SEVENTH FRAMEWORK PROGRAMME
THEME 1 – HEALTH

SCIENTIFIC AND TECHNOLOGICAL ISSUES IN 3RS ALTERNATIVES RESEARCH IN THE PROCESS OF DRUG DEVELOPMENT AND UNION POLITICS

START-UP

support action n° 201187

ecopa

Vrije Universiteit Brussel


START-UP

Table of Contents

1. ABBREVIATIONS 9
2. INTRODUCTION 14
3. CONCEPT AND PROJECT OBJECTIVES 15
4. OVERALL RESULTS 17
   4.1. Overall Executive Summary 17
   4.2. Overall Recommendations 21
   4.3. Road Maps 23
5. DETAILED RESULTS 27
   5.1. Expert Meeting 1 31
      5.1.1. Executive Summary 31
      5.1.2. Recommendations 32
      5.1.3. Report of Expert Meeting 1 34
         5.1.3.1. Introduction to the report 35
         5.1.3.2. Introduction to the Expert Meeting 35
         5.1.3.3. Summary of the discussed topics 35
            5.1.3.3.1. Research in alternative methods done in an "inefficient" way? 35
            5.1.3.3.2. Better use of appropriate biomarkers 37
            5.1.3.3.3. Investigation on new models: 3-D cultures, stem cells, embryotoxicology and teratogenicity 38
            5.1.3.3.4. Validated, regulatory-accepted, or scientific-validated tests: which should be used? 38
            5.1.3.3.5. Drug preclinical research: animal experiments still needed 39
            5.1.3.3.6. Clinical pharmacology and trials with humans 41
            5.1.3.3.7. Further considerations 43
   5.2. Expert Meeting 2 47
      5.2.1. Executive Summary 47
      5.2.2. Recommendations 49
      5.2.3. Report of Expert Meeting 2 51
         5.2.3.1. Introduction to the report 53
         5.2.3.2. Introduction to the Expert Meeting 53
            5.2.3.2.1. Background information on animal disease models, by Peter Maier, Forschung 3R, CH 53
            5.2.3.2.2. Background information on ecopa and START-UP project, by Vera Rogiers, *ecopa* 54
            5.2.3.2.3. Background information on Europe and 3R alternatives: programs, initiatives, projects and networks, by Bernward Garthoff, *ecopa* 54

5.2.3.2.4. Workshop Instructions, by Peter Maier, Forschung 3R, CH 55
5.2.3.3. Summary of the presentations of the individual working groups 55
5.2.3.3.1. Brain: Psychiatric diseases, by Christopher Pryce, Novartis, CH 55
5.2.3.3.2. Brain: degenerative diseases, Christoph Wiessner, Merck-Serono, CH 57
5.2.3.3.3. Inflammatory diseases, Päivi Alajuuma, Santen, FI & Janet Dawson, Novartis, CH 59
5.2.3.3.4. Oncology by Robert Cozens, Novartis, CH 61
5.2.3.3.5. Infectious disease, by Barry Wright, AstraZeneca, USA 63
5.2.3.3.6. Respiratory diseases by Alexandre Trifilieff, Novartis, CH 65
5.2.3.3.7. Metabolism, by Robert McArthur, McArthur and associates, CH 67
5.2.3.3.8. General plenum discussion, cross sectional issues: 69
5.3. Expert Meeting 3 73
5.3.1. Executive Summary 73
5.3.2. Recommendations 75
5.3.3. Report of Expert Meeting 3 76
5.3.3.1. Introduction to the report 79
5.3.3.2. Introduction to the Expert Meeting 79
5.3.3.3. Summary of the presentations within the different sessions 79
5.3.3.3.1. Session I: New approaches in regard to research and safety testing of bio-pharmaceuticals respectively cosmetics - Pharmaceuticals 79
5.3.3.3.2. Session II: New approaches in regard to research and safety testing of bio-pharmaceuticals respectively cosmetics - Dermato-Cosmetics 82
5.3.3.3.3. Session III: New approaches in regard to research and safety testing of bio-pharmaceuticals respectively cosmetics - Nanoparticles 84
5.3.3.3.4. Session IV: Immunotoxicology of Pharmaceuticals / Dermato-cosmetics 86
5.3.3.3.5. Session V: New approaches in regard to research and safety testing of bio-pharmaceuticals resp. cosmetics - Biotech products 91

5.3.3.3.6. Session VI: Research / Development / 3R-alternatives 94
5.3.3.3.7. Session VII: Research / Efficacy testing / 3R-Alternatives 96
5.3.3.3.8. Session VIII: Miscellaneous 98
5.3.3.4. Discussion 101
5.4. Workshop Refinement 107
5.4.1. Executive Summary 107
5.4.2. Recommendations 109
5.4.3. Report of the Workshop Refinement 111
5.4.3.1. Introduction to the workshop 113
5.4.3.1.1. Introduction lectures 113
5.4.3.2. Round Tables 115
5.4.3.2.1. Round Table 1: Housing and legal aspects 115
5.4.3.2.2. Round Table 2: Recent methodologies in refinement 119
5.4.3.2.3. Round Table 3: Refinement in drug development process 124
5.4.3.2.4. Round Table 4: Animal models in human diseases 128
5.4.3.3. Final Discussion 132
5.5. Workshop Reduction 137
5.5.1. Executive Summary 137
5.5.2. Recommendations 139
5.5.3. Report of the Workshop Reduction 140
5.5.3.1. Introduction to the workshop 142
5.5.3.1.1. Introduction lectures 142
5.5.3.2. Summary of the presentations within the different sessions 142
5.5.3.2.1. Session I: Bottlenecks in the Development of NCEs and NBEs (new chemical and biological entities) 142
5.5.3.2.2. Session II: Reduction Potential by Handling Statistics and Experimental Data 153
5.5.3.2.3. Session III: Reduction via Study Design and Methods in Drug Research and Development 160
5.5.3.2.4. Session IV: Transgenic Animals in Pharmaceutical Research and Drug Development 168
5.5.3.2.5. Plenary discussion 175
5.6. Workshop Replacement 181
5.6.1. Executive Summary 181
5.6.2. Recommendations 183
5.6.3. Report of the Workshop Replacement 184
5.6.3.1. Introduction to the workshop 186

5.6.3.2. Summary of the presentations within the different sessions 186
5.6.3.2.1. Session I: Replacement in the development of new chemical entities 186
5.6.3.2.2. Session II: Replacement in the development of new chemical entities 188
5.6.3.2.3. Free communications 208
5.6.3.2.4. Conclusions of the day, by Gisbert Sponer 210
5.6.3.2.5. Session III: Replacement in the development of new biological entities 210
5.6.3.2.6. Session IV: Replacement in quality control 219
5.6.3.2.7. Free communications 224
5.6.3.2.8. Round Table Discussion 227
5.6.3.3. Conclusions of the workshop, by Gisbert Sponer 229

1. ABBREVIATIONS

This is an alphabetical list of abbreviations used in the reports of the Expert Meetings and Workshops of the START-UP project.

ADME: Absorption, Distribution, Metabolism and Excretion
AED: Anti-Epileptic Drug
AhR: Aryl Hydrocarbon Receptor
ALS: Amyotrophic Lateral Sclerosis
APA: American Psychological Association
AST-2000: Advanced Skin Test
ATP: adenosine triphosphate
AUC: Area Under the Curve

BC: Benzalkonium Chloride
bIVF: Bovine In Vitro Fertilization
bIVM: Bovine In Vitro Maturation
BSP: Biological Standardisation Programme

CAF: Cancer-Associated Fibroblast
CAR: Constitutive Androstane Receptor
CASA: Computer Assisted Semen Analysis
CDER: Center for Drug Evaluation and Research
CHMP: Committee for Medicinal Products for Human Use
CHO: Chinese Hamster Ovary
CLSM: Confocal Laser Scanning Microscopy
CNS: Central Nervous System
COPD: Chronic Obstructive Pulmonary Disease
CPMP: Committee for Proprietary Medicinal Products
CRO: Contract Research Organisation
CSS: Clinical Candidate Selection
CT: Computer Tomography
CTA: Cell Transformation Assay
CYP: Cytochrome P450

DG RTD: Directorate-General for Research
DMPK: Drug Metabolism and Pharmacokinetics
DMSO: Dimethylsulfoxide
DNFB: 2,4-dinitro-1-fluorobenzene

ECVAM: European Centre for Validation of Alternative Methods

EDQM: European Directorate for the Quality of Medicines & HealthCare
EEG: Electro Encephalography
EFPIA: European Federation of Pharmaceutical Industries and Associations
EFSA: European Food Safety Authority
EGFR: Epidermal Growth Factor Receptor
EGP: Endogenous Glucose Production
ELISA: Enzyme Linked ImmunoSorbent Assay
EMEA: European Medicines Agency
EMMA: European Mouse Mutant Archive
EPAA: European Partnership for Alternative Approaches to Animal Testing
ERK: Extracellular Signal-Regulated Kinases
ESC: Embryonic Stem Cells
eSI: *ecopa* Science Initiative
EST: Embryonic Stem cell Test
EST-1000: Epidermal Skin Test
Estiv: European Society of Toxicology In Vitro
EUCOMM: European Conditional Mouse Mutagenesis
Eudralex: EU legislation in the pharmaceutical sector
EUPRIM-NET: European Primate Network

FACS: Fluorescence Activated Cell Sorting
FDA: Food and Drug Administration
fMRI: functional Magnetic Resonance Imaging

GABA: Gamma-AminoButyric Acid
GLP: Good Laboratory Practice
GOP: General Operating Procedure
GST: glutathione s-transferase

HCA: HeteroCyclic Amine
hERG: Human Ether-a-go-go Related Gene
HIV: Human Immunodeficiency Virus
HLC: Henry's Law Constant

IARC: International Agency for Research on Cancer
ICCVAM: International Coordinating Committee on the Validation of Alternative Methods
ICH: International Conference of Harmonzation
ILSI: International Life Sciences Institute
IMI: Innovative Medicines Initiative
IND: Investigational New Drug
IPS: Infused Pluripotent Stem cells
IVF: *In Vitro* Fertilisation
IVM: *In Vitro* Maturation
IVTIP: In Vitro Testing Industrial Platform

