECD TG 482 (1986)</td>
<td></td>
</tr>
<tr>
<td>Genotoxicity</td>
<td>Saccharomyces cerevisiae gene mutation assay</td>
<td>in vitro</td>
<td>OECD TG 480 (1986)</td>
<td></td>
</tr>
<tr>
<td>Genotoxicity</td>
<td>Saccharomyces cerevisiae mitotic recombination assay</td>
<td>in vitro</td>
<td>OECD TG 481 (1986)</td>
<td></td>
</tr>
<tr>
<td>Hematotoxicity: acute neutropenia</td>
<td>Colony forming unit granulocyte macrophage (CFU-GM) assay</td>
<td>in vitro</td>
<td>Submitted to EMEA</td>
<td></td>
</tr>
<tr>
<td>Immunotoxicity/Skin Sensitization</td>
<td>Local lymph node assay (LLNA)</td>
<td>in vivo</td>
<td>ICCVAM (1999)</td>
<td></td>
</tr>
<tr>
<td>Immunotoxicity/Skin Sensitization</td>
<td></td>
<td></td>
<td>Updated OECD TG 429 (2010)</td>
<td></td>
</tr>
</tbody>
</table>

<sup>1</sup> Updated OECD TG 429 (2010)
<table>
<thead>
<tr>
<th>Category</th>
<th>Assay Description</th>
<th>In/Ex vivo</th>
<th>Reference(s)</th>
<th>Regulation/Agency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonradiolabeled LLNA</td>
<td>LLNA for Potency Categorization</td>
<td>In vivo</td>
<td>ICCVAM (2009)</td>
<td>UN GHS (2009); US agencies (anticipated 2011)</td>
</tr>
<tr>
<td>Phototoxicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyrogenicity</td>
<td>Human whole blood IL-1</td>
<td>In vitro</td>
<td>ESAC 19 (2006)</td>
<td>EMEA; European Pharmacopoeia; US agencies</td>
</tr>
<tr>
<td></td>
<td>Human whole blood IL-6</td>
<td>In vitro</td>
<td>ICCVAM (2008)</td>
<td>EMEA; European Pharmacopoeia; US agencies</td>
</tr>
<tr>
<td></td>
<td>Human cryopreserved whole blood IL-1</td>
<td>In vitro</td>
<td>ESAC 23 (2006)</td>
<td>EMEA; European Pharmacopoeia; US agencies</td>
</tr>
<tr>
<td></td>
<td>PBMC IL-6</td>
<td>In vitro</td>
<td>ESAC 21 (2006)</td>
<td>EMEA; European Pharmacopoeia; US agencies</td>
</tr>
<tr>
<td></td>
<td>MM6 IL-6</td>
<td>In vitro</td>
<td>ESAC 22 (2006)</td>
<td>EMEA; European Pharmacopoeia; US agencies</td>
</tr>
<tr>
<td></td>
<td>Limulus amebocyte lysate (LAL) test</td>
<td>In vitro</td>
<td></td>
<td>European Pharmacopoeia (5.6); US Pharmacopoea (85)</td>
</tr>
<tr>
<td></td>
<td>Micromass embryotoxicity assay</td>
<td>Ex vivo</td>
<td>ESAC 11 (2002)</td>
<td></td>
</tr>
<tr>
<td>Test Category</td>
<td>Assay Details</td>
<td>In Vitro</td>
<td>In Vivo</td>
<td>Referee</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------------------------------------------------</td>
<td>----------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Method Description</td>
<td>Type</td>
<td>Standards</td>
<td>EU Test Method</td>
<td></td>
</tr>
<tr>
<td>------------------------------------</td>
<td>-------</td>
<td>------------------------------------------------</td>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td>Epiderm skin irritation test (with MTT reduction)</td>
<td>In vitro</td>
<td>ESAC 30 (2007)13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. All in vitro and ex vivo methods listed are new methods proposed to reduce or refine animal use also listed.
2. The methods are not intended for initial dose setting, but in vivo test required to complete assessment.
3. The screening assay is to be used as a part of a tiered testing strategy.
4. Recommended for use as initial step in a Top-Down Approach to identify ocular irritants (EU Cat A1, OECD Category I) and EPA Category I for water-soluble chemicals and as initial step in a Top-Down Approach to identify the irritants (EPI-200, OECD Category I) for water-soluble substances and water-soluble substances containing metals. Ocular irritants are further refined with respect to classification and labelling.
5. Can be used for screening to distinguish water-soluble surfactant-free preparations and certain types of surfactant-containing formulations that are not labelled as irritants (e.g., EPA Category IV, RV Not Labelled). PNLAS are listed from all relevant categories (e.g., EPA Category I, II, III, IV, OECD Category I) for hazard classification and labelling under EU, US and HMG-governed systems. High dose high safety (HDS-HS) for non-irritant substances and formulations. High dose (positive at 100% to 10% dose) for substances not labelled as irritants. Can be used as a screening test to identify ocular irritants in a Top-Down approach, as part of weight of evidence approach. Negative results need to be tested in another test method.
6. Recommended for use as initial step in a Top-Down Approach to identify ocular irritants and water-soluble irritants (EU Cat A1, OECD Category I) and EPA Category I for water-soluble chemicals. Further refinement with respect to classification and labelling is recommended.
7. Routines include use of avoiding data only.
8. Retrospective DTT data can be used in a weight-of-evidence approach to identify potential ocular irritants.
9. IUCN can be used for hazard classification when dose response information is not available.
11. Subject to product-specific validation to demonstrate equivalence to the rabbit pyrogen test (RPT).
12. Recommended as screening test or as part of sequential testing strategy: only positive test results accepted in the 2007 endorsement.
1.6. The legislative implementation of the 3RS principle

As already discussed in Paragraph 1.3.2, the principles outlined by Russell and Burch (1959) received little attention until the animal welfare movement became a more important societal concern (Zurlo, 1998), and the Universal Declaration of the Rights of Animals (UNESCO, 1978) was promulgated. Since then, active campaigns have been undertaken against manufacturers of cosmetics, household and personal care products and pharmaceuticals in attempts to stop animal testing, and significant pressure on government agencies took place to apply more stringent regulations on chemicals (Zurlo, 1998). As a consequence, several regulatory authorities have supported the principle of the 3Rs and several laws and directives regarding replacement, reduction, and refinement alternatives in scientific research throughout the world over the past decades. There are some primary events that had a profound impact on the research and development of alternative methods in toxicological testing, and are driving the need to examine the use of new and developing science in addition to traditional animal model approaches of safety testing worldwide. They are independent events, significantly different in their intended goals, that however have converged from vastly different directions and resulted in a common outcome: to drive the eventual replacement of animal testing (Wilcox and Goldberg, 2011).

The first European Directives concerning safety in biotechnology experimentation date back to 1986 (Directive 86/609/EEC) (Council of the European Communities, 1986). Such Directive has several relevant aspects for the field of alternatives and, at least in theory, it represented a very strong driver for alternative methods. However, in practice there was very little enforcement of it, and that is why it was recently revised in 2010 (Directive 2010/63/EU) (European Parliament, 2010), with the introduction of inspections and other reinforcements (Hartung, 2010, 2011). Regarding the aspects related to the field of alternatives, in these documents it is stated that: "An experiment shall not be performed if another scientifically satisfactory method of obtaining the result sought, not entailing the use of an animal, is reasonably and practicably available". The procedure which uses the minimum number of animals and which causes the least pain, suffering, distress or lasting harm must be chosen, and all experiments shall be carried out under general or local anaesthesia. Moreover, Article 23 of the 1986 Directive states: "The Commission and
Member States should encourage research into the development and validation of alternative techniques which could provide the same level of information as that obtained in experiments using animals but which involve fewer animals or which entail less painful procedures, and shall take such other steps as they consider appropriate to encourage research in this field." This article resulted to be very important, since it encouraged a substantial funding program for alternatives and their validation by the EU. More than € 250 million, indeed, was spent by the European Commission, and a similar amount was estimated to have been spent by the Member States (Devolder, et al., 2008; Hartung, 2010, 2011; Rusche, 2003).

After that first Directive, the EU has shown a strong interest in animal well-being and has promulgated directives concerning animal protection in breeding (Directive 98/58/EC) (Council of the European Union, 1998), transport (Directive 95/29/EC) (Council of the European Union, 1995), slaughter (Directive 93/119/EC) (Council of the European Communities, 1993a), and the ill-treatment of animals as well (e.g. art. 727 of the Italian Penal Code). Also the Italian Legislative Decree 116/92 (Italy, 1992) can be added to these regulations, since it imposes some specific rules for the protection of animals used for experimental or other scientific purposes, and considers as an offence any experimentation on higher animals or any that affects their genetic identity, in particular if the experiment will inflict unnecessary pain and suffering (Passantino, et al., 2004).

Of primary importance was also the adoption in 2003, by the EU institutions, of one of the most rigorous legislation enforcing the use of alternative methods, the 7th amendment (Directive 2003/15/EC) (European Parliament, 2003) of the 1976 European Cosmetics Directive (Directive 76/768/EEC) (Council of the European Communities, 1976; Hartung, 2008a). It is stated in the 7th amendment: "Currently, only alternative methods which are scientifically validated by the ECVAM or the OECD and applicable to the whole chemical sector are systematically adopted at Community level. However, the safety of cosmetic products and their ingredients may be ensured through the use of alternative methods which are not necessarily applicable to all uses of chemical ingredients. Therefore, the use of such methods by the whole cosmetic industry should be promoted and their adoption at Community level ensured, when such methods offer an equivalent level of protection to consumers" (Hartung, 2011). Importantly, such Directive set out a timetabled schedule for banning animal testing for personal care products, permitting only the use of replacement alternatives. A complete ban on the use of animals for testing cosmetic products and their
ingredients on European territory, came into force on 11 March 2009. The marketing ban of products or cosmetic ingredients tested on animals became effective on the same date, except with regard to the tests involving repeated-dose toxicity, toxicokinetics and reproductive toxicity (the most complex endpoints) for which a deadline of ten years after entry into force of the Directive (11 March 2013) is foreseen.

Similarly to the 7th amendment on cosmetics, also the Regulation (EC) No 1907/2006, better known as REACH Regulation (European Parliament, 2006, 2007) and adopted by the European Council and the European Parliament in December 2006, aims at avoiding animal testing and taking preference of alternative methods to animal testing as far as possible (Lilienblum, et al., 2008). This regulation is the first one in Europe to combine the issue of the 3Rs (therefore the need for the replacement, reduction, and refinement of animal experiments) with the need to implement the knowledge about the hazard of chemicals and the willing to harmonize the testing requirements for existing chemicals for which there is a lack of safety data and new ones (Zuang and Hartung, 2005). Because of its primary importance in the treated issue, REACH Regulation will be discussed in more details in Chapter 2.

1.7. Numbers of animals used

As a result of all the above mentioned laws and regulations, the development of test methods for characterizing human exposure to chemicals of toxicological concern have proceeded rapidly and will continue to do so in the foreseeable future (Wilcox and Goldberg, 2011). However, despite all the steps made in the field of in vitro models aiming at the reduction of the total number of animals used for experimental and other scientific purposes, that number is still quite high. Indeed, as reported in the 6th Commission Report on the Statistics on the number of animals used for experimental and other scientific purposes, that includes data from all the 27 Member States of the EU, in 2008 it was just above twelve million (European Commission, 2010a).
Mammals are often used as experimental species and, as also stated in the Report, rodents and rabbits represented more than 80% of the total number used in Europe. Among them, mice are by far the most commonly used species, accounting for 59% of the total use, followed by rats with 18%. Cold-blooded animals represent instead almost 10% of the total, and are considered as the second most used group, while the third largest one was represented by birds (a little over 6% of the total) (cf. Figure 1.2).

**Figure 1.2. Percentages of animals used by classes in the Member States.**

Among the total number of animal used for experimental purposes in many different areas, toxicological and other safety evaluations for different products and environmental tests represented around the 8,7% of the total, i.e. more than one million of animals (cf. Figure 1.3).
Figure 1.3. Purposes of experiments.

Of this total, animals used for toxicological or other safety evaluations of products or devices for human and veterinary use represented the largest number (almost 51%). The percentage of animals used for toxicological evaluation additives in food for human consumption, cosmetics and household products, is very small (1.18%) when compared to the other product groups, thanks to new regulations, while in the case of evaluation of industrial and agricultural products, the numbers are respectively represented by 7.1% and 7.9% of the number of animal used for that purpose. There are then other toxicological and safety evaluations that represent more than 21% (cf. Figure 1.4).
There is a clear decrease in the number of animals used for toxicological tests for products intended for industry, for agriculture and for potential contaminants of the environment, in comparison to the data submitted in the previous statistical report (European Commission, 2007a). There is also a significant decrease, represented by a 65% drop, in the number of animals used for testing of products for cosmetics and toiletries, that should be seen in light of the legal requirement to phase out animal testing for cosmetics in Europe.

The development of alternative methods is thus of high priority in order to reduce the number of animals that are nowadays considered to be necessary for safety evaluation of human health and environment (Rogiers, 2005).
Chapter 2
REACH, a new chemicals policy for the EU

“REACH is all about protecting human health and the environment. The challenge is to have scientifically sound information on the potential hazards of substances whilst at the same time minimising unnecessary animal testing. One of the fundamental aims of REACH is to promote alternative methods for assessing hazards of substances and to see animal testing as a last resort. All parties involved should take this very seriously.”

- Geert Dancet, Executive Director of ECHA -

2.1. Introduction

REACH is a new European Community Regulation on chemicals and their safe use (Regulation (EC) No 1907/2006) (European Parliament, 2006, 2007), which concerns the Registration, Evaluation, Authorisation and Restriction of CHemicals (from what derived the acronym REACH), and replaces different European Directives and Regulations on chemicals (more than 40 existing pieces of legislation) with a single system, in order to optimize and improve the former legislative framework of the European Union (EU).

A large number of substances, indeed, have been manufactured and placed on the market in Europe for many years along the course of history, sometimes in very high amounts, but there was insufficient information about most of chemicals and the hazards that they pose to human health and the environment (Danish Board of Technology, 1996). With regard to these chemicals, an important distinction between "existing" and "new" chemicals, based on the cut-off date of 1981, was introduced with the Regulation (EEC) 793/93 of 23 March 1993 on the evaluation and control of the risks of existing substances (Council of the European Communities, 1993b). All the chemicals listed in the European Inventory of Existing Commercial Chemical Substances (EINECS) (created by the European Community Commission Decision 81/437/EEC and including more than 100,000 chemicals) (Commission of the European Communities, 1981) and reported to be on the European market between 1 January 1971 and 18 September 1981 were called "existing"; on the contrary, all those chemicals introduced after 1981 and listed in the European List of Notified
Chemical Substances (ELINCS) (Sauer, 2004) were defined as "new" chemicals. The problem was that, while all of these "new" substances, before being placed in the market, had to be tested quite rigorously, the "existing" ones, which represented 99% of the market, were not subject to such requirements. Therefore, it is clear how it was relatively costly to introduce a new substance on the market, and this of course represented a brake to the expansion and innovation of EU chemicals industry, because the research of new compounds was discouraged, in favour of the development and use of existing substances (European Commission, 2006). At the same time, because there was a general lack of knowledge on properties and uses of "existing" substances, the process of identification and risk assessment was slow and resource-intensive, thus not allowing the system to work efficiently and effectively. Moreover, also the introduction of risk management measures was delayed, leaving EU industries of chemicals behind their U.S. and Japanese counterparts. As a consequence, there was an obvious need to fill these information gaps to ensure that industry was able to assess hazards and risks of the substances, and to identify and implement the risk management measures to protect both humans and the environment (Commission of the European Communities, 1998; Sauer, 2004). Because of this awareness, a period of intense negotiations, discussions and highly controversial debate started, leading to the amendment of the new law in 2006, which has been considered for its importance, an historical date (Fuchs, 2009).