JRC: Joint Research Centre

KOMP: Knock-Out Mouse Project

LCM: Laser Capture Microdissection
LCSA: Loose-fit Coculture-based Sensitisation Assay
LDH: Lactate DeHydrogenase
LLNA: Local Lymph Node Assay
LNA: Locked Nucleic Acid
LOEL: Lowest Observed Effect Level
LPS: lipopoly saccharide
LTT: Lymphocyte Transformation Test

mAb: Monoclonal Antibody
MABEL: Minimum Anticipated Biological Effect Level
MAP: Mitogen Activated Protein
MAPK: Mitogen Activated Protein Kinase
MAPREC: Mutant Analysis by Polymerase chain reaction and Restriction-Enzyme Cleavage
MCR: Metabolic Clearance Rate
MDL: Molecular Design Limited
mEH: microsomal Epoxide Hydrase
MHC: Major Histocompatibility Complex
MHRA: Medicines and Healthcare products Regulatory Agency
microCT: X-ray microtomography
miRNA: micro RNA
MnSOD: Manganese SuperOxide Dismutase
MRI: Magnetic Resonance Imaging
MTD: Maximum Tolerant Dose
MTS: Soluble Tetrazolium Salt
MTT: 3-(4,5-dimethylthiazol-2-yl)-2,5-diphemiltetrazoliumbromide

NAC: N-acetylcysteine
NBE: New Biological Entity
NCE: New Chemical Entity
NCI: National Cancer Institute
NCP: National Consensus Platform
NGAL: Neutrophil Gelatinase-Associated Lipocalin
NMR: Nuclear Magnetic Resonance
NOAEL: No Observable Adverse Effect Level
NorCOMM: North American Conditional Mouse Mutagenesis project
NRU: Neutral Red Uptake, as in 3T3 NRU

OCABR: Official Control Authority Batch Release
OECD: Organisation for Economic Co-operation and Development

OMCL: Official Medicines Control Laboratory

PAH: Polycyclic Aromatic Hydrocarbons
PAMP: Pathogen-Associated Molecular Patterns
PbAE: Poly β Amino Ester
PBPK: Physiologically Based Pharmacokinetic
PCB: PolyChlorinated Biphenyls
PCDD: PolyChlorinated DibenzoDioxins or pentadioxin
PCDF: PolyChlorinated DibenzoFurans
PCL: Precision Cut Lung slice
PCR: Polymerase Chain Reaction
PD: Pharmacodynamic
PET: Positron Emission Tomography
Ph. Eur.: European Pharmacopoeia
PK: Pharmacokinetics
PK/PD: Pharmacokinetic/Pharmacodynamic
PK/TK: Pharmacokinetic/Toxicokinetic
POP: Persistent Organic Pollutants
PPAR: Peroxisome Proliferator Activated Receptor
PPK: Population Pharmacokinetics
PRT: Positive Reinforcement Training
PSTC: Predictive Safety Testing Consortium
PT: Photoxicity Test
PXR: Pregnane X Receptor

qPCR: Quantitative Polymerase Chain Reaction
QRA: Quantitative Risk Assessment
QSAR: Quantitative Structure Activity Relationship

RDT: Repeated Dose Toxicity
REACH: Registration, Evaluation and Authorisation of Chemicals
RHE: Reconstructed Human Epidermis
RNAi: RNA interference
ROC: Receiver Operation Characteristics

SACUC: Solvay Animal Care and Use Committee
SCC: Squamous Cell Carcinoma
SCCP: Scientific Committee on Consumer Products
SCID: Severe Combined Immune Deficiency syndrome
SDS: Sodium Dodecyl Sulphate
SHE: Syrian Hamster Embryo Cell Transformation
siRNA: small interfering RNA
SME: Small and Medium Enterprises
SPECT: Single Positron Emission Computed Tomography
SPF: Specific Pathogen Free

SPME: Solid Phase MicroExtraction
STREP: Strategic Targeted Research Project

TCDF: tetrafuran
TEQ: Toxic Equivalency
TK: toxicokinetic
TLR: Toll-Like Receptor
ToBI: Toxin Binding Inhibition

UDPGT: uridyldiphosphoglucuronyl transferase
UTR: UnTranslated Region

VICH: International Cooperation on Harmonisation of Technical Requirements for Registration of Veterinary Medicinal Products

WEC: Whole Embryo Culture
WBGT: Whole Body Glucose Test
WHO: World Health Organisation

2. INTRODUCTION

START-UP (Scientific and Technological issues in 3Rs Alternatives Research in The process of drug development and Union Politics) is a support action (n° 201187) within FP7-HEALTH-2007-1.3-2 about the "Bottlenecks in reduction, refinement and replacement of animal testing in pharmaceutical discovery and development". It is coordinated by *ecopa* (European Consensus Platform on 3R Alternatives to Animal Experimentation) with the VUB (Vrije Universiteit Brussel) as the second partner, being responsible for the scientific/administrative secretariat of the project. In particular, it is the Department of Toxicology providing the necessary support for *ecopa*.

The development of 3R-alternatives and their implementation in the safety assessment of the different product types present on the EU market is one of the major objectives of *ecopa*. This international not-for-profit organisation officially exists since December 2002 and has coordinated before the FP6 CONAM (Consensus Networking on Alternative Methods within Europe) project. *ecopa* exists through so-called NCPs (National Consensus Platforms), which are in fact scientific organisations all over Europe, consisting of representatives of the four major parties involved in the use of alternative methods versus experimental animals. These include animal welfare, industry, academy and regulatory bodies. Actually, 16 NCPs exist (14 full members, 2 associate members) being Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Hungary, Italy, Norway, Spain, Sweden, Switzerland, The Netherlands, Ireland and Poland.

Several of these NCPs have worked in this project intensively and constructively together, not only with eminent pharmaceutical experts in the meetings and workshops, but also with representatives of *EPAA* (European Partnership for Alternative Aproaches to Animal Testing), OECD (Organisation for Economic Co-operation and Development), the European Pharmacopoeia and ECVAM (European Centre for Validation of Alternative Methods), with young scientists from all over Europe, with national and EU regulatory bodies, Commission representatives and many others.

3. CONCEPT AND PROJECT OBJECTIVES

The START-UP project must be seen in the context of the actual situation in the EU (and other parts of the world) with respect to the use of 3R-alternative methods to refine, reduce and replace experimental animals in the development process of several product types and their use as "safety and efficacy guarantee" for human health. In this project focus therefore does not lie on alternative methods as such, but on their potential application in the Pharmaceutical Industry in order to improve the final outcome.

Indeed, the process of lead identification, lead optimisation and bringing a drug candidate to the stage of a pharmaceutical finally being approved for clinical use is a time-consuming process. Identifying "wrong" drug candidates therefore at an early stage during the drug development process and avoiding efforts in optimising under-performing candidates in terms of safety and efficacy are essential for the competitiveness of the European industry. Bringing in alternative methods at the right moment and at the right place could therefore be of tremendous benefit not only in terms of animal numbers, but also in terms of a more successful drug development outcome. Only a limited number of 3R-alternative methods have been officially validated and can be used for regulatory purposes. However, a much greater number of alternative methods exists today that can be applied successfully in basic research, in mechanistic studies, in pharmaco-toxicological studies, etc, all of importance for the Pharmaceutical Industry and at the same time for the experimental animals involved.

Basically, no restriction exists in this field as long as the 3R-methodologies used are scientifically sound and relevant and have elucidating and discriminative power at a particular stage of the drug development process. Therefore, a project such as START-UP was necessary, namely a coordinated initiative covering as much as possible all parties involved, being the scientific world, the Pharmaceutical Industry and the different stakeholders in order to achieve a major collaborative activity to get a good and realistic overview of the current use of experimental animals in the whole drug development process and to assess the possibility to implement new alternative strategies and tiered approaches in the different stages of the overall drug development process. The challenge consisted of identifying existing gaps, scientific and technological bottlenecks, ethical concerns and issues related to union politics. This exercise has been carried out successfully. This study not only provides information on the classical drugs, the so-called new chemical entities (NCEs), but also on the new biological entities (NBEs). The latter generation of biological drugs (antibodies, proteins…), nanotechnology and nanobiotechnology molecules is a growing field and creates new challenges as safety clearance of all these new types of substances seems to be more complex and sophisticated than of the classical chemical substances. Consequently in a number of cases using animals for hazard determination is not even relevant. On the contrary, up-to-date fingerprint techniques may offer possibilities to better target the problems and mechanisms involved, so that only relevant molecules on a limited number of animals of the relevant species need some testing in order to guarantee safety and efficacy.

The direct objectives of this project consisted of:

- gathering all relevant information, mentioned above, by organising two Expert Meetings with pharma- and biotech-experts and 3R-specialists (note that an additional third Expert Meeting was organised)
- prioritisation of this information within the three domains of Refinement, Reduction and Replacement
- organisation of three high-level Workshops, one on each of the 3Rs
- developing a Consensus Report between all parties involved on the outcome of the Expert Meetings and Workshops
- proposing Road Maps for the Commission

All these objectives have been met within the given limited timeframe of 2 years.

4. OVERALL RESULTS

The results are presented as an Overall Executive Summary, followed by the Recommendations and Road Maps showing the way forward and full length detailed reports of the 3 Expert Meetings and 3 Workshops. Also, the detailed report will be printed and this booklet will be largely distributed among the different stakeholders and within the large *ecopa* network as *ecopa* Expert Report n°4. An article is in preparation to be taken up in an international journal interested in 3R-alternatives and drug development.

4.1. Overall Executive Summary

The basis of the START-UP project was the general intention to cover all the issues of 3Rs-Bottlenecks in pharmaceutical Research and Development, as represented by the abbreviation, i.e. Scientific and Technological issues in 3Rs Alternatives Research in The process of drug development and Union Politics.
In order to have as much coverage as possible, the area was intensively analysed in expert meetings, predominantly of industry, but also of academia, and regulatory authorities. Later on, in the 2nd year of the project, these closed expert meetings were extended to three open workshops on each of the 3Rs.

All in all, out of a total of 223 participants, there were 109 industrial experts (out of 42 companies) representing pharmaceutical industry or associated institutions, thereby reflecting in particular all aspects of "pharmaceutical life" in R and D.
Started at a kick off meeting in Leverkusen/Germany at Bayer AG, the expert meetings in Madrid (Ministeria de Sanidad), Basle (Novartis Research Center) and Alicante (Pueblo Acantilado, at the biannual eSI meeting) were then followed by Workshops on Refinement (Istituto di Sanita, Rome, I), on Reduction (University of Innsbruck, A) and on Replacement (Budapest, H). These were collaborations of *ecopa*'s National Consensus Platforms (NCPs) of Italy, Finland and Poland, respectively, Austria and the Netherlands, respectively, Hungary and Germany. Academia was represented by 65 participants, regulatory authorities by 29, and animal welfare by 10. Also locally interested scientists actively participated.

The results were presented and found entrance into the discussions; the format varied, intentionally, from brain storming sessions and working group style to formal scientific presentation workshops or plenum style forums to enable free and interactive communication. All presentations and discussions are detailed in report form and are accompanied by an executive summary and a list of specific recommendations. The

major outcome is present here as an overall executive summary, followed by the most prominent recommendations and a road map.

Some of the topics discussed, might be subject to future projects within coming EU Framework Programmes.

- Collation of 3R-topics in pharmaceutical research

- Animal experiments are still needed and realistic progress is actually expected by intelligent combination of refinement, reduction and replacement methodologies / strategies. This is in particular relevant in animal disease models. *In vivo* and *in vitro* research and testing should go together and not be seen as two opposites.
- It was emphasised that an alternative method not necessarily needs to be formally validated, the fact that a test works is for the pharmaceutical industry of more importance.
- Data obtained from *in vitro* tests, carried out before *in vivo* experiments start, can efficiently filter compounds of interest. These pre-tests should be of a higher degree of sophistication and complexity than is the case now e.g. use of 3D-cultures, co-cultures, stem-cell derived models, organ-specific and differentiated cell cultures; more human cells use and more attention for the parameters measured e.g. it is unlikely that only one biomarker will cover the complexity of the living organism, therefore a set of specific biomarkers of clinical relevance increases the translational nature of the *in vitro* model used; these should be developed at least for key organs and new and potent tools should be involved (e.g. transcriptomics, metabonomics, biostatistics).
- When animals are involved, they should be of a relevant species for the question posed, otherwise experimentation should be deleted. The same is true for exposure to unrealistic high dosages/exposure scenarios.
- Important fields for further development are teratogenicity and embryotoxicity as these tests are necessary for every newly developed drug coming on the market; for exploration of new opportunities pharmacodynamics, and for better integration into single test programmes for pharmacokinetics, carcinogenesis, safety pharmacology and toxicology.
- In test development more focus should be on “risk assessment” than on “hazard assessment”.

- Concepts of cell system improvements

- These were high on the agenda. Stabilisation (e.g. by epigenetic modifications, miRNA interaction) of existing cell systems, and to use these for long-term testing has potential toxicity and efficacy testing. In addition, the fact that the heterogeneity of human population is not taken up by current *in vitro* tests deserves efforts to develop models capable of mimicking human variability.

- Concepts of data sharing and reporting of “negative” results

- These aspects are important in gaining more basic information and reducing replication experiments. They are of special importance in certain diseases.

- Essential for sharing data are data quality control, protocol standardisation and in particular protection of intellectual property. It was proposed to overcome this hurdle by establishing a “neutral” pan-European party entity.

- Aspects of lab animal husbandry, of best practice for lab animal keeping

- Emphasis was given to positive aspects such as better training of personnel and in particular of competent authorities; positive welfare of experimental animals e.g. via group housing, creation of possibilities for natural behaviour, environmental enrichment, consideration of positive reinforcement training in the case higher animals are involved.
- Proposals for central breeding of controlled and certified quality were particularly brought forward for primates and transgenic animals.
- Emphasis was also given to the importance of the microbiological quality of the animals, leading to better experiments and indirectly leading to less animal use.

- Furthering of model development, especially of non-invasive in vivo methodology

- This point came up in all meetings and workshops and supports the further transfer of non-invasive diagnostic methodologies (e.g. magnetic resonance imaging, micro CT) from human medicine to laboratory animals allowing not only diagnosis but also long-term monitoring of treatment. In particular, the combination of different non-invasive imaging techniques was seen as a possibility for refinement and reduction and at the same time for gaining better knowledge.
- In particular, in animal disease models this methodology is seen as a key improvement.

- Bottlenecks in biologics development

- Use of humanised models, knock-out animals and transgenic animals could help to make more appropriate use of animals as high target specificity is involved. Also transgenic cells/enzymes/*in vitro* models have relevance.
- More parameters should be combined in one animal study (e.g. safety pharmacology, pharmacokinetics, local toxicity, immunogenicity).
- Standardisation of animal strains, microbiological high quality of animals, use of well-defined environmental conditions and techniques are crucial reduction parameters in this field.

- Special case of vaccines quality control

- As in the EU, authorities request that all vaccines lately must be tested; high numbers of animals are consumed. Moving from this traditional quality control concept towards the monitoring of all crucial steps during production could save these animals. This so-called consistency approach was largely supported.
- In vaccines quality control, refinement strategies should be developed and implemented

- Implementation of existing 3R-methods should be encouraged by improving and global harmonising of the regulatory procedures. Also providing incentives to development and production is considered to be important.
- More attention should go to the neglected area of veterinary vaccines.

- Specific disease animal models with high burden to experimental animals.

- As animal pain models are not very predictive, well-controlled studies in man using micro dosing were proposed in order to be able to score pain in a realistic way.
- Cancer models are a special target for further improvement, also since by the development of biologics for this topic, the area is more covered. In oncology, genetically engineered models and primary tumour models were said to be productive. A refinement alternative could be the study of surrogate tissues from normal animals which usually exhibit the fully functioning pathways that are targeted. Also the importance of measuring *in vitro* specific biomarkers that can also be detected in the clinical situation came up.

- Analysis of Union politics, country / Member States politics

- Over expectations with respect to alternative methods should be avoided.
- Ethical issues and political restrictions were discussed with respect to human stem cell use. Heterogeneous opinions within the different Member States should be better harmonised.
- Member States should establish National Animal Welfare and Ethics Committees with well-trained personnel to give advice to the competent authorities and permanent ethical review bodies of establishments. Networking of these committees should play a role in the exchange and communication of best practices.
- Importance was given to a trans-sector, cross sector-cutting information stream by regulators and industrial partners.

- Refined analysis of general EU research strategies

- The general research strategies applied today at the EU level are a burden to potential applicants and the administration of EU Framework Programmes are seen as a hindrance to appropriate research in alternative methods. Less bureaucracy, better integration of research teams, eventual leadership of pharmaceutical industry, limitation of number of projects per team and need for new names of young scientists and a fresh outlook were all mentioned as possible improvements.

- Global harmonisation

- The importance of global harmonisation as the basis for further implementation of alternative methods came up in all meetings and workshops.
- A unified animal legislation and, in this context, specific actions addressed towards the political world were seen as important.

- Communication on new models across sectors, involving regulatory agencies and competent authorities should be enhanced.
- Dissemination and promotion of refinement/reduction techniques in drug development was seen as an important step forward.
- Global harmonisation is highly important and should be pursued even it is difficult and slow.
  Worldwide harmonisation should be brought in the execution of pharmaceutical registration and general concepts, also existing Animal Welfare in the different Member States should be better harmonised and the revised Directive 86/609 could help in this process.

In summary, the EU 7th RTD Framework Programme project START-UP has delivered a whole landscape of ideas and potential avenues for further research and development projects within the future EU Framework Programme in regard to 3Rs bottlenecks and EU industry competitiveness; these should be considered when drawing up new project calls in this area in the future. It has been demonstrated that only detailed discussions with experienced experts can lay groundwork for adequate analysis.
Also, these approaches, as laid out in more detail in the individual recommendations, have to be discussed with the experts involved, the Scientific Officers of the EU Commission, the European Parliament representatives as well as the Industry and the interested public. With the pool of experts brought together under START-UP, the furthering of the Road Maps attached, can be achieved.
Further workshops organised by the project partners involved, should spread the message, in order to come up with solutions for some bottlenecks where solutions are not easy to come by.

The Project Coordinator
and the Organisers Team

4.2. Overall Recommendations

In the detailed report, a list of specific recommendations is given for each Expert Meeting and the 3R-Workshops, making a total of 36 recommendations for further follow-up by the parties concerned and in particular by the Commission.
Here the eight most important recommendations are summarised.

1. Reduction and refinement are particularly possible in the field of animal disease models. It is recommended to maximize the number of non-invasive and early or surrogate endpoints within one model. Progress in non-invasive test development is seen in the further development of non-invasive imaging / diagnostic techniques transferred from human medicine to laboratory animals, and their intelligent combination.

2. Efforts should be focused on the development of batteries of sensitive and specific safety biomarkers with clinical relevance to be measured during the preclinical *in vitro* testing phase.

3. The difference in bottlenecks during the development of biopharmaceuticals versus small molecules pharmaceuticals should be better recognised and dealt with. In particular, the relevance of the animal model came up in the case of biopharmacuticals. The use of non-human primates (in a number of indicated cases), humanised models and transgenic animals seems relevant.

4. A lot of animals could be spared without loss of quality in the quality control of vaccines in Europe. Therefore it is highly recommended to study the possibility for drastic change
   - by a better control of the implementation of already existing refinement and reduction alternatives by all producers and regulatory bodies.
   - by providing the necessary incentives to apply these alternatives
   - by stimulating the development of new alternatives in this field
   - by applying the so-called Consistency Approach confirming production consistency.
   - by paying special attention to veterinary products

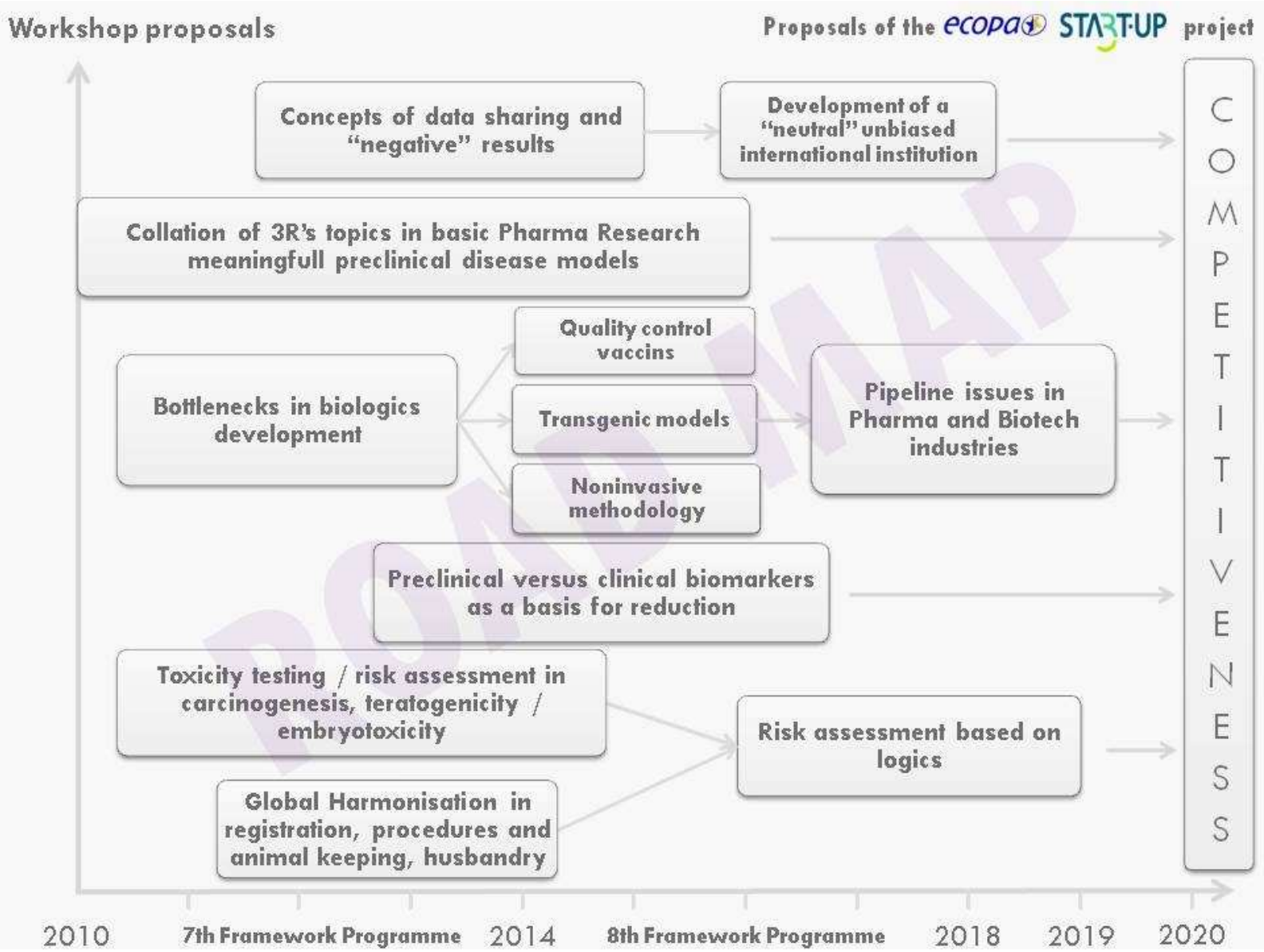
5. It is recommended to develop the possibilities of “data sharing” by creating the necessary working tool, namely the establishment of a “neutral” non biased body that could guarantee confidentiality and as such could take away the fear of losing competitiveness. In this way, also quality control of data and standards of protocols could be assured.
   Furthermore, it was felt that also the follow-up of “negative” results of high standard could contribute to the reduction process.