REACH Regulation, which officially came into force the 1st June 2007 (ECHA, 2007), replaced old legislations, encouraging innovation of safer substances and establishing a system for assessing both existing and new chemicals. Specifically, it has several aims, but the most important ones, concerning this thesis work, are to improve the protection of human health and the environment from the risk of hazardous chemicals, and to promote alternative methods for the assessment of hazards of such substances (a list of all the Regulation recitals and articles on this matter is provided in Appendix 1). Other aims of primary importance are maintaining and enhancing the competitiveness of the EU chemicals industry, which is a key sector for the economy of the whole EU, as well as promoting its integration with international efforts, ensuring the free circulation of substances on the internal EU market by increasing transparency, and preventing its fragmentation. Moreover, the Regulation aims to the progressive substitution of the most dangerous chemicals when suitable alternatives have been identified.
In order to do that, REACH Regulation places greater responsibility on industry to manage the risks posed by chemicals to the health and the environment, and to provide appropriate safety information on the substances to their users prior to production and marketing, while the role of authorities should be to ensure that industries meet their obligations and take actions on substances of very high concern or where there is a need for community action. On the contrary, the previous legislation required public authorities to identify and address possible safety issues for the chemicals on the market, and what was required was a comprehensive risk assessment, rather than a more specific one. This shift of responsibility towards producers and importers is a crucial element of REACH, and sometimes defined as a "self-responsibility approach" (Fuchs, 2009; Führ and Bizer, 2007). Moreover, another important point is that all manufacturers, importers and also for the first time downstream users of chemicals are required to collect information on the properties of their chemical substances and identify and manage the risks, facilitating their safe handling.

For these purposes, REACH required also the set-up of the European Chemicals Agency (ECHA), based in Helsinki, Finland, which started operations on 1 June 2008 and acts autonomously as the central point in the REACH system, coordinating and implementing role in the overall process. The Agency run the databases necessary to make the system operative, coordinates the evaluation of suspicious chemicals and it is also building up a public database where consumers and professionals can find hazard information. The Agency will therefore manage the technical, scientific and administrative aspects of the REACH system at a Community level, aiming to ensure that the legislation can be properly implemented and has credibility with all stakeholders.

2.2. Scope: REACH and chemicals

REACH Regulation is very wide in its scope, indeed it covers all substances (both new and existing ones) whether manufactured, imported, used as intermediates or placed on the market in the EU, either on their own, in preparations or in articles, unless they are radioactive, subject to customs supervision, or are non-isolated intermediates. These substances are excluded from the REACH Regulation because of some articles (Article 2, 56 and 67) and some of these exemptions are made on the basis of other equivalent legislation

A review of the scope of the Regulation has been foreseen by the Commission five years after entry into force, as stated in Article 138(6) of REACH: "By 1 June 2012 the Commission shall carry out a review to assess whether or not to amend the scope of this Regulation to avoid overlaps with other relevant Community provisions. On the basis of that review, the Commission may, if appropriate, present a legislative proposal."

### 2.3. How does reach work: REACH processes

REACH, as already said, stands for Registration, Evaluation, Authorisation and Restriction of Chemicals, and its entire process is organized in different phases and deadlines.

#### 2.3.1. Pre-Registration

The process began with a mandatory pre-registration deadline of December 2008. Manufacturers and importers of chemicals had to pre-register substances already present on the EU market (the so-called phase-in substances), in order to benefit from transitional arrangements, and also to share data with other registrants, avoiding therefore carrying out redundant tests. The pre-registration period was limited from 1 June 2008 to 1 December 2008, and a total of 180,000 pre-registrations were expected by about 27,000 companies for 30,000 substances. ECHA, however, received more than 2.7 million pre-registrations from about 65,000 companies for 144,000 substances (Hartung, 2011).
2.3.2. Registration

The 2nd scheduled phase was Registration, that first applied from 1 June 2008. REACH requires manufacturers and importers to obtain information on the physicochemical, health and environmental properties of their substances, to use them to determine how these substances can be used safely, and then to submit a registration dossier documenting data and assessments to the ECHA Agency. These requirements are used for substances produced or imported in quantities of 1 tonne or more per year per company (threshold much higher than the previous one of 10 kg for new substances). Quantities above 10 tonnes per year, instead, additionally require the submission of a Chemical Safety Report (CSR) in order to document the safety assessment of the substance (European Commission, 2006). Test requirements progressively increase with the volume of the specific chemical substance. For instance acute toxicity tests are required for substances produced or imported between 10 and 100 tonnes per year, subchronic toxicity tests for substances between 100 and 1000 tonnes per year and chronic toxicity/carcinogenicity studies for substances >1000 tonnes per year. Table 2.1, taken from Hengstler et al. (2006), shows all the required tests for the different categories. From a practical point of view, the detailed compilation of the required tests is one of the most important parts of the REACH proposal.

2.3.3. Evaluation

Once the registration dossier has been received, the Agency has then to perform two types of evaluation with different aims. Firstly, it has to perform a dossier evaluation, to check if information provided by industries is compliant with the requirements, and to assess the testing proposals made by the registrant, in order to ensure that the assessment of the chemical substances will not result in unnecessary testing, especially on animals. The Agency will also coordinate substance evaluation, which has to be conducted by the Member States to investigate chemicals of concern, in order to clarify suspicions of risks to human health or the environment by requesting further information from industry. This assessment will start in 2012 and it may lead authorities to the conclusion that action needs to be taken under the restrictions or authorisation procedures in REACH, or that information needs to be passed on to other authorities responsible for relevant legislation.
Table 2.1. Tests required by the REACH concept.

<table>
<thead>
<tr>
<th>Tonnage of production per year</th>
<th>Required tests</th>
<th>Changes in the concept</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1</td>
<td>Not covered by REACH</td>
<td>No tests for most compounds</td>
</tr>
<tr>
<td>≥1 to &lt;10</td>
<td>Irritation/corrosion in vitro, Short-term toxicity on daphnia</td>
<td>Tests required only for some compounds</td>
</tr>
<tr>
<td>≥10 to &lt;100</td>
<td>Eye/skin irritation, in vivo sensitisation, local lymph node acute toxicity, oral (acute toxicity inhalation, dermal) subacute toxicity (28 days), mutagenicity, bacteria cytogenicity, mammalian cells gene mutation, mammalian cells (reprotoxicity, development, 2-generation study), short-term toxicity on daphnia, short-term growth of algae, degradation, biotic, abiotic</td>
<td>Plus toxicity, acute, Plus biodegradation, Plus growth of algae</td>
</tr>
<tr>
<td>≥100 to &lt;1000</td>
<td>Eye/skin irritation, in vivo sensitisation, local lymph node (toxicokinetics) acute toxicity, oral (acute toxicity inhalation, dermal) subacute toxicity (28 days), subchronic toxicity (90 days), mutagenicity, bacteria cytogenicity, mammalian cells gene mutation, mammalian cells reprotoxicity, development reprotoxicity, 2-generation study, toxicity on daphnia (21 days), short-term growth of algae, acute toxicity on fish, toxicity on fish (life-cycle test), degradation, biotic, abiotic sedimentation, bioaccumulation, effect on earthworm, acute effect on microorganism, growth of plants</td>
<td>No test on reprotoxicity</td>
</tr>
<tr>
<td>≥1000</td>
<td>Eye/skin irritation, in vivo sensitisation, local lymph node (toxicokinetics) acute toxicity, oral (acute toxicity inhalation, dermal) subacute toxicity (28 days), subchronic toxicity (90 days), mutagenicity, bacteria cytogenicity, mammalian cells gene mutation, mammalian cells (carcinogenesis) reprotoxicity, development reprotoxicity, 2-generation study, toxicity on daphnia (21 days), short-term growth of algae, acute toxicity on fish, toxicity on fish (life-cycle test), degradation, biotic, abiotic sedimentation, bioaccumulation effects on earthworm, long-term effects on microorganism, effects on invertebrates, birds, plants</td>
<td></td>
</tr>
</tbody>
</table>

Italics in parenthesis indicate optional tests that are required when data indicate a risk.

a. Per producer or importer.
b. According to the EU Commission, October 2003.
c. According to the consolidated draft of the Council of the European Union.
d. Compound which meet the criteria defined in ANNEX (Ic).

2.3.4. Authorisation

REACH also provides for an authorisation system for the use of substances of very high concern (i.e. carcinogens, mutagens, toxic to the reproductive system, bio-accumulative and so on), because of their very serious and normally irreversible effects on humans and the environment. It is therefore essential to regulate them centrally, through a mechanism that ensures that the risks related to their actual uses are assessed, considered and then decided
upon by the Community. Therefore, the aim of the authorisation system is to ensure that risks associated with these substances are adequately controlled, and that they will be progressively substituted by safer substances or technologies, when technically and economically feasible, or only used where there is an overall benefit (i.e. socio-economic benefits) for the society that outweigh the risks.

2.3.5. Restriction

In addition, EU authorities may impose restrictions and prohibit or set conditions for the manufacture, placing on the market or use of certain dangerous substances or group of substances when unacceptable risks to humans or the environment have been identified. The use of certain dangerous chemicals is acceptable as long as appropriate risk management measures are implemented, such as the use of good ventilation or protective clothing. If measures at company level are not sufficient to keep the risks acceptable, restrictions and limitations can be imposed for substances in certain circumstances and for certain uses (for instance by consumers), or even complete bans on all uses (European Commission, 2006, 2007b). Thus, the aim of restrictions is to act as a safety net in order to manage risks that are otherwise not adequately controlled.


2.3.6. Classification and labelling inventory

Another important point for chemicals legislation in Europe has always been the requirement for industry to classify and label dangerous substances and preparations according to standard criteria. The aim of this requirement is ensuring that hazard classifications and consequent
labelling of all dangerous substances manufactured in or imported into the EU are available to all, in order to promote agreement on the classifications. Indeed, most divergences between classifications of the same substance should be removed over time, either through cooperation between notifiers and registrants, or by EU harmonised classifications for substances.

Regarding this matter, REACH builds on existing legislation, and the European Commission adopted on 27 June 2007 a new proposal for a Regulation on classification, labelling and packaging of substances and mixtures (Commission of the European Communities, 2007), that incorporates the classification criteria and labelling rules agreed at United Nations level, the so-called Globally Harmonised System of classification and labelling of chemicals (GHS) (ONU, 2009).

### 2.3.7. Communication in the supply chain

Communication in the supply chain is another important requirement of REACH. The aim is to ensure that not only manufacturers and importers, but also their customers (i.e. downstream users and distributors), have the information they need in order to use chemicals safely, and therefore information relating to health, safety and environmental properties, risks and risk management measures are required to be passed both down and up the supply chain. This will be done via the well-established and familiar Material Safety Data Sheet (MSDS) for all dangerous substances. Indeed, the provisions of the current Safety Data Sheets Directive (Directive 91/155/EEC) (Commission of the European Communities, 1991a) were carried over into the REACH. As more information on hazardous properties and information that challenges the quality of risk management measures will be available as a result of registrations, the quality of MSDSs will be improved, and also relevant exposure scenarios need to be annexed to them and have thus to be passed down the supply chain.
2.4. Timeline for REACH implementation

REACH aimed to ensure a smooth transition from the previous legislation, and therefore will be implemented gradually until May 2018. Appropriate deadlines for the repeal of various aspects of the previous legislation and for the phasing in of various provisions of REACH, indeed, has been set, with the aim of not wasting the work undertaken under the previous legislation.

The scheduled deadlines are listed in what follows:

1 June 2007: Entry into force of REACH.

1 June 2008: European Chemicals Agency became operational.

1 June 2008 to 1 December 2008: Pre-registration of so-called "phase-in substances".

30 November 2010: Registration deadline for substances produced or imported in quantities of 1000 tonnes and above, as well as carcinogens, mutagens and substances toxic to reproduction (CMR category 1 and 2) equal to or greater than 1 tonne/year, and substances classified as very toxic to aquatic organisms (R50/53) at and above 100 tonnes/year.

31 May 2013: Registration deadline for substances produced or imported in quantities equal to or greater than 100 tonnes/year.

31 May 2018: Registration deadline for substances produced or imported in quantities equal to or greater than 1 tonne/year (European Commission, 2006, 2007b).

The yearly tonnage refers to metric tons per manufacturer or importer, not to the total volume manufactured or imported (Rudén and Hansson, 2010).
2.5. Impact assessment

At the time when the REACH Regulation was negotiated and adopted (2003-2006), the European Commission conducted various studies to assess its possible impact. Before adopting its proposal on REACH in October 2003, the Commission published an Extended Impact Assessment (Commission of the European Communities, 2003), which assessed the benefits for health and the environment, as well as the costs for industry and the Agency.

2.5.1. Benefits

The most important benefits expected from REACH will be benefits to health and the environment, mainly due to the expected improvement of the risk management. A comprehensive quantitative assessment of the wide range of environmental and health impacts, however, is not possible, due to a lack of data.

An important benefit of REACH is that the hazards and risks of chemicals are more systematically identified. This allows for more effective risk management measures by industries and enterprise producing, handling and using these substances, and for end-users exposed to these substances, providing at the same time also a benefit for consumers, who will have access to more information on the potential hazards and risks of chemicals (European Commission, 2006, 2008). Authorities as well will benefit from the information gathered, as it should eventually lead to a better management of the risks associated with individual substances and uses, potentially yielding added benefits for the workforce exposed to these chemicals and for society at large, including an improved status of the environment. It will be possible indeed to ensure more effective implementation of the precautionary principle, being alerted at an earlier stage to potential risks and acting more rapidly to address the problem.

REACH will contribute also to a positive occupational impact (resulting in health benefits for workers, both in chemicals and downstream sectors) and public health impact, because it is well known that the exposure to chemicals is linked with a considerable number of diseases, including respiratory and bladder cancers, mesothelioma, skin disorders, respiratory diseases, eye disorders, asthma and others, even if most harmful effects are the result of many causes
acting together, such as genetics, lifestyle, radiation, diet, pharmaceuticals, chemicals (manufactured and natural), smoking and air pollution, including indoor and outdoor exposures. Moreover, sensitive groups, such as the elderly, children, the embryo, the sick, and pregnant women, may be affected at much lower doses than others. However, it is frequently very difficult to link diseases to particular chemicals and estimate the aggregate health impacts, but despite of this, the anticipated benefits are generally expected to be significant. In the Extended Impact Assessment of the Commission proposal (Commission of the European Communities, 2003), diseases caused by chemicals were assumed to account for some 1% of all types of disease in the EU. Assuming a 10% reduction in these diseases as a result of REACH (implying that 90% of the health impacts associated with chemicals are either related to historical exposures, will not be identified by REACH or cannot be tackled), a 0.1% overall reduction would result in the EU. This would be equivalent to around 4,500 deaths due to cancer being avoided every year. On the basis of a € 1 million value of life, the potential health benefits of REACH were then estimated approximately € 50 billion over a 30 year period. This figure was based on an illustrative scenario developed with the support of recognised international organizations, such as the World Bank and World Health Organisation, and it is not an estimate of the benefits of REACH, but rather an illustration of their potential scale.

It is clear that exposure to hazardous chemicals also damages the environment. However, due to lack of data, also in this case it is not possible to provide, as already mentioned, a comprehensive quantitative assessment of the impacts on the environment. Much of the information needed will only be available after the chemicals on the market today have been tested and registered in line with the requirements of REACH (Commission of the European Communities, 2003; European Commission, 2003). Anyway, it is possible to say that REACH will contribute to reduced pollution of air, water and soil, as well as to reduced pressure on biodiversity, even if on a long-term period, since chemicals get into the environment in a number of different ways. Improved control of persistent bio-accumulative and toxic substances is needed to ensure that these substances are prevented from polluting the environment, as once there they are very difficult to remove.

The proposal includes also a series of measures and requirements that will influence the direction of industrial innovation, both at product and process development level, as well as at organisational level. The modifications made to the chemical Research and Development
(R&D) regime by the introduction of the REACH system give more flexibility and they will facilitate R&D, especially in the longer-term, while in the short-term much will depend on the resource available. The new system may also lead to induced innovation in the chemicals industry towards the development of new and safer products and processes, and it may also create an incentive for the use and development of such new and safer substances, therefore encouraging innovation.

In summary, the data available indicate that there are significant health and environmental impacts associated with certain chemicals. A better knowledge of the properties of chemicals, acquired through REACH, can be expected to result in better safety and control measures, reducing exposure and hence, the impacts on human health and the environment, besides helping a better implementation of existing legislation.

### 2.5.2. Costs

With regards to costs, and considering a combination of the estimates of direct and indirect costs, the Extended Impact Assessment of the Commission’s proposal of October 2003 (Commission of the European Communities, 2003) has estimated the overall costs for chemical industry and its downstream users to fall in the range of €2.8 - 5.2 billion. These costs will be incurred over a period of 11 (the time to register all substances currently on the EU market) to 15 years (to allow for a longer adjustment period).

Specifically, the expected direct costs to the chemicals sector, which include testing and registration costs, together with Agency fees (paid by industry as contribution to the running of it) as well, were estimated to be approximately €2.3 billion over the 11 years soon after the entry into force of the Regulation. The estimate assumed a high level of sharing information and cooperative actions between stakeholders, and it was based on the assumed availability of validated computer-based methods (Quantitative Structure-Activity Relationships or QSARs), that should allow for a significant reduction in testing costs to the chemicals sector. These methods permit the prediction of physiochemical, environmental, or health effects based upon the molecular structure of a chemical, without the need for costly animal testing. However, the reduction in testing costs arising from the availability and use of QSARs is conditional on the validation, acceptance and use of such techniques on a large
scale and their practical application for regulatory purposes (Commission of the European Communities, 2003; European Commission, 2003). Since such estimations have been made, administrative costs for companies have been further decreased, due to changes such as reduced testing and reporting requirements and simplified registration procedures for low volume chemicals, exclusion of polymers from registration, and a major reduction in downstream user requirements. However, the general estimates remains the same because, at the same time, the costs of the Agency increased significantly, due to the fact that the Council and Parliament have added substantial new responsibilities, in particular for ensuring a harmonized approach to the evaluation of registration dossiers (European Commission, 2006).

Considerable attention has been focused also on the implications of the REACH system for downstream users of chemicals and the costs estimated for them. Predictions of costs were made on the basis of normal business behavior in response to changes in the market and the expert knowledge of the competitive situation of all the sectors and subsectors involved (Commission of the European Communities, 2003). The impact on downstream users results mainly from higher prices of chemicals, due to testing and registration costs, and from the costs related to the need to find substitutes for withdrawn chemical substances and preparations (European Commission, 2003).

As reported in the Extended Impact Assessment of 2003 (Commission of the European Communities, 2003), two scenarios for the costs of REACH to downstream users have been investigated: a "normal expectations" scenario, which is what under normal circumstances the Commission expects should happen and a "higher substitution costs" scenario, which cannot be excluded. Both scenarios are based upon estimated testing and registration costs of € 2.3 billion, i.e. including both testing and registration costs and Agency fees. In each case, a lower and upper estimate of the costs is derived for two time periods: 11 and 15 years.

The "normal expectation" case examines the impact of the introduction of REACH, where the implications for downstream users come only from the pass-through of testing and registration costs and the effects of the withdrawal of chemical substances on individual downstream users.

A "higher substitution costs" scenario illustrates the effects where the withdrawal of substances further increases the costs of substitution, through the cumulative effects of the
withdrawal of substances in terms of adaptation to the whole of the chemicals supply chain. In this case, it has been assumed that the efficiency of the chemicals industry is reduced marginally in proportion with the withdrawal of chemical substances. It also results in some increase in the market power of the suppliers of substitution substances. In this case, higher downstream user costs would be expected.

In the "normal expectation" case, the costs to downstream users of the introduction of REACH is assessed to be in the range € 2.8 - 3.6 billion. These costs will occur in the form of higher chemical prices resulting from the passing through of testing and registration costs and as a result of the additional substitution costs for downstream users of chemicals in finding potentially higher cost or less-effective replacements for those substances removed from the market. In the "higher substitution cost" scenario, the costs to downstream users of the introduction of REACH is assessed to be in the range € 4.0 - 5.2 billion.

In summary, costs to downstream users were estimated in the Commission’s Impact Assessment of 2003 at € 0.5 to 1.3 billion, under the assumption that 1 to 2 % of the substances would be withdrawn because continued production would no longer be profitable. Costs could rise to € 1.7 - 2.9 billion when industry would face higher substitution costs in the downstream supply chains (European Commission, 2006).

Table 2.2 and 2.3 from the Extended Impact Assessment (Commission of the European Communities, 2003) show testing and registration costs of REACH and estimated cost to industry and downstream users, respectively.
Table 2.2. *Testing and registration costs of REACH.*

<table>
<thead>
<tr>
<th>Net present value costs of REACH (€ millions, 3% discount factor)</th>
<th>&gt;1t/y</th>
<th>&gt;10t/y</th>
<th>&gt;100t/y</th>
<th>&gt;1000t/y</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registration costs</td>
<td>€ 100 mn</td>
<td>€ 100 mn</td>
<td>€ 100 mn</td>
<td>€ 200 mn</td>
<td>€ 500 million</td>
</tr>
<tr>
<td>Testing costs</td>
<td>€ 150 mn</td>
<td>€ 300 mn</td>
<td>€ 350 mn</td>
<td>€ 450 mn</td>
<td>€ 1250 million*</td>
</tr>
<tr>
<td>Safety data sheet costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>€ 250 million</td>
</tr>
<tr>
<td>Authorisation procedures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>€ 100 million</td>
</tr>
<tr>
<td>Reduced costs for new substances below 1t etc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(benefit of € 100 million)</td>
</tr>
<tr>
<td>Total testing and registration costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>€ 2,000 million</td>
</tr>
<tr>
<td>Agency fees (paid by chemicals sector)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>€ 300 million</td>
</tr>
<tr>
<td>Total costs (including Agency fees)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>€ 2,300 million</td>
</tr>
</tbody>
</table>

* assuming validation and acceptance of (Q)SARs can be applied within the timeframe envisaged leading to a cost reduction of € 0.949 million.

Table 2.3. *Summary of estimated costs to chemical industry and downstream users (these estimates include costs passed on from the chemicals sector to downstream users).*

<table>
<thead>
<tr>
<th>Net present value costs (€ billions at 3% discount factor)</th>
<th>Lower estimate</th>
<th>Upper estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal expectation</td>
<td>€ 2.8 billion</td>
<td>€ 3.6 billion</td>
</tr>
<tr>
<td>Higher substitution cost</td>
<td>€ 4.0 billion</td>
<td>€ 5.2 billion</td>
</tr>
</tbody>
</table>
2.6. REACH and animal testing.

Protecting the health and the environment from adverse effects due to dangerous chemical substances clearly includes also animals. The REACH Regulation indeed, as already said, has been designed also to reduce animal testing to the absolute minimum and, in line with the 3Rs principle (Russel and Burch, 1953), it requires to cause the minimum of distress and suffering to the animals when experiments are essential. It is clear how acquiring the necessary knowledge on the properties of substances necessarily requires some animal testing (European Commission, 2006). However, in order to minimise the number of animal tests, and therefore lessen its impact on animal use, the REACH Regulation provides a number of possibilities to adapt the testing requirements, using existing data and alternative assessment approaches (ECHA, 2009).

Unnecessary tests can be avoided mainly due to the obligation, for companies that are testing the same chemicals, to share their data, in order to ensure no duplication of animal testing. Available studies (such as studies from other countries, previous animal testing, available in vitro data, epidemiological studies, etc.) have to be shared and safety testing have to be carried out only when no data are available. An important point, in this regard, is that companies will face penalties if they don’t comply (Commission of the European Communities, 2003).

Furthermore, in order to decrease the duplication of animal testing, before performing new test on animals for testing high volume chemicals, the testing proposals must be approved by the Agency, and it would happen only if existing information and validated alternative methods are not sufficient. To further improved the situation, a public consultation period of 45 days before carrying out certain tests was also introduced, with the aim to verify whether the data is already available and consequently if the tests are necessary or not. On the contrary, for low volume chemicals, as far as possible, no animal testing will be required at all. In particular, REACH promotes the use of QSARs as a cheap alternative that does not involve animal testing. All these factors aim to ensure that the endpoints studied are relevant, that the scientific validity of the research is sufficiently high, and that the testing programme does not duplicate other studies (Commission of the European Communities, 2003; European Commission, 2006).
However, new tests must be optimized, standardized and validated before they can be accepted by regulatory authorities for replacing conventional animal studies, and these procedures normally take several years (Hartung, et al., 2003, 2004; Zuang and Hartung, 2005). Although several promising innovative techniques are available today, they have not yet been sufficiently validated with regard to their use in safety assessment, and a premature replacement of animal studies would impair the benefits of REACH, and would be counterproductive for the consumers’ health and block further progress in development of toxicological testing strategies (Hengstler, et al., 2006). Therefore, it is clear how the REACH concept will cause a transient increase in animal experiments in the years following its introduction, since many substances will have to be tested in conventional animal experiments.

When REACH was negotiated, between 2001 and 2005, several attempts were made to estimate the costs of the regulation, both financially and in terms of the number of animals used (Hartung and Rovida, 2009). Various estimates of the number of tests and consequently the number of laboratory animals required by REACH regulation were made to better address the problem. In general, it has been estimated an increase of 3% of animal testing (especially on rats and mice) in the EU for the first eleven years after the adoption of REACH. The Commission initially estimated the number of laboratory animals required to 2.6 million (range of estimate 2.1-3.9 million) (Van der Jagt et al., 2004), but this number was actually wrong, not taking into consideration the number of offspring produced during reproductive toxicity testing. Subsequent estimates, that include the offspring produced during testing, were obviously higher, reaching approximately 9 million laboratory animals (with a financial cost of approximately € 1.6 billion) (Höfer et al., 2004).

The scenario was obviously not bright. Moreover, two important publications of 2009 (Hartung and Rovida, 2009; Rovida and Hartung, 2009) questioned these estimates, suggesting even much higher numbers, mainly due to changes to the final legislation, such as the inclusion of reaction intermediates, and changes to the guidance for industry on how to test, together with an increased number of EU members. The study published by Rovida and Hartung (2009) suggested, indeed, that the testing required would involve 54 million vertebrate animals and that the costs would amount to € 9.5 billion. However, this study has been reviewed by the ECHA, who luckily concluded that the real figures are more likely to be the ones as originally provided in the Commission impact assessments (Commission of the
European Communities, 2003), and that the study overestimated the impact of the legislation by six times, because overestimating the subsequent factors:

1. The probable number of substances that will be registered under REACH and require a full data set, which has been overestimated in the paper by almost two fold, mainly because the assumptions are not justified and seem to be incorrect.

2. The probable number of tests and laboratory animals required, which has been overestimated approximately by six times, mainly because the availability of existing information and the possibilities for adapting the information requirements are not taken into account, and also because the rules for requesting testing in a second species are not interpreted correctly;

3. The probable costs for conducting the tests, which have been overestimated by approximately the same factor (ECHA, 2009).

For all these reasons, the REACH concept has been discussed controversially between different animal welfare groups, that press so that animal experiments could be replaced by \textit{in vitro} techniques on a much larger scale. For instance Greenpeace principally appreciates the REACH concept, but demands an acceleration of the introduction of non-animal alternatives (http://www.greenpeace.org/eu-unit/en/). It is clear how, after the initial period when an increased number of animals foreseen, such number would strongly decrease. It is supposed, indeed, that the lack of knowledge about a wide range of substances in use today should be adequately addressed and therefore, only a few new substances per year will have to be tested, leading to a decrease of the animals required (Hengstler \textit{et al.}, 2006). The primary importance of the REACH Regulation in the field of animal testing is therefore evident, even if clearly the benefits will occur over a longer time frame.
Chapter 3
Evaluation of the use of human Mesenchymal Stem Cells for acute toxicity tests

3.1. Introduction

Poisoning is generally recognized as a severe health problem. The Institute of Medicine (IOM) estimates that more than four million poisoning episodes occur annually in the United States (ICCVAM, 2006a; IOM, 2004). For this reason, *in vivo* toxicological testing designed to assess the acute oral toxicity of chemicals is required by regulatory authorities all around the world to provide classification and labelling warning of the possible consequences of oral exposure to a chemical. However, the utility of the data obtained in these *in vivo* tests for the purpose of predicting the effects occurring in humans has been often questioned (Balls, 1991; Basketter *et al.*, 1997; Langley *et al.*, 2005; York *et al.*, 1996; York and Steiling, 1998). Thus, the replacement of *in vivo* tests with *in vitro* alternatives has reached a high priority level. Recently, many *in vitro* methods for toxicity testing have been developed as alternatives to whole animal tests according to the 3Rs approach – Reduction, Replacement and Refinement – started in 1959 by Russell and Burch (1959). Clearly, the replacement and/or reduction of unnecessary *in vivo* tests would have significant animal welfare benefits and would also result in lower testing costs and time (Ukelis *et al.*, 2008). To this aim, several cell lines have been used for *in vitro* toxicity tests (Ekwall, 1983; Evans *et al.*, 2001; Gennari *et al.*, 2004), that are typically carried out with immortalized cell lines or primary cells, which are directly isolated from animal tissues. Immortalized cells are readily available and easily maintained, although they usually show anomalous behavior and phenotypes, which do not reflect the mechanism observed in their normal homologous cells. In particular, primary cells are considered a better option as model systems for predicting toxicological behavior, although they are limited in quantity and suffer from batch to batch variation. Two different cell lines, mouse fibroblast (BALB/c) 3T3 cells and Normal Human Keratinocytes
(NHK), have already been validated by ICCVAM for predicting cytotoxicity in vitro (ICCVAM, 2006a, 2006b). Both these validated methods not only take advantage of the Neutral Red Uptake (NRU) assay for determining the in vitro cytotoxicity of test substances, but also exploit the in vitro generated data for determining starting doses for in vivo acute oral systemic toxicity tests using the linear relationship between IC₅₀ values (50% Inhibitory Concentration) obtained from NRU assay and rodent oral LD₅₀ values (50% Lethal Dose) established by the Registry of Cytotoxicity (RC). The original linear regression formula, i.e. the so called RC millimole regression, was proposed by ZEBET, the German National Centre for the Documentation and Evaluation of Alternative Methods to Animal Experiments, as a method to reduce animal use by identifying the most appropriate starting doses for acute oral toxicity tests (Halle, 1998, 2003; Spielmann et al., 1999). Subsequently, in the framework of the ICCVAM Validation Study, two new regressions were developed in order to improve not only the RC millimole regression, but also its ability to accurately predict LD₅₀ values from IC₅₀ ones, and finally to make this approach relevant to test mixtures and substances whose molecular weight is unknown.

The 3T3 and NHK NRU test methods are today the in vitro basal cytotoxicity assays recommended for determining starting doses for acute oral toxicity tests using rats, although both these methods are not sufficiently accurate as stand-alone ones to correctly predict acute oral toxicity for the regulatory hazard classification purpose (ICCVAM, 2006b). The ICCVAM Test Methods Evaluation Report recommends that, for a new candidate in vitro basal cytotoxicity test, the correlation between in vitro and in vivo test methods must be quantitatively established by using at least twelve substances that cover all six hazard classification categories. In addition, the performance of the test method should be comparable to or better than the accuracy and reliability of the 3T3 and NHK NRU methods.

In this study we use, for the first time to our best knowledge, human Mesenchymal Stem Cells (hMSCs) as cell line in order to develop a novel in vitro model system for acute toxicity test of chemicals. The unique properties of hMSCs, such as unlimited proliferation ability, plasticity to generate other cell types and more readily available sources of human cells, clearly identify their potential benefits in toxicology (Davila et al., 2004). Specifically, the aim of this study was to evaluate the applicability of hMSCs as cell line for in vitro cytotoxicity tests to correctly predict LD₅₀ and the hazard category according to the Globally
Harmonized System of Classification – GHS (ONU, 2009), and finally to compare their behavior with the validated 3T3 and NHK NRU test methods.

3.2 Material and methods

3.2.1. Isolation of Mesenchymal Stem Cells

After obtaining informed consent, fifteen milliliters of bone marrow aspirate, taken from the iliac crest of two human donors, were diluted with equal volume of phosphate-buffered saline (PBS, Sigma, St. Louis, MO, USA). Mononuclear cells were isolated by a Ficoll-Paque density gradient centrifugation (Hystopaque 1077, Sigma, 30 min, 800 x g), and subsequently washed in PBS (10 min, 400 x g). Possible residual red blood cells were lysed using a commercial lysing solution (BD Pharm lyse, Milano, Italy). After centrifugation (5 min, 300 x g), the pellet was resuspended in complete culture medium constituted by αMEM (Sigma), 20% fetal bovine serum (FBS, Sigma), 1% penicillin/streptomycin solution (Sigma), and 2 mM L-glutamine (Sigma). Nucleate cells were counted by haemocytometer and plated at a density of $0.6-1 \times 10^6$ cells/cm$^2$ in uncoated Petri dishes. After overnight incubation at 37 °C under a humidified atmosphere containing 5% CO$_2$, non-adherent cells were removed and fresh medium was added. By eliminating the remaining non-adherent cells through complete exchange of the culture medium every 3 days, cultures were maintained for 2 weeks. Then, cells were harvested by the use of 0.1% trypsin and 0.04% EDTA for 6 min at 37 °C and replated at a density of $1.5 \times 10^4$ cells/cm$^2$. Medium was changed every 2 days, and cells at passage 6 were used for the study.