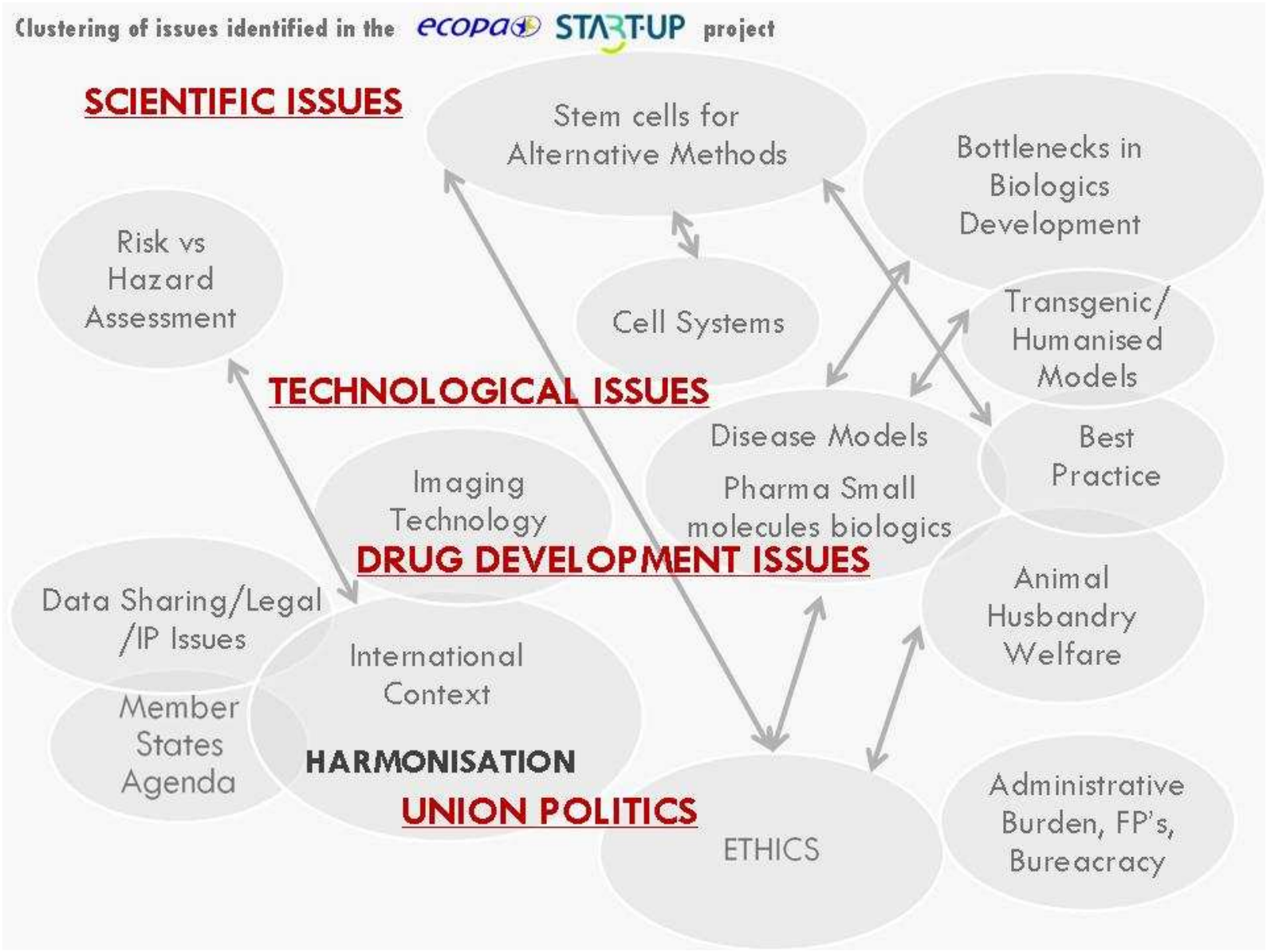
6. Animal reduction in drug development is possible by reducing the number of potential interesting molecules that undergo *in vivo* testing by better pre-screening for unwanted effects and deceiving efficacy.
   Therefore, more sophisticated *in vitro* models based on human cells and tissues should be developed and applied in pre-screening: 3D-models, co-cultures, epigenetically stabilised cell lines, stem-cell derived specific cell types,…

7. Promotion of positive welfare of experimental animals, besides minimalisation of suffering, is seen as a refinement priority and should include active improvement of the degree of animal welfare in- and outside experimental procedures, backed up by ethological studies on laboratory animals.

8. Global harmonisation is seen as one of the highest priorities for further success in the implementation of the 3Rs. It is thought that all different players internationally involved in drug development, human health, alternative methods development and animal welfare should be brought together to agree on the different procedures to be followed in registration toxicity and efficacy testing, and risk assessment, in the development of biologics and quality control of vaccines, and in the different stages of animal use during drug development, in particular in the case of animal disease models.

4.3. Road Maps


Workshop proposals
Proposals of the ecopa START-UP project
Concepts of data sharing and "negative" results
Development of a "neutral" unbiased international institution
Collation of 3R's topics in basic Pharma Research meaningfull preclinical disease models
Quality control vaccins
Bottlenecks in biologics development
Transgenic models
Pipeline issues in Pharma and Biotech industries
Noninvasive methodology
Preclinical versus clinical biomarkers as a basis for reduction
Toxicity testing / risk assessment in carcinogenesis, teratogenicity / embryotoxicity
Risk assessment based on logics
Global Harmonisation in registration, procedures and animal keeping, husbandry
COMPETITIVENESS
2010
7th Framework Programme
2014
8th Framework Programme
2018
2019
2020


Clustering of issues identified in the ecopa START-UP project
SCIENTIFIC ISSUES
Stem cells for Alternative Methods
Bottlenecks in Biologics Development
Risk vs Hazard Assessment
Cell Systems
Transgenic/ Humanised Models
TECHNOLOGICAL ISSUES
Disease Models
Best Practice
Imaging Technology
Pharma Small molecules biologics
DRUG DEVELOPMENT ISSUES
Animal Husbandry Welfare
Data Sharing/Legal /IP Issues
International Context
Member States Agenda
HARMONISATION
UNION POLITICS
ETHICS
Administrative Burden, FP's, Bureacracy

5. DETAILED RESULTS

In this section the results of the three Expert Meetings and three Workshop are presented. Each Expert Meeting and Workshop is dealt with separately. Each part exists of an executive summary and recommendations, as well as the full report, with agenda and persons composing the organising committee.

Expert Meeting 1

19 May 2008, Ministry of Health, Madrid, ES

5.1. Expert Meeting 1

5.1.1. Executive Summary

The first expert meeting was intended to cover in particular regulatory, political and industrial bottlenecks in the use and implementation of 3Rs methodologies and strategies with special emphasis on the pharma sector. A number of important issues were discussed in depth and are summarized here.

- Research on alternative methods done in an "inefficient" way:
Research of new alternative methods and its application to drug safety is done in a very inefficient way, despite the considerable amount of resources allocated by the EC. Need for breakthrough new ideas, better efficacy (i.e. less bureaucracy, better integration of research teams, active involvement, eventually leadership of pharmaceutical industry) and prioritisation in research and limitation on the number of projects a team can participate in, were identified as constraints to improve research outcome.

- Better use of appropriate biomarkers:
Integrated biomarkers are a need. Using only one biomarker is unlikely to cover the complexity of living organisms. Therefore, a reduced set of specific biomarkers, relevant for the clinical situation as well, should be developed for *in vitro* purpose.

- Need of investigation on new models:
More research on 3D-models as well stem cell research and differentiation aspects is needed, but too high expectations should be avoided. Apart from ethical and political restrictions to the use of human stem cells in research, this technology is not seen as having reached the maturity as to be really available for routine purposes. More suitable (as existing) models for teratogenicity and embryotoxicology, are needed and efforts for this purpose are encouraged.

- Validated, regulatory-accepted, or scientific-validated tests:
In industry, the fact that a test works is more important than whether it is formally validated or not. It will not be always possible to reach a full and "formal" validation for certain tests, despite its usefulness. There should be more focus on "risk assessment" than on "hazard assessment".

- Animal experiments still needed in drug preclinical research:
Experts acknowledged that animal experiments are still needed, despite their often-questioned predictability in preclinical research. Asking the right question to the right animal species is crucial. There is no point in doing animal experiments when the species

is not relevant to humans. Creating "humanised models" could help to make more appropriate use of the animals. With biological products (e.g. MAbs, chimeric proteins) this might be a determining factor.

***- Clinical pharmacology and trials with humans*:**
Clinical use of drugs without animal testing is to-day only possible in exceptional cases (life threatening diseases, no animal models available, adequate risk–benefit relation for the patient). The fact that the heterogeneity of human populations is not taken up by current *in vitro* tests deserves attention and efforts to develop models capable of mimicking human variability.

- Finally:
It was acknowledged that better collaboration and communication between academy and industry, availability of negative results harmonised and unified animal legislation and specific actions addressed towards the political world are needed.