3.2.2. Phenotypic characterization of stem cell cultures

Cells harvested by treatment with 0.1% trypsin and 0.04% EDTA were fixed in 4% paraformaldehyde for 20 min, washed with PBS, then resuspended in PBS containing 0.1% bovine serum albumin. Typically, $10^5$ cells were incubated in 100 µl solution for 20 min on ice with monoclonal antibody followed by two washes in PBS containing 0.1% bovine serum albumin and, when necessary, a further incubation with conjugated secondary antibody for 20
min at room temperature was performed. Fluorescein isothiocyanate-conjugated monoclonal antibodies against hCD34, hCD44 (Immunotech, Marseille, France), unconjugated monoclonal antibodies against hCD73, hCD90 (BD Pharmigen), and hCD105 (DakoCytomation, Glostrup, Denmark), and secondary fluorescein isothiocyanate-conjugated sheep F(ab′)2 fragment antimouse immunoglobulin G antibody (Chemicon International, Milano, Italy) were used (Mancuso et al., 2009; Pittenger et al., 1999). Flow cytometry was performed on a fluorescence-activated cell sorter (BD FACSCalibur, BD Biosciences, Franklin Lakes, NJ, USA). CellQuest Pro was the software used for flow cytometry data analysis (BD Bioscience, CA, USA).

3.2.3. Determination of doubling time

The ICCVAM Test Methods Evaluation Report (ICCVAM, 2006b) recommends to use a mammalian cell line (or primary cells) that divides rapidly with doubling times of less than 30 h under standard culture conditions. Human MSCs were plated in 20 Petri dishes (8 cm², Corning, Corning, NY, USA) at a density of 1 x 10⁴ cells/cm² and incubated at 37 °C in a humidified atmosphere containing 5% CO₂ in a complete culture medium supplemented with 20% fetal bovine serum. Cells from three plates were harvested after 5, 24, 48 and 72 h respectively, with the use of 0.1% trypsin and 0.04% EDTA for 7 min at 37 °C. Action of trypsin was stopped using the complete medium, and Petri dishes were washed with PBS. Cells were counted electronically using a Coulter Counter (Beckman Dickinson, Fullertan, CA, USA). Under these conditions the approximate doubling time for hMSCs is 24 h.

3.2.4. Chemicals and preparation

Test materials used for hMSCs NRU assay were purchased from Sigma–Aldrich (St. Louis, MO, USA). As suggested by ICCVAM, twelve of the thirty reference substances were chosen, two for each of the five GHS hazard categories and two unclassified ones (ICCVAM, 2006b). As positive control, sodium dodecyl sulfate was used. Chosen chemicals are reported in Table 3.1.
Table 3.1. The twelve of the thirty ICCVAM reference substances chosen for the study, divided for GHS hazard categories, along with the corresponding molecular weight and CAS-number.

<table>
<thead>
<tr>
<th>Chemicals</th>
<th>Molecular weight</th>
<th>CAS-number</th>
</tr>
</thead>
<tbody>
<tr>
<td>positive control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sodium dodecyl sulfate</td>
<td>288.38</td>
<td>151-21-3</td>
</tr>
<tr>
<td>LD\textsubscript{50} \leq 5 mg/kg (class 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mercury II chloride</td>
<td>271.5</td>
<td>7487-94-7</td>
</tr>
<tr>
<td>Cycloheximide</td>
<td>281.39</td>
<td>66-81-9</td>
</tr>
<tr>
<td>5&lt;LD\textsubscript{50} \leq 50 mg/kg (class 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sodium arsenite</td>
<td>129.91</td>
<td>7784-46-5</td>
</tr>
<tr>
<td>Sodium dichromate dihydrate</td>
<td>298</td>
<td>7789-12-0</td>
</tr>
<tr>
<td>50&lt;LD\textsubscript{50} \leq 300 mg/kg (class 3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cadmium II chloride</td>
<td>183.32</td>
<td>10108-64-2</td>
</tr>
<tr>
<td>Sodium fluoride</td>
<td>41.99</td>
<td>7681-49-4</td>
</tr>
<tr>
<td>300&lt;LD\textsubscript{50} \leq 2000 mg/kg (class 4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Propranolol HCl</td>
<td>295.8</td>
<td>3506-09-0</td>
</tr>
<tr>
<td>Atropine sulfate monohydrate</td>
<td>694.83</td>
<td>5908-99-6</td>
</tr>
<tr>
<td>2000&lt;LD\textsubscript{50} \leq 5000 mg/kg (class 5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Potassium chloride</td>
<td>74.56</td>
<td>7447-40-7</td>
</tr>
<tr>
<td>Trichloroacetic acid</td>
<td>163.39</td>
<td>76-03-9</td>
</tr>
<tr>
<td>LD\textsubscript{50} &gt;5000 mg/kg – unclassified (class 6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sodium hypochlorite</td>
<td>74.443</td>
<td>7681-52-9</td>
</tr>
<tr>
<td>Glycerol</td>
<td>92.09</td>
<td>56-81-5</td>
</tr>
</tbody>
</table>

3.2.5. Mesenchymal Stem Cells NRU assay

The NRU assay was performed according to the standard protocol of Borenfreund and Puerner (1985) modified by ICCVAM (2006a, 2006b). Human MSCs at 6\textsuperscript{th} passage were seeded in 96-well microtiter plates at the concentration of 3.5 \times 10^3 cells/100\mu l/well in αMEM culture medium (supplemented with 20% FBS, 1% penicillin/streptomycin and 2 mM...
L-glutamine, Sigma), and then were incubated (37 °C/5% CO₂) for 24 h, in order to assure adequate cell recovery and adherence. This optimal seeding density was determined by performing the NRU assay with different cell densities (1.5, 2, 2.5, 3, 3.5, 4, and 4.5 x 10³ cells) in order to identify the number of cells that would allow an exponential growth during the test.

After the 24 h-incubation, cells were treated with eight concentration levels of each test chemical in 100 µl treatment medium (αMEM supplemented with 5% FBS, 1% penicillin/streptomycin and 2 mM L-glutamine, Sigma). All chemicals were dissolved in αMEM culture medium without any solvent, and consequently the untreated vehicle control was incubated with only treatment medium. Cells were then incubated (37 °C/5% CO₂) for 48 more hours. After the incubation period, the chemical solutions were removed from all plates and the cells were washed with 250 µl/well of pre-warmed PBS. After that, 250 µl of NR medium (1 ml NR Stock Solution, 79 ml αMEM; Stock Solution: 0.4 g NR Dye, 100 ml milliQ H₂O) were added to all wells, including the blanks, and incubated for 3 more hours. The cells were briefly observed between 2 and 3 h after incubation in order to check the NR crystal formation. After 3 h, the NR medium was removed and the cells were carefully rinsed with 250 µl/well of pre-warmed PBS. The PBS was discarded and 100 µl/well of NR desorbing fixative (1% glacial acetic acid solution, 50% ethanol, 49% H₂O) was added, in order to elute the dye. Plates were shaken for 20 min to form a homogenous solution and NR absorption was detected at 540 nm in a microtiter plate reader/spectrophotometer (BioTek EL800 plate reader, BioTek Instrument Inc, Winooski, VT, USA).

### 3.2.6. Statistical analysis

Data from the microtiter plate reader were transferred in a spreadsheet template Microsoft Office Excell 2007® to determine cell viability (expressed as percentage of untreated controls) and to verify the test acceptance criteria established by ICCVAM for the assay (ICCVAM, 2006b). Specifically, IC₅₀ values for each test substance were calculated according to ICCVAM using the following rearranged Hill function, i.e. a four-parameter logistic mathematical model, by means of GraphPad Prism® 4.0 statistical software:
where $IC_{50}$ is the concentration producing 50% toxicity, $EC_{50}$ is the concentration producing a response midway between the Top and Bottom responses, Top is the maximum response (maximum survival), which has been fixed equal to 100, Bottom is the minimum response (maximum toxicity), which has been fixed equal to zero when 0% viability is reached during the experiments or it is left unconstrained when dose-responses do not achieve 100% cytotoxicity, $Y = 50$ (i.e. 50% response), and HillSlope, which has no units, describes the steepness of the curve. To evaluate the capability of the rearranged Hill function to quantitatively interpret the experimental data, the determination coefficient $R^2$ was used.

It should be noted that Eq. (1) is adopted according to ICCVAM (2006a, 2006b), where the rearranged Hill function is recommended as the most common equation for *in vitro* dose-response data, since all the dose-response information, rather than few points around the $IC_{50}$, is used. Specifically, using the rearranged Hill function it is possible to evaluate the slope of the dose-response curve, which is extremely important in predicting the toxicity of a substance at specific dose levels, and it shows how fast the response increases as the concentration is augmented.

The obtained $IC_{50}$ data were reported in terms of their average value ± SD of at least two independent experiments, which were carried out in six replicates (i.e. twelve values for each substance). The same data were used to perform a linear regression analysis using the corresponding $LD_{50}$ values provided in the ICCVAM Test Methods Evaluation Report (ICCVAM, 2006b). To quantitatively evaluate the performed regression analysis, the linear determination coefficient $r^2$ was used. The resulting regression was then compared to the 3T3 and NHK NRU regression ones, through a $F$-test performed using GraphPad Prism® 4.0 statistical software. The obtained $IC_{50}$ data were also used to predict $LD_{50}$ values and GHS hazard categories using the RC rat-only millimole regression and the RC rat-only weight regression, as recommended by ICCVAM (2006b).
3.3. Results

Phenotypic characterization of human bone marrow stem cells is considered first. Cells at passage six resulted to be negative for CD34 and CD45, and positive for CD44, CD105, CD90 and CD73 (data not shown). On the basis of these results, the adopted human bone marrow stem cells displayed a mesenchymal phenotype.

We have been then following the procedure for evaluating the potential use of hMSCs to estimate LD\textsubscript{50} values and to predict the GHS hazard category for each tested chemical.

The hMSCs test evaluation was performed using the twelve substances chosen as suggested by ICCVAM (2006b) and the cell viability related to each chemical was evaluated following the NRU assay. Recommendations from ICCVAM were accurately followed and the suggested acceptance criteria were completely fulfilled, otherwise the results were not taken into account and the assay was discarded. Such criteria are reported as follows:

- The average of the left and the average of the right columns of vehicle controls (VCs) do not differ by more than 15\% from the average of all VCs.

- At least one calculated cytotoxicity value $>0\%$ and $\leq 50\%$ viability and at least one calculated cytotoxicity value $>50\%$ and $<100\%$ viability must be present.

- The positive control (SDS) IC\textsubscript{50} must be within $\pm 2.5$ standard deviations (SD) of the average values established by the laboratory.

- The positive control (SDS) dose response must have an $R^2 \geq 0.85$ for the rearranged Hill model fit.

The concentration of each test chemical reflecting the 50\% inhibition of cell viability (IC\textsubscript{50}) was determined from the concentration response data through Eq. (1), as previously described in the Section 3.2.6. An example of the fit of the concentration-response data using Eq. (1) is shown in Figure 3.1.
Figure 3.1. Cell viability data. The cadmium II chloride dose-response curve for hMSCs after 48 h exposure. Points and error bars show the means ± SD, for the percent cell viability response of the 12 replicate wells at each of the eight concentrations. The line shows the fit of the concentration-response to the rearranged Hill function, reported in Eq. (1). The IC₅₀ found was 386.55 ng/ml.

The obtained IC₅₀ values (expressed as geometric mean) are reported in Table 3.2, along with the corresponding values of the rearranged Hill function determination coefficient (R²). Each IC₅₀ value was then transformed to mM units, and consequently linearly correlated with the corresponding LD₅₀ one from the RC (also reported in Table 3.2). The resulting regression was compared with the ones obtained during the 3T3 and NHK ICCVAM Validation Study for the same substances, as suggested by ICCVAM (2006b).
Table 3.2. Selected chemicals for candidate cytotoxicity test, rodent oral LD$_{50}$ values from the RC (Registry of Cytotoxicity) and IC$_{50}$ values obtained from our study with hMSCs expressed both in µg/ml and mmol/l, with the corresponding values of the rearranged Hill function determination coefficient ($R^2$) obtained when fitting viability data for each chemical as a function of its concentration levels through Eq. (1). (The IC$_{50}$ values for SDS represent the average of 20 different experiments previously determined in our laboratory).

<table>
<thead>
<tr>
<th>Chemicals</th>
<th>RC rodent oral LD$_{50}$ (mmole/kg)</th>
<th>IC$_{50}$ (µg/ml)</th>
<th>IC$_{50}$ (mM)</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sodium dodecyl sulfate</td>
<td>4.4663</td>
<td>53.85</td>
<td>0.19</td>
<td>0.96</td>
</tr>
<tr>
<td>Mercury II chloride</td>
<td>0.0037</td>
<td>43.10</td>
<td>0.16</td>
<td>0.97</td>
</tr>
<tr>
<td>Cycloheximide</td>
<td>0.0071</td>
<td>1.72</td>
<td>0.006</td>
<td>0.96</td>
</tr>
<tr>
<td>Sodium arsenite</td>
<td>0.3156</td>
<td>1.52</td>
<td>0.01</td>
<td>0.89</td>
</tr>
<tr>
<td>Sodium dichromate dihydrate</td>
<td>0.1908</td>
<td>1.15</td>
<td>0.004</td>
<td>0.95</td>
</tr>
<tr>
<td>Cadmium II chloride</td>
<td>0.4801</td>
<td>0.39</td>
<td>0.002</td>
<td>0.98</td>
</tr>
<tr>
<td>Sodium fluoride</td>
<td>4.29</td>
<td>65.51</td>
<td>1.56</td>
<td>0.98</td>
</tr>
<tr>
<td>Propranolol HCl</td>
<td>1.589</td>
<td>22.99</td>
<td>0.08</td>
<td>0.98</td>
</tr>
<tr>
<td>Atropine sulfate monohydrate</td>
<td>0.9204</td>
<td>445.14</td>
<td>0.64</td>
<td>0.98</td>
</tr>
<tr>
<td>Potassium chloride</td>
<td>34.9</td>
<td>5592.89</td>
<td>75.01</td>
<td>0.92</td>
</tr>
<tr>
<td>Trichloroacetic acid</td>
<td>30.59</td>
<td>1408.43</td>
<td>8.62</td>
<td>0.89</td>
</tr>
<tr>
<td>Sodium hypochlorite</td>
<td>138.7</td>
<td>1601.40</td>
<td>21.51</td>
<td>0.91</td>
</tr>
<tr>
<td>Glycerol</td>
<td>137.8</td>
<td>68391.33</td>
<td>742.66</td>
<td>0.98</td>
</tr>
</tbody>
</table>

The obtained regression line and its comparison with both the validated regressions are shown in Figure 3.2. F-test was then performed by comparing slopes and intercepts of the three regression lines. A significant level of p < 0.05 was used to test whether regressions were significantly different from one another, and the results show that hMSCs regression line is not statistically different from both 3T3 (slope p value = 0.8231, intercept p value = 0.4834) and NHK (slope p value = 0.7445, intercept p value = 0.5095).
Figure 3.2. Comparison of hMSCs, 3T3 and NHK regressions. hMSCs regression for the twelve chosen substances compared to both 3T3 and NHK regressions from the ICCVAM Validation Study. The hMSCs regression (black dashed line vs open squares) for twelve substances is \( \log LD_{50} = 0.66 \log IC_{50} + 0.3929 \) \((r^2 = 0.6303)\). The 3T3 regression (black dotted line vs open triangles) is \( \log LD_{50} = 0.7052 \log IC_{50} + 0.6644 \) for the same twelve substances \((r^2 = 0.7744)\). The NHK regression (black line vs open circles) is \( \log LD_{50} = 0.7246 \log IC_{50} + 0.6534 \) for the same twelve substances \((r^2 = 0.8024)\). The symbol \( r^2 \) refers to the linear determination coefficient. Data for 3T3 and NHK regressions came from the ICCVAM Validation Study.

According to the ICCVAM study (2006b), the accuracy of the NRU method and the associated IC\(_{50}\) - LD\(_{50}\) regressions can be assessed by calculating the predicted LD\(_{50}\) for each tested substance using its geometric mean IC\(_{50}\) in the following regressions, i.e. the RC rat-only millimole and the RC-rat only weight ones, respectively:

- \( \log LD_{50} \text{ (mmol/kg)} = 0.439 \log IC_{50} \text{ (mM)} + 0.621 \).
- \( \log LD_{50} \text{ (mg/kg)} = 0.372 \log IC_{50} \text{ (µg/ml)} + 2.024 \).
The LD₅₀ values obtained from the regression analyses are reported in Table 3.3 and are then used to predict GHS hazard categories, which are shown in Table 3.1 and also included in Table 3.3.

Table 3.3. Prediction of GHS Acute Oral Toxicity Category by the hMSCs, 3T3 and NHK NRU Test Methods with the RC Rat-Only Millimole and the RC rat-only weight regressions. Predictions for hMSCs were calculated in the present study, while 3T3 and NHK values come from the ICCVAM Validation Study (a – underprediction, b – overprediction).

<table>
<thead>
<tr>
<th>Chemicals</th>
<th>GHS class</th>
<th>RC rodent oral LD₅₀ (mg/kg)</th>
<th>Predicted LD₅₀ (mg/kg) with RC Rat-Only Millimole Regression</th>
<th>Predicted LD₅₀ (mg/kg) with RC Rat-Only Weight Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>hMSCs</td>
<td>3T3</td>
</tr>
<tr>
<td>Mercury II chloride</td>
<td>LD₅₀ ≤ 5 mg/kg (class 1)</td>
<td>1</td>
<td>506²</td>
<td>181a</td>
</tr>
<tr>
<td>Cycloheximide</td>
<td>LD₅₀ ≤ 5 mg/kg (class 1)</td>
<td>2</td>
<td>125²</td>
<td>48a</td>
</tr>
<tr>
<td>Sodium arsenite</td>
<td>5 &lt; LD₅₀ ≤ 50 mg/kg (class 2)</td>
<td>41</td>
<td>77²</td>
<td>57a</td>
</tr>
<tr>
<td>Sodium dichromate dihydrate</td>
<td>5 &lt; LD₅₀ ≤ 50 mg/kg (class 2)</td>
<td>50</td>
<td>109²</td>
<td>81a</td>
</tr>
<tr>
<td>Cadmium II chloride</td>
<td>50 &lt; LD₅₀ ≤ 500 mg/kg (class 3)</td>
<td>88</td>
<td>51</td>
<td>58</td>
</tr>
<tr>
<td>Sodium fluoride</td>
<td>50 &lt; LD₅₀ ≤ 500 mg/kg (class 3)</td>
<td>180</td>
<td>213</td>
<td>230</td>
</tr>
<tr>
<td>Propranolol HCl</td>
<td>300 &lt; LD₅₀ ≤ 2000 mg/kg (class 4)</td>
<td>470</td>
<td>403</td>
<td>325</td>
</tr>
<tr>
<td>Atropine sulfate monohydrate</td>
<td>300 &lt; LD₅₀ ≤ 2000 mg/kg (class 4)</td>
<td>639</td>
<td>2388²</td>
<td>1099</td>
</tr>
<tr>
<td>Potassium chloride</td>
<td>2000 &lt; LD₅₀ ≤ 5000 mg/kg (class 5)</td>
<td>2602</td>
<td>2073</td>
<td>1699²</td>
</tr>
<tr>
<td>Trichloroacetic acid</td>
<td>2000 &lt; LD₅₀ ≤ 5000 mg/kg (class 5)</td>
<td>4999</td>
<td>1758²</td>
<td>1445²</td>
</tr>
<tr>
<td>Sodium hypochlorite</td>
<td>LD₅₀ &gt; 5000 mg/kg (class 6)</td>
<td>10328</td>
<td>1196²</td>
<td>990²</td>
</tr>
<tr>
<td>Glycerol</td>
<td>LD₅₀ &gt; 5000 mg/kg (class 6)</td>
<td>12691</td>
<td>7006</td>
<td>4452²</td>
</tr>
</tbody>
</table>
Using the RC rat-only millimole regression, it is seen that the hMSCs test method correctly predicts the GHS category by 41.7% (5/12) of the tested chemicals, while the 3T3 NRU one by 33.3% (4/12) and the NHK one by 41.7% (5/12) respectively. More specifically, in vivo toxicity was underpredicted by 41.7%, 33.3% and 25% and overpredicted by 16.7%, 33.3% and 33.3% for the hMSCs, 3T3 and NHK, respectively. On the other hand, using the RC rat-only weight regression, it is seen that the hMSCs test method correctly predicts the GHS category by 41.7% (5/12) of the tested chemicals, while both the 3T3 and NHK NRU ones by 25% (3/12). In vivo toxicity in this case was underpredicted for all the cell lines by 41.7%, and it was overpredicted by 16.7%, 33.3% and 33.3% for the hMSCs, 3T3 and NHK, respectively. The major difficulty for hMSCs in predicting the right category, using both the RC rat-only regressions, was found particularly when considering the most toxic chemicals (GHS 1 and 2 classes), whose in vivo toxicity was invariably underpredicted. The same result holds also true for the first two GHS classes when considering the 3T3 and NHK NRU assays with both the RC rat-only regressions. On the contrary, the use of hMSCs allows one to overpredict the in vivo toxicity of only two of the selected chemicals, i.e. trichloroacetic acid (GHS class 5) and sodium hypochlorite (GHS class 6). Instead, the 3T3 and NHK validated methods resulted in an overprediction for all the slightly toxic chemicals. It should be noted that the 3T3 NRU method showed a good prediction only for the case of potassium chloride (GHS class 5) when using specifically the RC rat-only weight regression.

### 3.4. Discussion

The aim of the present study was to test a new cell line (hMSCs) for basal cytotoxicity assays. The approach was based on the regression model developed by Spielmann (1999), following the rules established by ICCVAM (2006b) after the 3T3 and NHK Validation Study. Recommendations from ICCVAM were accurately followed and the suggested acceptance criteria were completely fulfilled. The ICCVAM recommendations require that any new cell line to be tested should be demonstrated to meet or exceed the accuracy and reliability of the two already validated cell lines 3T3 and NHK. Our results show that the obtained hMSCs regression does not differ statistically from the regressions related to 3T3 and NHK cells validated by ICCVAM, based on the comparison of the slope and the intercept, which are essential when considering the test suitable for basal cytotoxicity assays.
The accuracy of the presented method was then characterized by identifying the chemicals (and their proportion with respect to all tested ones) for which the GHS hazard category was correctly predicted, while comparing the proposed cell line to the two validated ones. The results clearly show that, for the tested substances, hMSCs have the same or better ability than 3T3 and NHK to correctly predict the GHS hazard category.

This finding seems to be very promising, since the ability to test chemicals using a human system would increase the relevance and accuracy of predicting toxicological outcomes. Moreover, and that is why we choose to test stem cells, the ability of stem cells to differentiate into a variety of cell types and develop into organ system could allow them to replace transformed cell lines and primary cells for \textit{in vitro} studies, thus eliminating potential limitations and improving the relevance of predictive assay. Another factor to be considered is that the ability to derive stem cells from individual human subjects would offer unprecedented opportunities to analyze the contribution of genetic background that affect susceptibility to toxicity.

There are also additional aspects why hMSCs could be a better option than the two already validated cell lines. In this regard, it should be pointed out that 3T3 and NHK NRU methods are comparable, while the ICCVAM suggests the use of 3T3 cells to conduct these kinds of tests because of the lower costs compared to the use of NHK cells. On the other hand, 3T3 do not represent a human system that in general should be preferred. Along these lines, hMSCs might represent an ideal alternative to NHK, since the corresponding costs of related experiments are comparable to 3T3.

The proposed method will enable toxicity evaluation of novel chemicals in \textit{in vitro} studies, and could also be used for predicting starting doses for \textit{in vivo} toxicity studies, so that a substantial reduction of tested animals and experiment costs will result as well. Such starting doses can be effectively employed in the framework of several approved procedures for \textit{in vivo} acute oral toxicity testing, such as ATC (Acute Toxic Class), UDP (Up-and-Down Procedure) and FDP (Fixed Dose Procedure) methods, according to OECD (2001a, 2001b, 2008). In addition, computer simulations could be used appropriately to estimate, per substance, the number of animals necessary for the study and also the related survival rate.
It should be noted, however, that the proposed cell line, i.e. hMSCs, as well as 3T3 and NHK, are not able to correctly predict the right hazard category of the most toxic chemicals (GHS 1 and 2 classes).

A number of factors could be potentially responsible for such lack of accuracy. It is apparent in this regard that in vitro cell cultures are not able to mimic the kinetics and dynamics of substances related to an in vivo system. Indeed, when considering in vitro systems, absorption, distribution, metabolism and excretion processes, which control the exposure of the target tissue of the organism in vivo, are absent (Blaauboer et al., 1990; Flint, 1990). As a consequence, the toxicant concentrations to which in vitro and in vivo systems are exposed may not correspond (or correspond only for a limited fraction of time) to each other (Blaauboer, 2002a, 2002b; ICCVAM 2006a, 2006b). Thus, as for the chemicals requiring metabolic activation to display their toxicity, it is quite possible that their behavior is not correctly interpreted by the regression, which in turn is not able to properly predict in vivo toxicity (Clemedson et al., 2002).

Furthermore, it is well known that most basal cytotoxicity assays underestimate toxicity for chemicals known to act on specific receptor or cells (Clemedson, et al., 2002). This could be the case of mercury II chloride (GHS 1 class) and sodium arsenite (GHS 2 class), whose primary target organs in humans are kidney, central, and peripheral nervous systems (Aschner and Aschner, 1990; Cheng et al., 2011; HSDB, 2005; Huang et al., 2010; Lewis, 2000; Magos and Clarkson, 2006) and gastrointestinal tract, heart, brain, and kidneys (Das et al., 2010; Garcia-Chavez et al., 2003; HSDB, 2003; Lin et al., 2007; Manna et al., 2008), respectively, while the cell lines, i.e. hMSCs as well as 3T3 and NHK, used for toxicity assays and compared in this work, are derived from tissues different from the target ones. Thus, the toxicity estimated using those cell lines may be ascribed to different mechanisms.

Another important factor to be considered is the presence of serum proteins, which may or may not be included in cell culture systems. For example, hMSCs and 3T3 culture media contain serum, while this is not the case of NHK one. Indeed, a toxicant may or may not be bound to serum proteins, which could reduce its availability to target sites. Serum proteins create an optimal growth environment for the cells and preserve certain vital components, such as for example membrane and cytoskeletal ones, or key enzymes. Specifically, the underpredicted toxicity of the cycloheximide, a known inhibitor of protein synthesis, could be
explained, as supported by Geier et al. (1992), who demonstrated that, although total protein synthesis remain blocked, serum inhibits cell death induced by such toxicant, probably through stabilization of some crucial cell proteins vital for the cells.

Let us consider the case when the proposed hMSCs give rise to overpredicted toxicity. While certain factors discussed above may have an influence, the observed behavior can be related to the NRU assay itself, since it is based on the ability of viable cells to incorporate and bind the neutral red, which is a supravital dye that penetrate cell membranes of viable cells and accumulates in the lysosomes. Thus, if alterations of the cell surface or the sensitive lysosomal membrane due to the action of the toxicant result in a decreased uptake and binding of neutral red, a false reflection of low cell number and viability may be observed, which in turn gives rise to a toxicity overestimation (Barile, et al., 1994). This explanation does apply to the case of trichloroacetic acid, since a recent study shows that it can cause lysosomal membrane destabilization (Abdel-Hamid, et al., 2011). Sodium hypochlorite behavior as well may be justified, in our opinion, following the same considerations, since it displays mechanism of membrane degradation, mainly due to the chlorine action (Estrela et al., 2002; Simon et al., 2009).

In conclusion, we show for the first time that hMSCs can confidently be used to perform in vitro acute toxicity tests. Indeed, compared to traditional in vitro systems based on transformed or immortalized cell lines, hMSCs provide a more accurate modeling of in vivo conditions. It is expected that the study carried out in this research, even if on a relatively small number of toxicants, will be a step forward to current and future validation efforts to develop alternative methods to replace animal testing.
Chapter 4
Evaluation of ovine blood vessels as a possible alternative method for vascular toxicity: decellularization

4.1. Introduction

Heart and blood vessels are an important circuit inside the body both for the transport of nutrient and oxygen, through blood delivery, to all the cells of the organism, and for the removal of waste products of cellular metabolism from those cells. In serving circulatory functions, however, cells of the heart and vasculature are repeatedly exposed to blood-borne toxicants and their metabolic products as well, which compromise not only cardiovascular function, but interfere with specialized functions as well. What happened is that acute cardiovascular toxicity caused by chemicals and drugs involves cellular death and destruction of extracellular matrix (ECM) components, and it may also cause alterations of arterial pressure, mainly due to changes in contractility and blood flow, finally leading to an end-organ dysfunction (Partridge et al., 2005). That means that, as a consequence, damages to the cells of the heart and vasculature produced by chemicals are reported to be responsible also for other organ alterations, such as kidney (Kahan, 1989; Milner et al., 1991) or lung (Martin and Kachel, 1987; Martin et al., 1985; Rosenow et al., 1968; White et al., 1989) for example, which are highly vascularized (Chappey et al., 1995). Moreover, they can contribute to a variety of diseases that today represent a serious threat to human health, including elevated blood pressure (hypertension), hardening of the arteries (arteriosclerosis), abnormal heartbeat (cardiac arrhythmia), and decreased blood flow to the heart (coronary ischemia) (Laverty et al., 2011). That is why studying cardiovascular toxicity, virtually associated directly or indirectly with all chemicals and drugs that cause deteriorating effects in mammalian system, is so much important, and its accurate understanding and detection is vital both for the safeguard of the health of patients and for the development of new drugs and chemicals,
being therefore a critical issue in clinical practice, environmental health and drug development (Acosta, 2008; Knapton et al, 2010; Laverty et al., 2011).

*In vitro* technologies to study chemically-induced injury to the vascular systems cover a wide range of possibilities, such as perfused organ preparations (Crass et al., 1988; Hale and Poklis, 1986; Hein and Kuo, 1998; Khatter et al., 1989; Pilcher and Langley, 1986; Skrzypiec-Spring et al, 2007; Sutherland and Hearse, 2000), organ culture (Bachlav et al., 1999; De lima et al., 1999; Gotlieb and Boden, 1984; Ingwall et al., 1975; Speralakis and Shigenobu, 1974, Tanaka et al., 1987), tissue slices (Gandolfi et al., 1995; Kretz et al., 1999; Parrish et al., 1995), isolated muscle preparations (Abdel-Haq et al., 2000; Campbell et al., 2000; Conklin and Boor, 1998; Koyama et al., 1997; Togna et al., 1984), or cell suspensions (Andrieu-Abadie et al., 1999; Khalifa et al., 1991; Shyu et al., 2000) and cultures (Aszalos et al., 1984; Concklin et al., 1999; Estevez et al., 2000; He et al., 1999; Hendrickson et al., 1999; Kossenjans et al., 1996; Marra et al., 2000; Muthalif et al., 2001; Parrish and Ramos, 1997; Ramos et al., 1984; Wenzel and Innis, 1983), depending on what is the objective to be tested. Clearly, responses at the cellular level may be not representative of those observed in the intact organism and therefore multiple approaches need to be employed to properly assess the overall toxic response (Partidge et al, 2005). Common *in vitro* models for evaluating cardiovascular toxicity are listed in Table 4.1, as provided by Partidge et al. (2005).

What interests us the most with regard to our studies is specifically "vascular toxicity", therefore the toxicity of the vasculature, and the application of those models considering whole perfused blood vessels. The use of such preparations in toxicological studies is advantageous because the level of structural organization is similar to that encountered *in vivo*, and changes in physiological or pharmacological sensitivity, excitability and contractility can be readily evaluated. On the contrary, their most significant limitations are the small number of replicate preparations that can be processed at any time, and the short time available for isolation and placement of the tissue under physiological conditions (Partidge et al, 2005). We therefore thought about a model able to get through these limitations, while maintaining all the advantages the most as possible. In our opinion, the application of tissue engineering principles, which hold today great promises in several fields, and of the concepts of decellularization/recellularization could be applied to this matter.
Table 4.1. Common in vitro models for evaluating cardiovascular toxicity.