5.1.2. Recommendations

1. Research on alternative methods in the context of drug development should be conducted in a more efficient way. Less bureaucracy, better integration of research teams, active involvement (ev. leadership) of pharmaceutical industry, prioritization of research and limitation on the number of projects a team can participate should be undertaken.

2. There is a need for identifying more specific biomarkers which should be clinically relevant as well procedures to better integrate the information.

3. More research on 3D as well as stem cell differentiation is needed, but too high expectations should be avoided. The use of human stem cells in research is not perceived as having reached the maturity as to be used for routine purposes.

4. The use of scientifically substantiated tests can contribute to reduce the number of animals even if they are not formally validated. There should be a focus preferably on "risk assessment", instead of "hazard assessment".

5. Animal experiments are still needed despite their predictive value in preclinical research is often questioned. Asking the right question to the right animal species is crucial. There's no point in doing animal experiments when they are not relevant to humans.

6. The heterogeneity of human populations is not properly taken up by *in vitro* tests and deserves efforts to develop models capable of mimicking human variability. *In vitro* models that mimic human variability should be developed.

7. Better communication and collaboration between academy and industry, availability of negative results, harmonised and unified animal legislation and specific actions addressed towards the political world should be undertaken

5.1.3. Report of Expert Meeting 1

Start-Up Expert Meeting 1
19 May 2008, Ministry of Health, Madrid, ES

Program:

10:30 – 10:40	Welcome. José Castell, *ecopa*, ES
10:40 – 10:50	Introduction. Vera Rogiers, *ecopa*, BE
10:50 – 11:00	Working Instructions. Bernward Garthoff, *ecopa*, DE
11:00 – 14:00	Brainstorming Sessions
11:00 – 12:00	Session: Preclinical Research
12:00 – 13:00	Session: Preclinical Safety, Kinetics, Metabolism
13:00 – 14:00	Session: Requirements of Clinical Pharmacology and Clinical Trials
14:00 – 15:30	Lunch
15:30 – 16:30	Working Teams: Identifying of first recommendations
16:30 – 17:30	Plenary Session: Passing final recommendations
17:30	End of the Meeting

Scientific committee:
Bernward Garthoff, treasurer *ecopa*, Bayer, DE
Vera Rogiers, chair *ecopa*, VUB, BE
José Castell, vice-chair *ecopa*, Hospital Universitario La Fe, ES

5.1.3.1. Introduction to the report

This first Expert Meeting was conceived as a preparatory meeting for the forthcoming Workshops of which each one should address an individual "R" within the context of drug development, and will constitute the core of the START-UP project. Invited participants were major players in the field from academia, pharmaceutical industry and regulatory bodies and were asked to define existing bottlenecks, provide ideas and suggestions about the overall process of drug development in respect to the 3R's.

5.1.3.2. Introduction to the Expert Meeting

After the welcome by the local organiser José V. Castell (ES, *ecopa* Vice-President), participants were informed by Vera Rogiers (BE, President of *ecopa* and coordinator of START-UP) about the nature and goals of START-UP project, and the way it was structured with a series of specialized meetings with the goal of gathering relevant information concerning 3Rs burning issues in drug development. The ideas and suggestions drawn out from this first meeting were loaded into the forthcoming workshops, each one dealing with one of the 3Rs. By the end of START-UP, the ideas and recommendations are to be forwarded to the Commission with proposals for further research in the field of 3Rs with respect to bottlenecks in safety and efficacy testing of pharmaceuticals, together with Road Maps for its implementation.

Detailed instructions about the development of the meeting were provided by Bernward Garthoff (DE, *ecopa* Treasurer) as well, and emphasis was laid on the fact that the represented expertise of companies, governmental and academic institutions should convene and work together, to identify the bottlenecks and make recommendations to overcome them. Issues of relevance for the institutions/companies present should be accompanied by others of broader implications. The meeting and the foreseen report was said to be semi-confidential, meaning that it was not open to the public in general, but made available to those involved in the START-UP project.

5.1.3.3. Summary of the discussed topics

5.1.3.3.1. Research in alternative methods done in an "inefficient" way?

The first issue addressed by the participants was the perception that research on alternative methods was done in a rather inefficient way. Despite the considerable EC funding there hasn't been too much success in the past twenty years in the field of developing new alternatives. Alone in the FP6 already € 63.3 million were spent on the development of 3R-alternatives. Additional € 30.4 million are planned in FP7. In addition, during the past 10-20 years, millions were spent by governments, companies

and private foundations, but very few methods and tests came out which are widely used in pharmaceutical research and development. What are the reasons for this?

- Concerning the EU-funded research, the increasing bureaucratic complexity needed to apply, develop and report was recognised. It sequesters considerable efforts from the coordinator and partners to fulfil the formalities and to get the funds delivered in due time. Pharmaceutical companies often step out of these projects and refrain themselves to participate because of this bureaucracy. Pharmaceutical industry should be again involved in ongoing research and development of 3R methods.

- Most of the projects are coordinated by academics, and there is a big difference between the ways research is done at academic institutions compared to the pharmaceutical industry. The lack of clearly defined goals that could be realistically completed along the duration of the proposal, together with a lack of flexibility and agility to adopt decision, should the research be reformulated, contribute to this "low efficiency". Pharmaceutical companies often perceive the goals of the projects as very academic and often "unrealistic" in terms of calendar.

- The expectations, according to the proposal memorandum, are often unrealistic in the sense that it is generally expected that a simple model could solve virtually any complex problem. A lot of money has been devoted to fund models that only can provide imprecise answers, at best. This seems to be unwise invested money. It even happens that during the lifetime of a project the expected results have become of general scientific knowledge, but still the project continues and often receives financing for a couple of years more. It is surprising that for the vast majority of projects, their outcome undergoes little if any evaluation to determine whether they have made a real contribution to *in vitro* research, in view of the state-of-the-art of alternative methods.

- There is a certain redundancy in the names of researchers and teams involved in projects. An example was given of a partner being present in 7 projects and not delivering the results promised, but still being financially supported in the different EU projects. Researchers involved in *in vitro* alternatives are well aware of each other. Having a strong *in vitro* research network normally show advantages, but one should be cautious not to fall into a lack of constructive criticism. In this context, it was emphasised the need for new names of young scientists and a fresh outlook. There is a need for a better integration of research teams and an active involvement (eventually leadership?) of pharmaceutical industry.

- There is a disconnection between the different on-going research projects on alternative methods. Previous results of EU-sponsored projects are frequently ignored both by researchers and evaluators, resulting in an unnecessary and

redundant repetition of research. There is a clear need for efficacy in research and prioritisation and funding, based on previous project outcomes is a key issue.

- There is a clear need for breakthrough new ideas. Intelligence and research must be invested. As a last comment it was said that too many of the EU projects were focused on replacement, nearly forgetting refinement and reduction.

5.1.3.3.2. Better use of appropriate biomarkers

The concept and use of biomarkers is not new. Diseases as well as injury mechanisms tend to be multifactorial, with many factors playing a relevant role. While it is easy to look for one factor, and subsequently identify a biomarker, it is nearly impossible to do so for all factors putatively involved in a pathological process. Using only one biomarker is unlikely to cover the complexity of living organisms. However trying to go too complex is not working either, as tests should keep a certain degree of simplicity to be easily applied and at the same time generate data that could be understood and interpretable.

The problem is that nowadays is that a lot of *in vitro* work is done, often using too simple tests and models, and that *in vivo* studies come in too late. There should be again more attention for *in vivo* studies but in an intelligent way, namely identifying relevant biomarkers both *in vitro* and in animals, and using them when scientifically justified and with the aid of sensitive biomarkers also relevant for the clinical situation. Indeed, while attention is paid to the complexity of the phenomena in the animal models, species specific differences between animals and humans sometimes seem to be forgotten in this process.

Thus, there is a need for a reduced set of more specific biomarkers which should be relevant for the clinical situation as well. Actually only few fulfil this criterion. For the kidney for example, it has been possible to identify 5 to 7 biomarkers which represents a relevant step forward with respect to safety. Ideally, it should be also possible for the heart, lungs and liver, and efforts to identify such biomarkers should be addressed with priority, and it is likely to be the case for all other organs where attention hasn't been focused on yet. Appropriate use of biomarkers will help to exclude compounds likely to elicit adverse effects in man.

New and potent tools are now available to identify better and more relevant biomarkers. Following the boom of transcriptomics, and having identified the limitations and the scope of applicability, metabonomics now appear as a powerful strategy to identify parameters intrinsically associated to a pathological process, and that are coincidental *in vitro*, *in vivo* and in clinics. The need for identifying more specific biomarkers will be liked to the use of appropriate procedures (biostatistics) to better integrate the information in order to make it more predictive.

5.1.3.3.3. Investigation on new models: 3-D cultures, stem cells, embryotoxicology and teratogenicity

The relation between *in vitro* and *in vivo* models and the relevance of *in vitro* generated data to the situation *in vivo* was discussed by the experts. The question posed was how representative the results of the former model are for the latter and vice-versa? The example was given of *in vitro* systems lacking cell-cell communication and the 3-dimensional structure of organs. It is recognised that progress has been made in developing culture systems aim to reproduce this complexity, but it was underlined that most of the current alternative methods available today cover a specific endpoint but do not reflect the specific situation *in vivo*. Experts present at the meeting emphasised that mechanisms and processes should first be identified and described *in vivo*, in the whole complex organism, and only then translated to *in vitro* systems, possibly starting with the use of lower organisms as surrogate of the more complex animal models.

Attention was drawn to the fact that research on stem cells is still hindered in some European countries. In Germany and Switzerland there are restrictions concerning the use of human stem cells and the research that can be carried on with them. Interestingly, experts were quite unanimous when recognised that stem cell technology, although very promising, is not yet available for routine purposes. Oversized expectations are created when SMEs advertise their technology as being mature and readily available for screening purposes. More research is still needed to generate differentiated cell models out of stem cells, but should be encouraged.

A significant number of animals used in toxicological research are in the field of long term safety studies. Pharmaceutical companies make most use of animals in long-term toxicity studies; teratogenicity and embryotoxicology are among the heavy consuming animal assays. Therefore, more efforts are needed to develop suitable alternatives because these tests are necessary for every newly developed drug coming on the market, and still no good alternatives are available. There was an EU call in FP6 with respect to repeated-dose toxicology, and no groups applied, showing the problem in this difficult field. More efforts are needed to develop suitable alternative tests for this purpose.

5.1.3.3.4. Validated, regulatory-accepted, or scientific-validated tests: which should be used?

The issue of whether either validated, regulatory-accepted, tests should be used, or scientific-validated tests could be part as well of routine testing was raised and extensively discussed among participants. In industry, the fact that a test works is more important than whether it is formally validated. There is a lot on-going research on drugs being done without animals. These investigations are often mechanistically oriented and not carried out for safety assessment reasons. In this context, alternative methods work quite well in pharmacology research, while they do not work that well in risk assessment,

probably due to the fact that the target is generally better identified and known in pharmacology (known target) than in toxicology (unknown target).