<table>
<thead>
<tr>
<th>Method</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfusion</td>
<td>Maintains structural organization</td>
<td>Small number of replicates</td>
</tr>
<tr>
<td>Langendorff</td>
<td>High correlation with drugs associated with TDP in clinical setting</td>
<td>Time constraints during isolation and preparation</td>
</tr>
<tr>
<td>Vascular rings</td>
<td>PV relationship can be maintained</td>
<td>Drugs may alter coronary vasculature impacting rate, output and contraction</td>
</tr>
<tr>
<td>Organ culture:</td>
<td>Cell-cell-substratum interactions</td>
<td>Small number of replicates</td>
</tr>
<tr>
<td>Isolated vessels</td>
<td>Effects of endothelium can be removed</td>
<td>Long term effects cannot be monitored</td>
</tr>
<tr>
<td>Fetal hearts</td>
<td>Simple, cheap and robust</td>
<td>No metabolism</td>
</tr>
<tr>
<td>Frog or mouse whole hearts</td>
<td>Preserve tissue architecture</td>
<td>Phenotypic changes in VSMCs cultured with FBS</td>
</tr>
<tr>
<td>Isolated muscle:</td>
<td>Examination of tension development</td>
<td>Oxygenation depends on diffusion</td>
</tr>
<tr>
<td>Atrial</td>
<td>Plasma membrane can be removed</td>
<td>Stability of contraction/relaxation reliability short lived</td>
</tr>
<tr>
<td>Ventricular</td>
<td>Isotonic, isometric force development can be examined</td>
<td>Xenobiotic metabolic effects not accounted for</td>
</tr>
<tr>
<td>Papillary</td>
<td>Confining effects of drugs on heart or pulmonary circulation removed</td>
<td>Need to co-culture with liver or intestinal slices to observe biotransformation of drugs</td>
</tr>
<tr>
<td>Vascular beds</td>
<td>Direct interspecies comparison</td>
<td>Variants such as temperature, enzymes or pH/osmolarity can critically affect experiments</td>
</tr>
<tr>
<td>Tissue slices:</td>
<td>Easy preparation and standardization</td>
<td>Reduction of surface receptor binding sites can occur with isolation</td>
</tr>
<tr>
<td>Single cell suspension:</td>
<td>Multiple cell functions studied at same time</td>
<td>Cell architecture and heterogeneity maintained</td>
</tr>
<tr>
<td>Embryonic</td>
<td>Use of voltage clamp possible</td>
<td>Cell architecture and heterogeneity maintained</td>
</tr>
<tr>
<td>Adult</td>
<td>Reliably pre-clinical model of long QT</td>
<td>No activators of ERG available</td>
</tr>
<tr>
<td>HERG channel assay:</td>
<td>Assay for APD prolongation or shortening</td>
<td>Isolation of cells difficult</td>
</tr>
<tr>
<td>Isolated cardiac tissue</td>
<td>In vitro results easily compared to in vivo</td>
<td>Cloned ion channels lack the native conditions and structural architecture</td>
</tr>
<tr>
<td>Heterologous hERG</td>
<td>Experimental manipulations difficult to achieve in vivo</td>
<td>Time consuming</td>
</tr>
<tr>
<td>Purkinje Fibers</td>
<td>Ion concentration</td>
<td></td>
</tr>
<tr>
<td>Papillary muscle</td>
<td>Primary cultures can be established</td>
<td>Fibroblast contamination can be problematic</td>
</tr>
<tr>
<td>Transmural wedge</td>
<td>Readily proliferate under appropriate conditions</td>
<td>Interactions between different cell types not easily determined</td>
</tr>
<tr>
<td>Cell culture:</td>
<td>Effect of different ECM components can be examined</td>
<td>Tissue architecture not maintained</td>
</tr>
<tr>
<td>Endothelial</td>
<td>Low abundant proteins/complexes can be enriched</td>
<td>Pure fractions can be tedious and difficult to obtain</td>
</tr>
<tr>
<td>Myocyte</td>
<td>Complexity of sample can be reduced</td>
<td>Quality control of fractions time consuming</td>
</tr>
<tr>
<td>Smooth muscle</td>
<td></td>
<td>Possibility of fraction cross-contamination</td>
</tr>
</tbody>
</table>

Several groups today focus their studies on the decellularization of tissues and organs (Bader et al., 1998; Badylak et al., 1989, 1995; Booth et al., 2002; Chen et al., 1999, 2004; Cortiella et al., 2010; Freytes et al., 2004; Gilbert et al., 2005; Grauss et al., 2005; Kasimir et al., 2003; Korossis et al., 2002; Kropp et al., 1995; Ott et al., 2008; Petersen et al., 2010; Rieder et al., 2004; Schenke-Layland et al., 2003; Taylor, 2009; Uygun et al., 2010; Wainwright et al., 2010), including blood vessels (Conklin et al., 2002; Dahl et al., 2003; Schmidt et al., 2000; Uchimura et al., 2003), in order to obtain matrices for tissue engineering, providing a high number of quite different protocols and cover a wide range of techniques (Badylak et
The term "decellularization" refers to a series of methods by which the cellular membrane surrounding an isolated organ, or at least a part of it, is removed and the cytoplasmic and nuclear components are solubilised, leaving the complex mixture of structural and functional proteins that constitute the ECM intact, so minimizing any adverse effect on its composition, biological activity, and mechanical integrity. A decellularization protocol generally begins with the lysis of the cell membrane using physical treatments or ionic solutions, it is followed by separation of cellular components from the ECM using enzymatic treatments, solubilisation of cytoplasmic and nuclear cellular components using other detergents, and finally removal of cellular debris and residual chemicals from the tissue. All of these steps can be coupled with mechanical agitation to increase their effectiveness (Crapo et al., 2011; Gilbert et al., 2006). On the contrary, "recellularization" (or repopulation) refers to a regeneration process which occurs by contacting a decellularized organ or tissue (essentially the biological scaffold) with a population of regenerative cells (e.g. stem cells, either undifferentiated, partially or fully differentiated). These cells can be introduced (seeded) into the scaffold either by perfusion or, alternatively or in addition, by injection into one or more locations and at different densities depending on organ/tissue size (Badylak et al., 2011; Ott and Taylor, 2007).

On the basis of these concepts, the path we aim to follow is decellularizing blood vessels, taken from animals at the local abattoir, in order to obtain suitable biological scaffolds, and recellularize them with human mesenchymal stem cells harvested from human bone marrow, in order to obtain a vessel which could be the most as possible comparable to a human one. Then, the idea is to set up a novel human toxicity test, based on perfusion, for different substances, as requested by the REACH Regulation. From our best knowledge, indeed, this kind of approach has never been used for the evaluation of in vitro vascular toxicity, and could have several advantages for increasing knowledge in this field, while at the same time it seems to be in line with the very important concepts of improving alternative methods and of refinement, reduction and replacement of animal models. For the moment, only the first part of the project, i.e. the decellularization process, has been performed, and the corresponding results we obtained from histological examination of different protocols brought us to find a robust and effective method of decellularization, as reported in what follows.
4.2. Material and methods

4.2.1. Artery preparations

Descending and abdominal tracts of aorta samples from adult sheep were obtained from a local abattoir (CO.AL.BE. dei F.lli Contu & C.s.n.c., Cagliari, Italy). Tissue was immediately transported to the laboratory on ice, being careful that the warm ischemic time was no more than 1 h from the time of tissue extraction to processing or storage at -20°C. Aorta samples, after arriving at the laboratory, were always cleaned from fat and adherent tissue with a scalpel, and samples of approximately 2.5 cm in length were prepared, considering vessels with an internal diameter of 0.8-1 cm (cf. Figure 4.1). Artery segments were immediately washed for about 30 minutes at +4°C in a PBS solution (PBS, Sigma, St. Louis, MO, USA) containing 1% penicillin/streptomycin (Sigma), in order to avoid the risk of bacterial contamination. After that different kinds of treatments were performed.

4.2.2. Decellularization protocols

Different kinds of treatments, testing several substances alone and/or in suitable combination and varying the contact time, always at Room Temperature (RT, approximately 25°C) under stirring conditions to facilitate cell removal were performed. In all of the performed tests, a segment of fresh sample was immediately fixed in formalin to serve as control (CTRL). The treatments were always performed in a 100 ml-becker and the amount of solution was calculated considering approximately 15 ml per cm of length, in order to completely cover the sample. At least three repetitions per type of Protocol were performed independently.

Protocol A (A1, A2, A3, A4) uses a PBS solution containing 1% SDS (Sodium Dodecyl Sulfate, Sigma) for 3-6-12-24 h.

Protocol B (B1, B2, B3, B4) uses a PBS solution containing 1% Triton X-100 (Sigma) for 3-6-12-24 h.

Protocol C uses a PBS solution containing 0.01% Trypsin/EDTA (Sigma) for 24 h.
**Protocol D** uses a PBS solution containing 1% SDS for 6 h, a PBS solution containing 1% Triton X-100 for 16 h, and a PBS solution containing 0.01% Trypsin/EDTA for 6 h.

**Protocol E**, taken from the literature (Sawa *et al.*, 2005), uses a PBS solution containing 1% SDS for 48 h; the samples are then washed in a PBS solution for 24 h, treated in a PBS solution containing 1% Triton X-100 for 48 h, and washed again for about 72 h. All the solutions were changed every 24 h. Subsequent Protocols F and G were modified from Protocol E.

**Protocol F** uses a PBS solution containing 1% SDS for 24 h; the samples are then washed in a PBS solution for 24 h, treated in a PBS solution containing 1% Triton X-100 for 24 h, and washed again for about 72 h. All the solutions were changed every 24 h.

**Protocol G** uses a PBS solution containing 1% SDS for 48 h; the samples are then washed in a PBS solution for 12 h, treated in a PBS solution containing 1% Triton X-100 for 24 h, and washed again for about 48 h. All the solutions were changed every 24 h.

### 4.2.3. Histological analysis

Immediately after treatments, samples were fixed in formalin and afterwards dehydrated in an ascending series of alcohols, embedded in resin (Technovit 7100, Heraeus Kulzer GmbH & Co. KG, Wehrheim, Germany), and sectioned at a thickness of 3.5 µm (LKB 2218 BROMNA Historange Microtome). Sections were transferred to slides, oven-dried (45°C for 24 h), and finally stained with Hematoxylin and Eosin staining, Masson’s Trichrome staining, Silver Impregnation, Verhoeff and Weigert stainings, depending on the cases. The stained sections were then examined under a light microscope (Leica DM IL, Leica Microsystems, Germany) for the presence of cell nuclei and for analyzing morphology of collagen and elastic fibres. Histological sections were taken from both control and treated samples in each decellularization run and compared in order to verify tissue decellularization.
**Figure 4.1. Aorta samples.** (a) and (b) show aorta samples before cleaning procedure from fat and adherent tissues; (c) show the samples after the cleaning procedure; (d) shows the 2.5 cm length sample.

### 4.3. Results

Fresh porcine aorta (CTRL) samples showed the typical features of this vessel, with presence of cellular nuclei immersed in a matrix of collagen and elastic fibres. Nuclei are evidenced with Hematoxylin and Eosin staining (cf. **Figure 4.2**), while collagen fibres may be seen with both Masson's Trichrome staining (cf. **Figure 4.3**) and Silver Impregnation (cf. **Figure 4.4**), and elastic fibres may be highlighted with both Verhoeff staining (cf. **Figure 4.5**) and Weigert staining (cf. **Figure 4.6**).
Attempt to obtain decellularized scaffolds from ovine arteries were performed by means of different treatments, always at RT and under stirring conditions, as reported in the Material and Methods paragraph. Of all the tested treatments, only two of them resulted in an effective decellularization procedure, based on histological results and comparison with CTRL samples. After treatments with Protocols A and B, and under high powered light microscopic examination of the Hematoxylin and Eosin-stained specimens, results show that the cellular material was not solubilised, and nuclei were still present in the ECM. Therefore, these protocols were both considered to be ineffective. Only results from Protocol A are shown, because the ones related to Protocol B are essentially the same (cf. Figure 4.7).

Also Protocol C resulted to be ineffective. Indeed, even after the treatment with Trypsin/EDTA and examination with Hematoxylin and Eosin-staining, some of the nuclei were still present in the ECM. Moreover, in this case, an altered structure of the ECM structure was observed (cf. Figure 4.8).

On the contrary, Protocol D, after examination of the Hematoxylin and Eosin-stained specimens, resulted in a complete absence of nuclei. However, from the different staining performed, a clear destruction of the ECM was observed (cf. Figure 4.9), and therefore the Protocol is considered to be ineffective as well.

Literature Protocol E was instead considered to be effective, based on the absence of cellular or nucleic components in both the luminal surface and the underlying matrix scaffolding (cf. Figure 4.10) after examination of the Hematoxylin and Eosin-stained specimens. The staining revealed that the vascular matrices processed with the described decellularization technique were indeed all acellular. Moreover, the structure of the ECM resulted well preserved.

Unfortunately, attempts to decrease the time required for Protocol E, i.e. treating samples with Protocol F, did not provide successful results, since cellular nuclei were still present after the treatment (cf. Figure 4.11).

Finally, Protocols G was instead considered to be effective, on the same bases reported for Protocol E. The matrix resulted to be indeed well preserved and acellular (cf. Figure 4.12).
Further histological analysis were performed for the case of these successful protocols (Protocols E and G), thus confirming a well preserved decellularized vascular ECM. It was shown indeed that collagen (Masson’s Trichrome Staining, cf. Figure 4.13) and elastic fibres (Verhoeff Staining, cf. Figure 4.14) kept their original porous morphology and structure. Only results from Protocol G are shown, because those ones related to Protocol E are essentially the same.

4.4. Discussion

The aim of the present study was to find the most effective and least destructive procedure, among different protocols, to be used in order to obtain the most as possible preserved decellularized matrix from ovine blood vessels (in the specific case aorta). In all the cases, sample agitation was guaranteed since, as it is reported in literature, physical methods of that type, can increase the effectiveness of the process, facilitating cellular lysis and removal of the cell content from the ECM. However, such physical treatments are generally insufficient to achieve a complete decellularization, and therefore they must be combined also with enzymatic (e.g. trypsin) and chemical ones (e.g. ionic and non ionic solutions or detergents), which disrupt cell membranes and the bonds responsible for inter- and extracellular connections (Gilbert et al., 2006). For this reason, we performed different kinds of treatments, in order to compare the decellularization potential of three of the most used decellularization agents, alone and/or in combination under different contact times.

Firstly, we chose to use an ionic detergent, the Sodium Dodecyl Sulfate, since it results to be very effective for removal of cellular components from tissues, even more than other agents. Since SDS is an ionic detergent, it solubilizes cytoplasmic and nuclear cellular membranes and tends to denature cytoplasmic proteins. We chose to use a non ionic detergent as well, the Triton X-100, since such detergents are quite extensively used in these kinds of treatments because of their relative mild effects on tissue structure. Lipid-lipid and lipid-protein interactions are then disrupted, while leaving protein-protein interactions intact (Crapo et al., 2011; Gilbert et al., 2006; Seddon et al., 2004). Finally, we performed also enzymatic treatments, using Trypsin/EDTA. Trypsin is a proteolytic enzyme that cleaves peptide bonds (Voet et al., 2002), while EDTA is a chelating agent that disrupt cell adhesion to ECM,
thereby facilitating cell removal, while tending to disrupt the ECM ultrastructure (Gilbert et al., 2006; Waldrop et al., 1980).

The evaluation of the protocol efficiency to verify the removal of cellular material from the tissues and also the preservation of the ECM is essential. It can be assessed by several methods, including for example histological staining, necessary to examine for the presence of nuclei or various cytoplasmic and extracellular molecules, such as collagen, adhesion proteins like fibronectin and laminin, GAGs, elastic fibres, etc. Other methods comprise immunohistochemistry, to analyze specific intracellular proteins, such as actin and vimentin, inspection for the presence of DNA using fluorescent molecules, electron microscopic methods to examine the presence of remnant nuclear material or cytoplasmic debris, and also mechanical testings for the check of the presence and integrity of structural proteins within the scaffold (Crapo et al., 2011; Gilbert et al., 2006). The present study focused on the histological analysis of the different applied protocols, using hematoxylin and eosin to firstly examine the presence or absence of nuclei, and alternative staining to verify the structure of collagen and elastic fibres, which are two of the main components of a blood vessel.

We can conclude from our results that a combination of different agents is preferable to the use of a single one. Indeed, we demonstrated that, in the case of ovine aorta, a treatment with a single agent (cf. Protocol A, B and C) do not result in decellularization, since nuclei were not removed from the ECM (cf. Figures 4.7 and 4.8).

Instead of further increasing the time of treatment with each one of the single agents, these results drove us to test a protocol that may be considered as a combination of Protocols A, B, and C (Protocol D). The reason that brings us to such a choice is that it is probably better to combine different agents in order to act simultaneously in different directions and obtain a more successful removal of the nuclear material. Even if, after this treatment, the nuclei of the cells resulted to be absent, the structure of the ECM, as shown in Figure 4.9, resulted to be definitely distorted, compared to the CTRL structure (cf. Figures 4.2, 4.4 and 4.6). Such distortion is obviously not acceptable from the decellularization point of view, because its goal is indeed the solubilisation of cytoplasmic and nuclear components, while leaving intact the complex mixture of structural and functional proteins that constitute the ECM (Gilbert et al., 2006). In fact, the final aim of such process is to obtain a biological scaffold to be recellularized, and it is known that composition and ultrastructural characteristics of the ECM
play an important role in cells’ attachment, thus regulating their ability to migrate into and attach to specific locations (Brown et al., 2006), and also influence tissue-specific phenotypic differentiation (Badylak et al., 2011; Cortiella et al., 2010; Gong et al., 2008; Sellaro et al., 2010). In our opinion, the distortion observed in the results could be mainly due to the combined presence of Trypsin/EDTA. As already anticipated indeed, Trypsin/EDTA tend to disrupt the ECM ultrastructure, affecting ECM proteins such as collagen (Waldrop et al., 1980), which is one of the main constituents of the matrix (especially in blood vessels) (Wagenseil and Mecham, 2009). Such opinion is even supported from results obtained after treatment with Protocol C, which is the only other one were Trypsin/EDTA was used. Also in this case, indeed, as shown in Figure 4.8, a distortion of the matrix occurred, even on a small scale. Based on this conclusion and on the fact that for our further purposes we need to obtain a vessel the most as possible similar to an operating one, we decided to avoid the use of Trypsin/EDTA, even at lower concentration and for shorter periods of time.

Successful and effective treatments resulted to be those ones where Protocol E and G were applied. Nuclei resulted to be totally absent (cf. Figures 4.10 and 4.12) and the ECM structure is much better organized, with fibres (mainly collagen and elastic fibres) more linear, regular and close to each other (cf. Figures 4.13, and 4.14). Protocol G was modified from Protocol E, which has been taken from the literature (Sawa et al., 2005), in order to improve it, while better preserving ECM structure and saving processing time. The obtained results were satisfying and encouraging. By cutting by 50% the time required for the action of Triton X-100, indeed, an appropriate decellularization was also obtained, with a complete solubilisation of the nuclear material. The attempt to decrease by 50% the time required for Protocol E (Protocol F), in order to shorter the contact of the detergents with the ECM, was unfortunately not successful. We decided, therefore, to cut the time required only by one of the used agents. The choice was to consider Triton X-100 because the SDS appears to be more effective than Triton X-100 in removing nuclei, while preserving tissue mechanics, aspect which is of course essential for blood vessels functioning. Moreover, its addition to a decellularization protocol can make the difference between complete and incomplete cell nuclei removal (Crapo et al., 2011). Alternative staining demonstrated that Protocol G was able to maintain an appropriate ECM structure, both in terms of collagen and elastic fibres (cf. Figures 4.13 and 4.14 in comparison with Figures 4.3 and 4.5), and therefore we can conclude that, even if further works are required in order to reach the final goal of our study,
these results are already quite encouraging and promising, since we have improved quite long
decellularization protocols in terms of time and costs, thus assuring at the same time a better
preservation of the extracellular matrix, which is essential for the next steps of the process.

This decellularized ECM from ovine aorta could be used in a subsequent phase of the study
as a biological scaffold where seeding human MSCs in order to obtain a vessel which could
be the most as possible comparable to a human one, and with the final aim of setting up a
novel in vitro human toxicity test based on perfusion. Our choice to use human MSCs for this
purpose has to be related to the unique properties of this cell type, such as unlimited
proliferation ability and being a readily available source of human cells, but mainly because
of their capability to generate other cell types (Davila et al., 2004; Laustriat et al., 2010;
Trosko and Chang, 2010; Wobus and Loser, 2011). With regard to blood vessels’ population
of cells, for example, recent in vitro-studies presented the differentiation ability of MSCs to
differentiate into cells with an endothelial phenotype as well as into smooth muscle cells
(Gong and Niklason, 2008; Hashi et al., 2007). Although further studies have to be made for
completely fulfil the high expectations in this field, MSCs remains very promising and, by
means of their selective isolation and expansion in vitro and specific stimulation (growth
factors, culture media, etc.), endothelial and vascular muscle cells can be propagated and
used for the repopulation of a decellularized vessel (Aper et al., 2009).

The new in vitro test could then be used for the testing of different substances, as requested
by the REACH Regulation, being therefore of primary importance and in line with what is
requested today by the legislation, and in our opinion could have several advantages. Firstly,
difficulties in obtaining vessels from patients or cadavers would be avoided, enlarging
therefore the possible number of replicates that could be processed at any time. By taking into
account the alternative of using animal vessels, the repopulation of the matrix, considered
merely as a biological scaffold, with human cells would solve the problem of species-specific
differences, and extrapolation and assumptions, which is one of the main problems arising
from animal experimentation. However, since discrepancies can arise between in vivo and in
vitro testing (Bhogal, et al., 2005; Blauiboer, 2002a, 2002b; Flint, 1990; Piersma, 2006;
Zucco, et al., 2004), experimental animals cannot be avoided at all, as in toxicology testing in
general, but the proposed method could be considered as a reduction alternative, in line with
the 3Rs principle proposed by Russel and Burch (1959), since it could serve as a first line of
inspection and the number of animals required for the tests would be considerably reduced.
Moreover, it should be noticed that, the possibility to obtain stem cells from individual human subjects and therefore the future possibility of obtain an *in vitro* vascular system which is specific for specific patients, would offer unprecedented opportunities to analyze the contribution of genetic background that affects susceptibility to toxicity and particularly vascular susceptibility.

**Figure 4.2. Hematoxylin and Eosin staining for the CTRL samples.** Results from Hematoxylin and Eosin staining, which is one of the most common methods used in histology and anatomy, confirm the presence of cellular nuclei in CTRL samples. The nuclei of the cells are stained blue-purple with Hematoxylin, while eosinophilic structures, such as cytoplasm, connective tissue and other extracellular substances are counterstained bright pink with Eosin. Different magnifications are shown.
Figure 4.3. Masson's Trichrome staining for the CTRL samples. Results from Masson's Trichrome staining, a three-color staining protocol commonly used in histology for the detection of fibres in tissues, confirm the typical features of CTRL samples. Collagen fibres are stained blue, while nuclei are stained dark purple and the background - acidophilic cytoplasm and muscle fibres - is stained pink. Different magnifications are shown.
Figure 4.4. Silver impregnation for the CTRL samples. Results from Silver Impregnation, another staining protocol commonly used in histology for the detection of fibres in tissues, confirm the typical features of CTRL samples. Fibrillar collagen, which represents the structural support for cells, is stained in yellow, while reticular fibres, the more immature form of the collagen, are stained in black. Different magnifications are shown.

Figure 4.5. Verhoeff staining for the CTRL samples. Results from Verhoeff staining, a technique used to detect elastic fibres in tissues, confirm their typical features in CTRL samples. Elastic fibres are stained in black. Different magnifications are shown.
Figure 4.6. **Weigert staining for the CTRL samples.** Results from Weigert staining, another technique used to detect elastic fibres in tissues, confirm their typical features in CTRL samples. Elastic fibres are stained in purple. Different magnifications are shown.

Figure 4.7. **Hematoxylin and Eosin staining for the Protocol A-treated samples.** Results show that the cellular material was not solubilised after treatment with Protocol A, and nuclei were still present in the ECM. Nuclei are stained blue-purple.
Figure 4.8. Hematoxylin and Eosin staining for the Protocol C-treated samples. Results show that after the treatment with Protocol C and examination with Hematoxylin and Eosin-staining, some of the nuclei (stained in blue) were still present in the ECM, and an altered structure of the ECM structure was observed. Different magnifications are shown.
Figure 4.9. Different stainings for the Protocol D-treated samples. (a) Hematoxylin and Eosin staining; (b) Silver Impregnation; (c) Weigert staining. Results show that the treatment with Protocol D resulted in a complete absence of nuclei; however, a clear destruction of the ECM was observed.
Figure 4.10. *Hematoxylin and Eosin staining for the Protocol E-treated samples.* Results show that the treatment with Protocol E resulted in the absence of cellular or nucleic components in both the luminal surface and the underlying matrix, and the structure of the ECM resulted well preserved. Different magnifications are shown.

![Figure 4.10](image1.jpg)

10X 20X

Figure 4.11. *Hematoxylin and Eosin staining for the Protocol F-treated samples.* Results show that cellular nuclei (stained in blue) were still present after the Protocol F- treatment. Different magnifications are shown.

![Figure 4.11](image2.jpg)

20X 40X
Figure 4.12. *Hematoxylin and Eosin staining for the Protocol G-treated samples.* Results show that the treatment with Protocol G resulted in the absence of cellular or nucleic components in both the luminal surface and the underlying matrix, and the structure of the ECM resulted to be well preserved. Different magnifications are shown.
Figure 4.13. Masson’s Trichrome Staining for the Protocol G-treated samples. Results show that, after the treatment with Protocol G, collagen fibres kept their original morphology and structure.
Figure 4.14. *Verhoeff Staining for the Protocol G-treated samples.* Results show that, after the treatment with Protocol G, elastic fibres kept their original morphology and structure.
Conclusions

Today, modern society is becoming constantly more dependent on the use of a wide range of different chemicals, that provide substantial highly appreciated benefits, but at the same time also have the potential to cause damages to the environment and human health. Therefore, in order to ensure an adequate protection, it is essential to evaluate their potential to cause adverse effects and to provide reliable and adequate information on them through a process named risk assessment. Toxicological testing is therefore an important part of the regulatory safety assessment for chemicals worldwide. Currently, test methods rely largely on animal models, even if the utility of the data obtained with \textit{in vivo} tests has been often questioned, due to scientific, economic and ethical concerns. These factors, together with legislative changes (among the others the REACH Regulation), are driving toxicologists to explore the potential of non-animal alternative approaches to assess toxicity in the pharmaceutical and chemical industries. Recently indeed, many \textit{in vitro} methods for toxicity testing have been developed as alternatives to whole animal tests, according to the 3Rs approach – Refinement, Reduction, and Replacement – started in 1959 by Russell and Burch. Clearly, especially replacement and/or reduction of unnecessary \textit{in vivo} tests, that today has reached a high priority level and has been heavily promoted, would have significant animal welfare benefits and would also result in lower testing costs and time.

The activity performed during the PhD falls within the framework of the risk assessment of industrial/commercial chemicals, and specifically in the REACH Regulation, giving specific relevance to the 3Rs concept. It concerns, indeed, the development of novel \textit{in vitro} models in order to test the potential toxicity of drugs and chemicals, following two different lines of research, both of them based on the use - for the first time - of human Mesenchymal Stem Cells (MSCs) isolated from bone marrow. From our best knowledge indeed, both animal and human MSCs have never been adopted for developing \textit{in vitro} model systems for acute toxicity tests, but their unique proprieties, such as unlimited proliferation ability, plasticity to generate others cell types and even being a more readily available sources of human cells, clearly identify their potential benefits in toxicology. Moreover, they would permit to overcome limitations related to ESCs and iPS cells, such has ethical concerns and potentially tumorigenic status.
The research activity has been firstly focused on the application of a selected *in vitro* toxicity method, the NRU assay, to detect chemical toxicity on human MSCs. Specifically, the aim of this study was to evaluate the applicability of human MSCs as cell line for *in vitro* cytotoxicity tests to correctly predict LD$_{50}$ and the hazard category of chemicals according to the Globally Harmonized System of Classification – GHS. By comparing the behavior of MSCs with the already validated 3T3 and NHK NRU test methods (validated by ICCVAM in 2006), the results clearly show that the tested cell line can confidently be used to perform *in vitro* acute toxicity tests, providing a more accurate modeling of *in vivo* conditions. Indeed, the ability to test chemicals using a human system would increase the relevance and accuracy of predicting toxicological outcomes. Moreover, the ability of stem cells to differentiate into a variety of cell types and develop into organ system could allow them to replace transformed cell lines and primary cells for *in vitro* studies, thus eliminating potential limitations and improving the relevance of predictive assay. Another factor to be considered is that the ability to derive stem cells from individual human subjects would offer unprecedented opportunities to analyze the contribution of genetic background that affect susceptibility to toxicity. There are also additional aspects why human MSCs could be a better option than the two already validated cell lines. In this regard, it should be pointed out that 3T3 and NHK NRU methods are comparable, while the ICCVAM suggests the use of 3T3 cells to conduct these kinds of tests because of the lower costs compared to the use of NHK cells. On the other hand, 3T3 do not represent a human system that in general should be preferred. Along these lines, human MSCs might represent an ideal alternative to NHK, since the corresponding costs of related experiments are comparable to 3T3. Furthermore, the proposed method will enable toxicity evaluation of novel chemicals in *in vitro* studies, and could also be used for predicting starting doses for *in vivo* toxicity studies, so that a substantial reduction of tested animals and experiment costs will result as well. For all these reasons, it is expected that the study carried out, even if on a relatively small number of toxicants, and even considering that, however, the proposed cell line, as well as 3T3 and NHK, is not able to correctly predict the right hazard category of the most toxic chemicals (GHS 1 and 2 classes) due to a number of potential factors, will be a step forward to current and future validation efforts to develop alternative methods to replace animal testing.

Next, the research activity was mainly focused on the research of new *in vitro* methods in order to test, specifically, the vascular toxicity of chemicals. This research is based on the
application of emerging tissue engineering concepts in the field of toxicology. For the moment, only preliminary studies were conducted, which regarded the application of decellularization techniques to blood vessels harvested from animals at the local abattoir, in order to obtain biological scaffolds. These scaffolds could be subsequently used for recellularization processes using human MSCs, and setting up of perfused in vitro human toxicity tests. Of course, it is necessary to be sure that the animal vessel to be used as scaffold is completely acellular. The study presented in this thesis focused, therefore, on the histological evaluation of the different decellularization protocols performed, in order to find the most effective and least destructive procedure to be used in order to obtain the most as possible preserved decellularized matrix from ovine blood vessels (in the specific case aorta). Firstly, the samples were subjected to agitation because physical methods, such as agitation, can increase the effectiveness of the process, facilitating cellular lysis and removal of the cell content from the ECM. However, physical treatments like this are generally insufficient to achieve a complete decellularization, and therefore they must be combined also with enzymatic (e.g. trypsin) and chemical ones (e.g. ionic and non ionic solutions or detergents), which again disrupt cell membranes and the bonds responsible for inter- and extracellular connections. Among the different tested protocols, the obtained results let us to conclude that an optimized effective protocol has been achieved using a combination of two different chemicals, a ionic detergent (SDS) and a non ionic one (Triton X-100). Results show, indeed, that after this treatment the scaffold resulted to be completely acellular and an appropriate ECM structure was maintained, both in terms of collagen and elastic fibres, characteristics that are the goal of the perfect decellularization protocol. Therefore, we can conclude that, even if further works are required in order to reach the final goal of our study and setting up a new in vitro test for vascular toxicity, these results are already quite satisfying and promising, since we have improved a quite long decellularization protocol in terms of time and costs, assuring at the same time a better preservation of the extracellular matrix, which is essential for the next steps of the process.

We can generally conclude that the activities performed during this thesis work represent a step forward to current and future validation efforts to develop alternative methods to replace animal testing, perfectly fitting in one of the most current hot topic in toxicological research and being in line with what is today highly required from scientific, economic and ethical points of view, as well as from European legislation.
Appendix 1

Specific References in the Regulation to animal testing

Alternatives to animal testing are strongly promoted throughout the REACH text. All the Regulation recitals and articles where the matter is handled are report in what follows (European Parliament, 2006, 2007).

Recital (1). Promotion of the development of Alternative Methods for the assessment of hazards of substances. This Regulation should ensure a high level of protection of human health and the environment as well as the free movement of substances, on their own, in preparations and in articles, while enhancing competitiveness and innovation. This Regulation should also promote the development of alternative methods for the assessment of hazards of substances.