There are quite a number of useful alternative methods developed and being used in pharmaceutical companies, but these are kept within the individual companies. They first are validated within the company and become accepted for the testing of a class of drugs. The usefulness of such methods rely on a large number of compounds tested in house. Researchers are confident on the consistency of the data these models deliver, despite they are not "formally" validated, or accepted at the regulatory level. In industry, the fact that a test works is more important than whether it is "regulatory" validated or not.

"Validation" of an alternative test is no priority for industrial companies, as this is a long-lasting, complex, resource-consuming and expensive process, in particular when the test under consideration is of limited use. The process of formal validation is complex and it was emphasised than in many cases it may be far from being feasible. It will not be always possible to achieve a full and "formal" validation for certain tests, despite its undoubtful utility. Regulatory institutions seem to put more emphasis on the validation of tests than pharmaceutical companies do.

The FDA (Food and Drug Administration) encourages companies to include *in vitro* test data in the dossiers but it is only seen as an "add-on" to the *in vivo* data. Acceptance of well supported *in vitro* data will contribute to the reduction of animal consumption.

5.1.3.3.5. Drug preclinical research: animal experiments still needed

All participants agreed that animal experiments are still needed, despite the fact that the predictive value of animal models in preclinical research is often questioned. Some experts underlined the fact that pharmacologists tend to work using exclusively *in vitro* systems without any animal experimentation until later stages may be inappropriate. The complexity of the animal models is also needed at early research stages, because a number of changes or findings cannot be seen in *in vitro* experiments. Furthermore, information about pharmacokinetics and drug efficacy can only be reliably obtained in animal experimentation. *In vitro*, there are a number of models and endpoints that can serve as supportive evidence: pharmacokinetic models, knock-out models, presence of transporters, CYPs (Cytochrome P450 family), etc. These can support the safety of compounds, but hardly their efficacy.

- The predictive value of using animal models in preclinical research is often questioned. Most of the inconsistencies are due to the use of an inappropriate animal model or biomarker. Frequently, the targets are identified wrongly, because we do not properly understand the pathophysiological mechanisms involved so far. Only then, better *in vitro* models can be developed. Importance should be given to identifying molecular markers, such as a receptor or a gene, which make it easier to acquire a model.

- It is very important to choose the right species and to ask the right questions. There's no point in doing animal experiments when they are not relevant to humans. Therefore much emphasis has to be given to assess the relevance of a given animal model prior to run the experiments. The relevant species should be selected on the basis of predictability for humans.

- When it comes to reducing and to refining, it should be realised that a number of *in vivo* tests used today are considered not appropriate, such as the uterotrophic assay as a general marker for *in vivo* endocrine disrupters. When it comes to refinement, in particular non-invasive methodology is important, but some systems are over-claimed and cannot be used for all the endpoints mentioned.

- More attention should be devoted to the issue of choosing the right target. Most of the targets are multi-tissue related and in different tissues different targets can be present (also silent targets). Today one jumps too fast into one single target.

- *In vivo* and *in vitro* research should go together and not be seen as two opposites.

"Humanised animal models" & transgenics. One of the difficulties with animal models is that compounds do not always have the same effects in the model species as in humans. This could be solved by creating transgenic refined models (i.e. humanised models), more representative of the human situation. It is, however, not known what the effects are on the total physiology of the test animals in general when only limited parts e.g. one organ is humanised. The use of transgenic models also can help through, for example, the use of transgenic enzymes.

Regulatory toxicology. A frequent problem with *in vitro* tests is that they tend to be oversensitive. Examples on this exist: in genotoxicity studies, for instance, *in vitro* tests have been used already for a long time showing that they are very sensitive, but not very specific; *in vitro* tests for phototoxicity give large numbers of false positives (i.e. positive in *in vitro*, while negative in *in vivo* tests); a similar situation is observed for *in vitro* clastogenicity assays. Oversensitivity could perhaps be reduced by using more complex models (i.e. 3D models, like reconstructed human skin that can help to analyse false positives in the case of irritancy or phototoxicity).

In vitro methods should be re-focused to address the issue of "risk assessment", instead of only "hazard assessment".

Drug metabolism and pharmacokinetics. Studies on pharmacokinetics and drug metabolism must be carried out at early stages of drug development, because the data

generated is important to choose the right animal model for further efficacy studies. Knowledge of pharmacokinetics is important before compounds undergo clinical trials.

While there are well consolidated methods to investigate *in vitro* metabolism of drugs that generate highly predictive data of the *in vivo* situation, *in vitro*/*in silico* modelling of drug pharmacokinetics is still in a juvenile phase of development. *In vivo* experiments remain necessary, because, for example volume distribution and transporter studies are not possible by only using computer simulations and/or *in vitro* methods. Consequently, *in vivo* animal experimentation seems in this case unavoidable.

Experts made a series of comments that could help to make a more rational use of animals:

- There is doubt whether a full set of regulatory pharmacokinetic data is needed at an early stage of development. It seems possible to combine pharmacokinetic and good quality toxicokinetics data that can be obtained during routine toxicity testing (i.e. in the course of repeated-dose toxicity studies).

- In many companies, pharmacokinetic studies are not coordinated with toxicology studies. Combining pharmacology and toxicology studies should also lead to reduction in the use of animals. This would require some adaptation of regulatory guidelines, since for toxicology studies high doses are requested for margin safety assessment, which is not the case for pharmacology studies (except for the therapeutic window determination). Next to this, it is thought that studies at unrealistic high doses (in toxicology) often are not of help.

- Miniaturisation of certain common methodologies could possibly also lead to reduction and refinement. For example, investigations using smaller volumes for urine and blood are realistic and feasible possibilities using the nowadays very sensitive analytical methods

- More attention should be given to species selection in pharmacokinetics. For the time being, rat and dog are chosen as testing species because of availability, regulatory requirement, cost and historical data in hands of companies. There is too much of a “tick-box attitude” and not an attitude to determine whether these studies are really helpful or valid.

5.1.3.3.6. Clinical pharmacology and trials with humans

Would it ever be acceptable to go to clinical trials without animal testing? When biotechnology-derived drugs are being developed, frequently there is no suitable animal model. Can those drugs then be accepted by only using *in vitro* methods? The question was discussed in detail, and the experts pointed out several considerations:

- Performing clinical trials without (or with reduced) animal testing is possible in two exceptional cases. First, when one is dealing with life threatening diseases and there are no animal models available (reduced animal testing is accepted in certain anticancer products, vaccines, biologics, mAb, etc.). It is also important to take into account the risk–benefit relation for the patient.

- However, distinction can be made between the use of a given animal model to study efficacy and to study drug safety.

- Second, when a compound is under consideration for which it is not possible to get a suitable animal model. In such cases, the combination of human *in vitro* and animal *in vivo* data for pharmacology and toxicology testing may be acceptable. A report with these data is presented instead of the conventional animal data dossier.

- Another example is mAbs directed to human tumour cells, inducing apoptosis. This is very human specific. If there is a good proof of concept *in vitro* that the mAb only acts on tumour cells and not on normal cells, a good rationale and good communication with authority, then it seems reasonable to perform clinical trials without testing on animals first.

- A very interesting issue raised by experts is the fact that the heterogeneity of human populations is not taken up by current *in vitro* tests. This can even be a problem with "pure-bred" test animal species, which are more homogeneous than natural populations, including humans. *In vitro* models capable of mimicking human variability should be developed.

Toxicity testing of monoclonal antibodies. The Tegenero case. Monoclonal antibodies (mAbs) are the newest therapeutic approach to many diseases (e.g. malignant tumours, AIDS, Hepatitis). With the use of humanised mAbs the possible species differences become very important, and thus there is a big concern about choosing the relevant models. Discussion about the use of appropriate homologous systems is an issue, because of its intrinsic difficulties (sometimes it might be possible to use transgenic models *in vivo)*. Because of the molecular targets addressed by mAbs, it really depends case by case. This is probably the type of compounds where having an integrated *in vitro* / *in vivo* approach might be determinant to assess its safety.

Monoclonal antibodies were once tested on humans with disastrous consequences (the Tegenero case). The six volunteers all got very sick upon administration of the mAb. What did go wrong? Apparently, the possible differences in pharmacology were not taken enough into consideration. The cynomolgus monkeys were pharmacologically responsive, but not very sensitive, something that was noticed from the beginning (pointed out one of the participants). Should this have been taken into consideration, the tested dose in man would have been much lower. It was very risky to test this compound with six persons simultaneously. One wonders why they did not start with one person

with a lower dose, then gradually increasing the dose. Unfortunately, the *in vitro* tests were predictive only after they had been modified to address this problem.

The special case of cosmetics. From March 2009 onwards cosmetic ingredients are not anymore allowed to be tested on animals within the EU. Exceptions are repeated dose toxicity, developmental toxicity and toxicokinetics. From March 2013 onwards all animal testing is banned and products tested, even outside Europe, with the help of animals are completely banned from the EU market. This could result in more actual testing of new ingredients and more testing outside Europe. Obviously both of these possibilities do not help in saving animals.

Companies from the USA and Japan come to Europe to look for alternatives which they can and do use. The difference is that in these countries the use of alternatives is not obligatory, while in Europe it is. They can use those tests that are useful / valid and that have been developed and tested in Europe.

The fact that there will be soon a ban on the testing of cosmetics is, at least partly, a consequence of extensive lobbying by animal welfare organisations often supported by scientists. Most people only think of "make-up products" when cosmetics are mentioned. Nowadays, cosmetics consist of a broad group of products extensively used and having in many cases a preventive action on disease (for example by applying UV-filters to prevent skin cancer by sun exposure). This makes clear that safety control is important in cosmetics and science should keep being involved with the development of alternatives. However, lobbying for using a methodology which is not yet fully scientifically available is dangerous and bad practice. Pharmaceuticals should be spared of this type of lobbying.

5.1.3.3.7. Further considerations

Several ideas were finally posed by experts:

Improved animal act regulations. There should be more concern on the burden of pain that animals suffer from during tests, than to the killing itself. Now we see too many bits and pieces, which make it easy to go to another country when certain tests or experiments are not allowed in one's home country. Experts indicated that there should be a clear basic regulation (laws) for the whole of Europe based on the following:

- European Convention for the Protection of Vertebrate Animals Used for Experimental and Other Scientific Purposes (adopted 1991), including ETS No 170 (adopted 2005)
- Appendix A: accommodation and care of animals (adopted 2007)
- Appendix B: Tables for the presentation of statistical data on the use of animals for scientific purposes
- Directive 86/609/EEC which is under revision.

Ideal world. Harmonisation and standardisation of methods would make inter-laboratory comparisons of tests and procedures easier. However, this is difficult to achieve in industry, as some people are reasoning that this will have a negative impact on innovation. Communication and collaboration between academy and industry have to be improved. Academia is nowadays also interested in protecting knowledge as industry is, and this should facilitate collaboration.

Concerning acute (single-dose) toxicology, the OECD (Organisation for Economic Co-operation and Development) states that a number of animals could be saved. But this doesn't seem to be important for pharmaceuticals anymore, but has still a considerable impact on other industries, i.e. for Chemicals, Agrochemicals, Pesticides and Food industry.
In practical terms, the introduction of alternatives in regulatory toxicology can only be reached nowadays through the ICH (International Conference of Harmonisation) process. The latest progress to this respect is remarkable (e.g. ICH Guidelines draft S2, M3 …).

Meaningful political implication.

- Politicians should be taught about the real use of alternatives with their limitations, as important as to teach the public and the patients. Contacting high-level European politicians (Commission, Parliament) and representatives from regulatory bodies (EMEA (European Medicines Agency), ICH) to improve awareness of the issues and points of concern, could be of great value. Ideas like a one-day face to face meeting were suggested.
- Participation of representatives in such meetings of EU representatives, EFPIA (European Federation of Pharmaceutical Industries and Associations) and EMEA, should be encouraged to proactively drive the ICH-process with respect to alternatives.
- It may be a recommendation that an "Alternatives *ad hoc* group" be created at EFPIA, to better represent the alternative "scene" in the European industry association. Currently, the competence and expertise for the 3Rs are not focused.

Availability of "negative results" to scientists. One of the unsolved problems still is that negative results regarding certain classes of drugs / outcome in a particular animal model, are almost never published, but they can provide important information, and avoid the use of unnecessary animal experiments. Only when both positive and negative results become available real progress could be made. It is very important to have access to the knowledge about negative results and to get a better understanding on why a model doesn't work. There is a lot to be learnt from mistakes before certain *in vitro* models are "killed".

Expert Meeting 2

5 September 2008, Novartis, Basel, CH

5.2. Expert Meeting 2

5.2.1. Executive Summary

The second expert meeting was intended to (partly) cover the bottlenecks in animal disease models used during the drug development process of New Chemical Entities (NCE) by the pharmaceutical industry. Because of declines in clinical success rates of NCE (lack of efficacy, differences in human metabolism, unexpected toxicity) and lack of predictive validity of animal disease models, initiatives to improve research strategies for drug development and efficacy testing were taken including further development of animal disease models, going together with further development and implementation of the 3Rs.
Animal models include initiation and progression of the diseases involved. These either are experimentally induced or expressed via genetically modified experimental animals. The models are incorporated at an early stage in drug research followed by efficacy, safety and toxicity studies.
In depth discussions took place between 28 researchers, actively working with disease animal models in different European, globally active, pharmaceutical companies. In seven workout/disease groups participants were focusing on animal models representing psychiatric diseases, degenerative diseases in the brain, inflammatory/autoimmunity diseases, oncology, infectious diseases, respiratory diseases and metabolic diseases.
Participants were asked to describe the state of the art in their research field and to redefine and update 3Rs bottlenecks/strengths in the animal disease models in their field of expertise and how the 3Rs could be further developed. Disease animal models are not standardised and protocols are optimised to the specific needs and symptoms of a disease, often developed only for internal use only. This decreases the chance to build up experiments on published data and requires considerable resources in each laboratory to establish own standardised protocols. In the past, mainly symptomatically and only when possible mechanistically based disease animal models were in use. Disease animal models allow drugs interacting at different progression stages of a given disease to be tested and their relevance is determined, if available, by drugs already in clinical use. Based on these principles, a summary is given here.

- Further development of the 3Rs:
• The direct replacement of disease animal models by *in vitro* tests seems not to be realistic today since in most diseases, many different cell types are participating and mechanisms involved in the dynamic progression of a disease are often quite complex or not known. However, with data obtained from pre-tests in more complex *in vitro* approaches, e.g. the use of isolated organ, co-cultures or 3D cultures, a better filtering of compounds before *in vivo* experiments start, could be achieved.

• There is a need to increase the work at the interface of preclinical and clinical trials. The more translational the animal models are the less the number of animals that need to be used. Preclinical scientists need to maximize the understanding of what is been done in the clinic, while the clinical researchers should have more knowledge of the preclinical testing.
• Optimisation of protocols through laboratory interactions and workshops is necessary to lead to a reduction of the number of control tests /positive controls /overall sample size. Expert groups from industry and academia should work out recommendations for efficacy profiling in animal models for different diseases, with respect to study design (numbers of subjects, models to use, most relevant efficacy read-outs), and the possibility to use and develop target-based models including relevant pharmacodynamic markers. CRO's (Contract Research Organisations) may have even more knowledge regarding protocols in disease animal models, because they have multiple companies as clients that are approaching similar type of models or strategies. Up to now, this knowledge was not transferred between companies (particularly if a new model is viewed as being a competitive advantage). This information about study design and protocols together with data on compounds should be collected in a database and be used by the contributors. This exchange of information between companies requires new strategies for dealing with shared intellectual properties.
• Maximising the number of read outs per animal would reduce the number of animals used, e.g. in psychological disease models, imaging, PK/PD (pharmacokinetic / pharmacodynamic) measuring and multiple behaviour measures could be done in one animal; in arthritis models, swelling data, MRI (Magnetic Resonance Imaging) data, *micro*CT (X-ray microtomography), blood sampling, synovial fluid and urine samples can be taken of the same animal, while histology of the joints, metabonomics and genomics data can be gathered post mortem. Whenever possible translational, mechanism-based read-outs in animal efficacy testing should be chosen via comparison to corresponding clinical data from patients.
• The screening paradigms should be carefully considered to avoid unpredictive models (reduction), e.g. use of toxin models evaluating the neuroprotective potential of compounds for Parkinson's Disease.
• Non-publishable (negative) data should be made available to the researchers working in a particular disease area. Published negative data must fulfill similar quality standards as the positive ones. Journals/scientific associations should be encouraged to provide a forum within which negative data can be presented and discussed.
• Academia should be involved in developing new *in vitro* models that are predictive of clinical outcomes and that would be accepted by regulatory agencies and companies. This includes the use of isolated organ studies, the development of 3D cell cultures e.g. in oncology or to use human stem cells to grow whole tissue/organs. Human stem cells would eliminate the problem of species differences and improve the predictive ability of the tests. But very few studies have been done to specifically compare the data from such cultures with those of the actual efficacy studies *in vivo*.
• Non-invasive techniques should be encouraged, also for infectious disease models (bacterial infections). The longitudinal observations not only reduce the number of

animals in experiments, but might provide more sensitive/relevant read outs and allow to determine the humane endpoints more exactly.
• Population pharmacokinetic models (PPK, *in silico*) are a powerful technique which can be applied for different diseases. Its utility is particularly valuable in dose setting in the absence of pre-existing animal models (or as a potential method to replace animal models). Much of the development work is already done especially in the human oncology area.
• In general, education and awareness of the 3Rs should be improved e.g. scientists should be trained in the proper use of statistics, and the application of power analyses to determine the minimum number of subjects needed in an experiment.

- Specific disease animal models with high burden to experimental animals:
There are disease areas in which animals are subjected to high and long-lasting burdens.
• Inflammation models are inherently painful, since pain is a cardinal sign of inflammation and analgesics cannot always be used since these could adversely affect the outcome of the model. Even opioids have been shown to have an effect on cell migration in inflammatory models. Some valuable recommendations were presented e.g. tissue chamber model. In pain models, direct studies in man (see also microdosing) would be preferable to let them score their pain because many animal pain models are not very predictive.
• In oncology models, genetically engineered models and primary tumour models (e.g. tumour tissue biopsies taken directly from patients and transplanted into animals) were put forward as being more predictive of the effects seen in the clinic. A refinement strategy could be to apply PK/PD in surrogate tissues, e.g. tissues of normal animals, instead of looking at tumour bearing animals, since many of the pathways that are targeted are fully functional in healthy tissues.

5.2.2. Recommendations

1. Development of human pathways/mechanism based animal disease models (translational animal models, less symptomatic models).

2. Optimisation (quality standards) and harmonisation of study protocols (based on clinical applications) for individual disease models.

3. Maximising the number of non invasive and early or surrogate endpoints (read-outs) within a model (lower burden, more relevant for human diseases, less analgesics, no pain related pathophysiological disturbance of read outs).

4. Development and use of more complex *in vitro* systems (human cells, functional organ specific tissues, co-cultures, slices, *ex vivo*) in which predictive clinical effects can be investigated e.g. human stem cells derived.

5. Exchange of non-publishable protocols and data (type of compound) derived from animal disease models among pharmaceutical industry and publication of negative data.

6. Reduction or replacement of animal efficacy studies and long-term carcinogenicity studies (carried out at later stages of drug development) by intelligent combination of information derived from studies performed during the drug research phase.

7. Development and use of human population pharmacokinetic models instead of non-existing (or not reliable) animal disease models.

5.2.3. Report of Expert Meeting 2

START-UP Expert Meeting 2
5 September 2008, Novartis, Basel, CH

3Rs Bottlenecks in Animal Disease Models

Program:

09:15 – 09:20	Welcome. Peter Maier, *ecopa*, CH
09:20 – 09:30	Introduction. Vera Rogiers, *ecopa*, BE
09:30 – 09:40	Europe and alternatives: programs, initiatives, projects and networks. Bernward Garthoff, *ecopa*, DE
09:40 – 09:50	Instructions. Peter Maier, *ecopa*, CH
09:50 – 10:10	Coffee Break
10:10 – 11:00	Workout Groups: 1) Psychiatric Diseases, 2) Degenerative Diseases, 3) Inflammatory Disease 4) Oncology, 5) Infectious Diseases, 6) Respiratory Diseases, 7) Metabolic Diseases, 8a) Bone disease, 8b) Cardiovascular diseases State of the art: Study design, Reproducibility, Role of the model, Problems regarding 3Rs
11:00 – 12:30	Representatives from individual groups: 5 min presentation, 2 min discussion per group
12:30 – 13:15	Lunch
13:15 – 13:45	Walking tour within the Novartis Campus
13:45 – 13:50	Instructions. Peter Maier, *ecopa*, CH
13:50 – 14:40	Workout Groups 1) Psychiatric Diseases, 2) Degenerative Diseases, 3) Inflammatory Diseases, 4) Oncology, 5) Infectious Diseases, 6) Respiratory Diseases, 7) Metabolic Diseases, 8a) Bone disease, 8b) Cardiovascular diseases Implementation and further development of the 3Rs, Recommendations: for a particular disease area, for a particular model, new concepts, new approaches
14:40 – 15:10	Coffee break
15:10 – 16:10	Representatives from individual groups: 5 min presentation, 2 min discussion per group
16:10 – 16:20	Discussion
16:20 – 16:30	Questionnaire: additional personal recommendations for upcoming 3R Workshops 2009 and Road Map
16:30 – 16:35	Conclusions. Peter Maier, *ecopa*, CH
16:35-18:00	Individual informal discussions in groups

Scientific committee:
Peter Maier, Board Member *ecopa*, Forschung 3R, CH
Bernward Garthoff, treasurer *ecopa*, Bayer, DE
Vera Rogiers, chair *ecopa*, VUB, BE

5.2.3.1. Introduction to the report

The 2nd Expert Meeting, as preparatory meeting for the Workshops on 3Rs of the START-UP projects, addressed specifically disease models, used in pharmaceutical research and development. This is due to the general situation of complex disease symptoms only to be shown in whole organisms and the lack of meaningful alternative methods, especially *in vitro* methods. This situation is predominant in the areas of CNS and disruptive diseases such as arthritis and cardiovascular diseases. Because the experts present were mainly active in these areas, the focus of this report is on these disease areas and models. Thereby, a sound basis for improvements is laid by the START-UP project.