Recital (13). Cosmetics. This Regulation should apply without prejudice to the prohibitions and restrictions laid down in Council Directive 76/768/EEC of 27 July 1976 on the approximation of the laws of the Member States relating to cosmetic products in so far as substances are used and marketed as cosmetic ingredients and are within the scope of this Regulation. A phase-out of testing on vertebrate animals for the purpose of protecting human health as specified in Directive 76/768/EEC should take place with regard to the uses of those substances in cosmetics.

Recital (33). Joint submission and sharing of information to reduce testing on vertebrate animals. Joint submission and the sharing of information on substances should be provided for in order to increase the efficiency of the registration system, to reduce costs and to reduce testing on vertebrate animals. One of a group of multiple registrants should submit information on behalf of the others according to rules which ensure that all the required information is submitted, while allowing sharing of the costs burden. A registrant should be able to submit information directly to the Agency in certain specified cases.
Recital (37). Tests to comply with requirements of protection of laboratory animals (86/609/EEC) and GLP (Good laboratory Practice). If tests are performed, they should comply with the relevant requirements of protection of laboratory animals, set out in Council Directive 86/609/EEC of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes, and, in the case of ecotoxicological and toxicological tests, good laboratory practice, set out in Directive 2004/10/EC of the European Parliament and of the Council of 11 February 2004 on the harmonisation of laws, regulations and administrative provisions relating to the application of the principles of good laboratory practice and the verification of their application for tests on chemical substances.

Recital (40). Commission, Member States, industry and other stakeholders should continue to contribute to promotion of alternative test methods. The Commission, Member States, industry and other stakeholders should continue to contribute to the promotion of alternative test methods on an international and national level including computer supported methodologies, in vitro methodologies, as appropriate, those based on toxicogenomics, and other relevant methodologies. The Community's strategy to promote alternative test methods is a priority and the Commission should ensure that within its future Research Framework Programmes and initiatives such as the Community Action Plan on the Protection and Welfare of Animals 2006 to 2010 this remains a priority topic. Participation of stakeholders and initiatives involving all interested parties should be sought.

Recital (47). In accordance with Directive 86/609/EEC, it is necessary to replace, reduce, and refine testing on vertebrates. In accordance with Directive 86/609/EEC, it is necessary to replace, reduce or refine testing on vertebrate animals. Implementation of this Regulation should be based on the use of alternative test methods, suitable for the assessment of health and environmental hazards of chemicals, wherever possible. The use of animals should be avoided by recourse to alternative methods validated by the Commission or international bodies, or recognised by the Commission or the Agency as appropriate to meet the information requirements under this Regulation. To this end, the Commission, following consultation with relevant stakeholders, should propose to amend the future Commission Regulation on test methods or this Regulation, where appropriate, to replace, reduce or refine
animal testing. The Commission and the Agency should ensure that reduction of animal testing is a key consideration in the development and maintenance of guidance for stakeholders and in the Agency's own procedures.

Recital (49). Sharing of information, in particular to reduce animal testing. In order to avoid duplication of work, and in particular to reduce testing involving vertebrate animals, the provisions concerning preparation and submission of registrations and updates should require sharing of information where this is requested by any registrant. If the information concerns vertebrate animals, the registrant should be obliged to request it.

Recital (50). Sharing of information, in particular to reduce animal testing. It is in the public interest to ensure the quickest possible circulation of test results on human health or environmental hazards of certain substances to those natural or legal persons which use them, in order to limit any risks associated with their use. Sharing of information should occur where this is requested by any registrant, in particular in the case of information involving tests on vertebrate animals, under conditions that ensure a fair compensation for the company that has undertaken the tests.

Recital (64). To prevent unnecessary testing, interested parties have 45 days to comment on testing proposals. In order to prevent unnecessary animal testing, interested parties should have a period of 45 days during which they may provide scientifically valid information and studies that address the relevant substance and hazard end-point, which is addressed by the testing proposal. The scientifically valid information and studies received by the Agency should be taken into account for decisions on testing proposals.

Article 2(4b). Application. (4) REACH regulation shall apply without prejudice to (b) Directive 76/768/EEC (cosmetics) as regards testing involving vertebrate animals within the scope of that Directive.
Article 2(6b). Application. (6) The provisions of Title IV shall not apply to the following preparations in the finished state, intended for the final user: (b) cosmetic products as defined in Directive 76/768/EEC.

Article 13(1). General requirements for generation of information on intrinsic properties of substances. (1) Information on intrinsic properties of substances may be generated by means other than tests, provided that the conditions set out in Annex XI are met. In particular for human toxicity, information shall be generated whenever possible by means other than vertebrate animal tests, through the use of alternative methods, for example, in vitro methods or qualitative or quantitative structure-activity relationship models or from information from structurally related substances (grouping or read-across). Testing in accordance with Annex VIII, Sections 8.6 and 8.7, Annex IX and Annex X may be omitted where justified by information on exposure and implemented risk management measures as specified in Annex XI, section 3.

Article 13(2). General requirements for generation of information on intrinsic properties of substances. (2) These methods shall be regularly reviewed and improved with a view to reducing testing on vertebrate animals and the number of animals involved. The Commission, following consultation with relevant stakeholders, shall, as soon as possible, make a proposal, if appropriate, to amend the Commission Regulation on test methods adopted in accordance with the procedure referred to in Article 133(4), and the Annexes of this Regulation, if relevant, so as to replace, reduce or refine animal testing. Amendments to that Commission Regulation shall be adopted in accordance with the procedure specified in paragraph 3 and amendments to the Annexes of this Regulation shall be adopted in accordance with the procedure referred to in Article 131.

Article 13(4). General requirements for generation of information on intrinsic properties of substances. (4) Ecotoxicological and toxicological tests and analyses shall be carried out in compliance with the principles of good laboratory practice provided for in Directive 2004/10/EC or other international standards recognised as being equivalent by the Commission or the Agency and with the provisions of Directive 86/609/EEC, if applicable.
Article 14(5b). *Chemical safety report and duty to apply and recommend risk reduction measures.* (5) The chemical safety report need not include consideration of the risks to human health from the following end uses: (b) in cosmetic products within the scope of Directive 76/768/EEC.

Article 25(1). *Objectives and general rules.* (1) In order to avoid animal testing, testing on vertebrate animals for the purposes of this Regulation shall be undertaken only as a last resort. It is also necessary to take measures limiting duplication of other tests.

Article 26(1c). *Duty to inquire prior to registration.* (1) Every potential registrant of a non-phase-in substance, or potential registrant of a phase-in substance who has not preregistered in accordance with Article 28, shall inquire from the Agency whether a registration has already been submitted for the same substance. He shall submit all the following information to the Agency with the inquiry: (c) which information requirements would require new studies involving vertebrate animals to be carried out by him;

Article 26(3). *Duty to inquire prior to registration.* (3) If the same substance has previously been registered less than 12 years earlier, the Agency shall inform the potential registrant without delay of the names and addresses of the previous registrant(s) and of the relevant summaries or robust study summaries, as the case may be, already submitted by them. Studies involving vertebrate animals shall not be repeated. The Agency shall simultaneously inform the previous registrants of the name and address of the potential registrant. The available studies shall be shared with the potential registrant in accordance with Article 27.

Article 27(1a). *Sharing of existing data in the case of registered substances.* (1) Where a substance has previously been registered less than 12 years earlier as referred to in Article 26(3), the potential registrant: (a) shall, in the case of information involving tests on vertebrate animals; and (b) may, in the case of information not involving tests on vertebrate animals, request from the previous registrant(s) the information he requires with respect to Article 10(a)(vi) and (vii) in order to register.
Article 30(1). Sharing of data involving tests. (1) Before testing is carried out in order to meet the information requirements for the purposes of registration, a SIEF participant shall inquire whether a relevant study is available by communicating within his SIEF. If a relevant study involving tests on vertebrate animals is available within the SIEF, a participant of that SIEF shall request that study. If a relevant study not involving tests on vertebrate animals is available within the SIEF, a SIEF participant may request that study.

Within one month of the request, the owner of the study shall provide proof of its cost to the participant(s) requesting it. The participant(s) and the owner shall make every effort to ensure that the costs of sharing the information are determined in a fair, transparent and non-discriminatory way. This may be facilitated by following any cost sharing guidance which is based on those principles and is adopted by the Agency in accordance with Article 77(2)(g). If they cannot reach such an agreement, the cost shall be shared equally. The owner shall give permission to refer to the full study report for the purpose of registration within two weeks of receipt of payment. Registrants are only required to share in the costs of information that they are required to submit to satisfy their registration requirements.

Article 30(3). Sharing of data involving tests. (3) If the owner of a study as referred to in paragraph 1 which involves testing on vertebrate animals refuses to provide either proof of the cost of that study or the study itself to (an) other participant(s), he shall not be able to proceed with registration until he provides the information to the other participants(s). The other participant(s) shall proceed with registration without fulfilling the relevant information requirement, explaining the reason for this in the registration dossier. The study shall not be repeated unless within 12 months of the date of registration of the other participant(s), the owner of this information has not provided it to them and the Agency decides that the test should be repeated by them. However, if a registration containing this information has already been submitted by another registrant, the Agency shall give the other participant(s) permission to refer to the information in his registration dossier (s). The other registrant shall have a claim on the other participant(s) for an equal share of the cost, provided he makes the full study report available to the other participant(s), which shall be enforceable in the national courts.
Article 38(2f). Obligation for downstream users to report information. (2) The information reported by the downstream user shall include the following: (f) except where the downstream user is relying on the exemption in Article 37(4)(c), a proposal for additional testing on vertebrate animals, where this is considered necessary by the downstream user to complete his chemical safety assessment.

Article 40(2). Examination of testing proposals. (2) Information relating to testing proposals involving tests on vertebrate animals shall be published on the Agency website. The Agency shall publish on its website the name of the substance, the hazard end-point for which vertebrate testing is proposed, and the date by which any third party information is required. It shall invite third parties to submit, using the format provided by the Agency, scientifically valid information and studies that address the relevant substance and hazard endpoint, addressed by the testing proposal, within 45 days of the date of publication. All such scientifically valid information and studies received shall be taken into account by the Agency in preparing its decision in accordance with paragraph 3.

Article 56(5a). General provisions. (5) In the case of substances that are subject to authorization only because they meet the criteria in Article 57(a), (b) or (c) or because they are identified in accordance with Article 57(f) only because of hazards to human health, paragraphs 1 and 2 of this Article shall not apply to the following uses: (a) uses in cosmetic products within the scope of Directive 76/768/EEC;

Article 67(2). General provisions. (2) Paragraph 1 shall not apply to the use of substances in cosmetic products, as defined by Directive 76/768/EEC, with regard to restrictions addressing the risks to human health within the scope of that Directive.

Article 117(3). Reporting. (3) Every three years the Agency, in accordance with the objective of promoting non-animal testing methods, shall submit to the Commission a report on the status of implementation and use of non-animal test methods and testing strategies used to generate information on intrinsic properties and for risk assessment to meet the requirements of this Regulation. The first report shall be submitted by 1 June 2011.
Article 117(4b). Reporting. (4) Every five years, the Commission shall publish a general report on: (a) the experience acquired with the operation of this Regulation, including the information referred to in paragraphs 1, 2 and 3 and; (b) the amount and distribution of funding made available by the Commission for the development and evaluation of alternative test methods. The first report shall be published by 1 June 2012.

Article 138(3). Review. (3) The report, referred to in Article 117(4), on the experience acquired with the operation of this Regulation shall include a review of the requirements relating to registration of substances manufactured or imported only in quantities starting at one tonne but less than 10 tonnes per year per manufacturer or importer. On the basis of that review, the Commission may present legislative proposals to modify the information requirements for substances manufactured or imported in quantities of one tonne or more up to 10 tonnes per year per manufacturer or importer, taking into account the latest developments, for example in relation to alternative testing and (quantitative) structure-activity relationships ((Q)SARs).

Article 138(9). Review. (9) In accordance with the objective of promoting non-animal testing and the replacement, reduction or refinement of animal testing required under this Regulation, the Commission shall review the testing requirements of Section 8.7 of Annex VIII by 1 June 2019. On the basis of this review, while ensuring a high level of protection of health and the environment, the Commission may propose an amendment in accordance with the procedure referred to in Article 133(4).

Annex I – GENERAL PROVISIONS FOR ASSESSING SUBSTANCES AND PREPARING CHEMICAL SAFETY REPORTS.

4.1. Step 1: Comparison with the Criteria. This part of the PBT and vPvB assessment shall entail the comparison of the available information, which is submitted as part of the technical dossier, with the criteria given in Annex XIII and a statement of whether the substance fulfils or does not fulfil the criteria.
If the available information is not sufficient to decide whether the substance fulfils the criteria in Annex XIII, then other evidence like monitoring data available for the registrant and giving rise to an equivalent level of concern shall be considered on a case-by-case basis.

If the technical dossier contains for one or more endpoints only information as required in Annexes VII and VIII, the registrant shall consider information relevant for screening for P, B and T properties to decide whether further information needs to be generated to fulfil the objective of the PBT and vPvB assessment. In case the generation of further information is necessary and would require testing on vertebrate animals, the registrant shall submit a testing proposal. However, such further information does not need to be generated if the registrant implements or recommends sufficient risk management measures and operational conditions that enable derogation according to Section 3 of Annex XI from testing relevant for PBT and vPvB assessment.

**Annex VI – INFORMATION REQUIREMENTS REFERRED TO IN ARTICLE 10**

**Step 1 - Gather And Share Existing Information.**

The registrant should gather all existing available test data on the substance to be registered, this would include a literature search for relevant information on the substance. Wherever practicable, registrations should be submitted jointly, in accordance with Articles 11 or 19. This will enable test data to be shared, thereby avoiding unnecessary testing and reducing costs. The registrant should also collect all other available and relevant information on the substance regardless whether testing for a given endpoint is required or not at the specific tonnage level. This should include information from alternative sources (e.g. from (Q)SARs, read-across from other substances, *in vivo* and *in vitro* testing, epidemiological data) which may assist in identifying the presence or absence of hazardous properties of the substance and which can in certain cases replace the results of animal tests.

In addition, information on exposure, use and risk management measures in accordance with Article 10 and this Annex should be collected. Considering all this information together, the registrant will be able to determine the need to generate further information.

**Step 4 - Generate New Data/Propose Testing Strategy.**

In some cases it will not be necessary to generate new data. However, where there is an information gap that needs to be filled, new data shall be generated (Annexes VII and VIII),
or a testing strategy shall be proposed (Annexes IX and X), depending on the tonnage. New tests on vertebrates shall only be conducted or proposed as a last resort when all other data sources have been exhausted.

In some cases, the rules set out in Annexes VII to XI may require certain tests to be undertaken earlier than or in addition to the standard requirements.
Literature Cited


Gong, Z, and Niklason, LE. (2008)."Small-diameter human vessel wall engineered from bone marrow-derived mesenchymal stem cells (hMSCs)." The FASEB Journal 22(6):1635-1648.


ICCVAM. (2003). "ICCVAM guidelines for the nomination and submission of new, revised and alternative test methods." NIH Publication No. 03-4508, National Institute for Environmental Health Sciences, Research Triangle Park, NC.


Marra, DE, Simoncini, T, and Liao, JK. (2000). "Inhibition of vascular smooth muscle cell proliferation by sodium salicylate mediated by upregulation of p21 (Waf1) and p27 (Kip1)." Circulation 102(17):2124-2130.


Moore, JC, Fu, J, Chan, YC, Lin, D, Tran, H, Tse, HF, and Li, RA. (2008). "Distinct cardiogenic preferences of two human embryonic stem cell (hESC) lines are imprinted in their proteomes in the pluripotent state." Biochemical and Biophysical Research Communications 372(4):553-558.


There are a number of people that I would like to thank for their help and guidance in the past three years. First of all, I would like to sincerely thank Prof. Giacomo Cao for giving me the opportunity to carry out the PhD activity. Then, I really want to express great thanks to Dr. Luisa Mancuso for her essential support, for both scientific and personal advices, for her patience and availability, and for all the fun we have had in the last three years. I am grateful also to all the other colleagues for their valuable support and help. Among them, I would like to thank in particular Clara and Gianluca for all the chatting and exchange of opinions during lunch times, and Sandro for all the advices and teasings at the same time. I would also like to thank Dr. Cristinetta Porcu for her technical assistance in histology. Last but not the least, special thanks go to my family, who have always been there to support and encourage me in all my endeavours.