5.2.3.2. Introduction to the Expert Meeting

5.2.3.2.1. Background information on animal disease models, by Peter Maier, Forschung 3R, CH

After a welcome from PM and thanks given to the sponsors, the background for the title of this 2nd meeting (*3Rs Bottlenecks in Animal Disease Models*) was presented.
The focus on bottlenecks in disease models derives from several observations or concerns raised in the past concerning the drug development process. These are:
- Continuing decline in clinical success rates of newly developed drugs. In some disease areas, this may be attributable to a lack of efficacy, but, in other areas, it is more commonly based on species-specific metabolism or unexpected toxicity.
- Biases on preclinical assessment due to the perceived pressure to demonstrate efficacy in preclinical models within drug discovery are described and target-based discovery might increase the effects of bias (Lindner, 2007).
- Apparent lack of predictive validity of animal models in some disease areas (McArthur and Borsini, 2006), typically in diseases where the disease mechanism is either poorly understood or very difficult to model in an animal species.

Using observations such as these, some have raised serious doubts on the use of animal models as a whole.
On the other hand initiatives are going on which improve research strategies for drug development and efficacy testing. This includes changes in the use of animal disease models. Modelling of human disease in a variety of animal models has provided a distinct advantage in the search for new therapies and in disease areas where animal models are adept at predicting efficacy. Further refinement of the models may allow better prediction of clinical PK and toxicology.
This further development of models has to go together with further development and implementation of the 3Rs. This is a challenge and an opportunity for reaping benefits of every kind - scientific, economic and humanitarian. This workshop was intended to examine some of the challenges that are faced as academia, industry and government try

to move towards increasingly useful and refined animal models as well as to look at some of the areas where progress has been made in minimizing the need for animal models.

5.2.3.2.2. Background information on ecopa and START-UP project, by Vera Rogiers, *ecopa*

After the introduction, VR explained to the industry experts the role of *ecopa* and the ultimate aim of the START-UP project.

5.2.3.2.3. Background information on Europe and 3R alternatives: programs, initiatives, projects and networks, by Bernward Garthoff, *ecopa*

An overview of what is done in Europe with respect to 3R alternative methods was presented; which projects and initiatives are undertaken and which networks are active in this field.
The diversity of European bodies that are dealing with issues of 3R-alternatives is impressive. These are organised at different levels, for example in the EU there are organisations like Estiv (European Society of Toxicology *In vitro*), IVTIP (*In vitro* Testing Industrial Platform), *ecopa* (European Consensus Platform on 3R Alternatives), ECVAM (European Centre for Validation of Alternative Methods) at JRC (Joint Research Centre) and efsa (European Food Safety Authority) consultation processes (e.g. also addressing the new Directive 86/609). In the member and associated states there are local national platforms, for example the NCPs (national consensus platforms) connected with *ecopa*, like SET (Germany), Norecopa (Norway), IPAM (Italy), 3R Research Foundation (Switzerland)… Indeed, an initiative like *ecopa* is especially specialised in networking between the different countries and between the different stakeholders concerned with animal testing.
Programs such as the EU R+D Framework Programmes 5, 6 & 7 sponsor projects in alternative methodology with increasing amounts. The total funds in 2008 have run up to €237 Mio. The attention given to alternative methods has especially gained in importance since the REACH legislation. Examples of such projects are Esnats, ReProTect, START-UP, CarcinoGENOMICS…, which are mostly projects between academy and industry participants.
The European Commission itself has launched initiatives, such as EPAA (The European Partnership for Alternative Approaches to Animal Testing), IMI (The Innovative Medicines Initiative) and DG RTD (Directorate-General for Research) at the same time.
The START-UP project is an FP7 project, which aims at making a roadmap to be used as a guideline for future EU work programmes. The first Expert Meeting took place in Madrid, in the Ministry of Health, and was also a closed meeting with invited experts.
This expert meeting here is the second one and the third one will be held in Alicante, Spain.

5.2.3.2.4. Workshop Instructions, by Peter Maier, Forschung 3R, CH

After the introductions, PM explained the objectives of the meeting, namely to redefine and update 3Rs bottlenecks/strengths in Animal Disease Models in the preclinical phase of drug development, in specific disease areas and animal models and with respect to the type of drug candidates.
The information gained will contribute to the development of a Road Map for further implementation of the 3Rs within coming EU Framework Programs and National Initiatives.
Participants were divided into 7 working groups, focusing on themes such as psychiatric and degenerative diseases in the brain, inflammatory/autoimmunity diseases, oncology, infectious diseases, respiratory diseases and metabolic diseases. These are all areas where animals are used to model aspects of the disease to a major extent.
First the state of the art of each of these fields was discussed and bottlenecks in 3Rs development were identified.
Afterwards discussions were leading to recommendations about possible solutions for some of the problems raised and where the 3Rs could be further developed.
The outcome of the discussions and conclusions were presented to and discussed by the other participants.

5.2.3.3. Summary of the presentations of the individual working groups

5.2.3.3.1. Brain: Psychiatric diseases, by Christopher Pryce, Novartis, CH

5.2.3.3.1.1. State of the art, identification of 3Rs bottlenecks:

There are four major psychiatric diseases in terms of incidence: depression, anxiety disorders, schizophrenia and sleep disorders. These diseases are severe and common, e.g. by 2020 it is estimated that depression will be the second most prevalent disease, second only to heart disease.
This is a very complex group of diseases. They are difficult to understand and treat in humans and increased understanding and future treatment are dependent on animal studies.
Traditionally modelling of psychiatric disorders has sought to recapitulate its key behavioural features as described in diagnostic classification schemes such as DSM-IV or ICD-10 (APA (American Psychological Association), 1994; WHO (World Health Organisation), 1992). An example of such a traditional model of depression, for example, is the “forced swim test”. During this test, rats or mice are placed in a jar of water from which they swim and struggle to escape. When exposed to this procedure subsequently they become inactive; an inactivity that has been interpreted as feeling of helplessness and despair analogous to the lack of activity and feelings of lack of worth and helplessness reported in depressed patients (McArthur and Borsini, 2006). Clinically

active pharmaceuticals are known to counteract this inactivity and consequently this model has been used extensively to compare the effects of new chemical compounds against those of clinically active drugs as a preclinical measure of efficacy before the compound is progressed for clinical development. Another interpretation, however, is that the animals stop swimming to save energy, until something happens so that they can escape, which makes it an adaptive response/strategy. This example shows that interpretation of read-outs of animal models based on symptoms only, as performed in the past, is very difficult.

The limitations of diagnostic criteria and the present emphasis of developing more clearly defined research criteria of psychiatric disorders (e.g., Cooper, 2003; Hyman, 2007), and the virtual impossibility of modeling all aspects of a complex disorder in a single animal, academics and industrial researchers are focusing on modeling specific phenotypic or behavioural changes of the disorder and correlating them with changes in brain function and molecular alterations, that is the "endophenotypic approach" e.g., Cryan and Holmes, 2005; van Os and Tamminga, 2007).

A representative example of this is the "fear conditioning test". It is based on a specific brain region, the amygdale, which shows higher activity in psychiatric diseases, e.g. anxiety, in humans (Bishop, 2007). The homologous region exists in the brain of rodents. By understanding the behavioural neurobiology of the amygdale and fear conditioning in rats and mice, the importance of the altered activity of the amygdale in human psychiatric diseases, and potential drug targets for modification of this altered activity, can be studied. The fear conditioning test in rodents can then also be used as a predictive translational test to do compound screening (Delgado et al., 2006). Thus, more robust interpretations of the behavioural changes induced by the model of the disorder and its pharmacological modification can be made, thereby improving its predictive validity (cf., Millan, 2008).

The following major challenges and bottlenecks were identified:

A major *challenge* is the development of brain function-related or phenotypic disease models requiring basic and translational research specific in symptomatology or biomarkers, so that the relationship between the changes in animal behaviour and expected changes in patients following drug treatment can be accurately predicted.

A *bottleneck* is the subtle, but potentially important procedural variations introduced by different labs when establishing a particular model or procedure such as the forced swim or fear conditioning tests described previously. These subtle changes are introduced pragmatically in response to environmental realities and instrumentation availability. For example, many companies produce their own versions of apparently the same equipment, and the choice of buying equipment from one company or another, or building one's own equipment, is determined by many reasons. The models and procedures, however, are "standardised" through the demonstration that clinically-effective standards are active and therefore internal and external consistency can be achieved. Notwithstanding this standardisation through pharmacological isomorphism (Miczek, 2008), it is at times difficult to reproduce and compare results obtained between laboratories. It was proposed to organise a workshop, which would not necessarily aim to dictate or impose fixed protocols across laboratories, but to compare protocol details that would allow identification of critical parameters that would help increase the validity of the animal

model of the human disorder to predict the efficacy of new pharmaceuticals developed for its treatment.

5.2.3.3.1.2. Implementation, further development of the 3Rs and Recommendations:

Optimisation of protocols through laboratory interactions and workshops is necessary to lead to a reduction of the number of control tests and positive controls as well as a reduction of the overall sample size.
Maximizing the number of read outs that is obtained from each animal would also reduce the number of animals used. Imaging, PK/PD measuring and multiple behaviour measures could possibly all be done in one animal. For example, when a specific behaviour is being measured, at the same time the general locomotion could be measured. Traditionally locomotion and fear condition freezing are measured in different tests, whereas these could be measured in one test with automated systems.
There is a need to increase the work at the interface of preclinical and clinical trials. The more translational the animal models are the less the number of animals that need to be used. Preclinical scientists need to maximize the understanding of what is been done in the clinic, while the clinical researchers should have more knowledge of the preclinical testing. This should lead to less negative clinical trials and maybe to more rapid drug development and reduce the number of animals that need to be used.

5.2.3.3.2. Brain: degenerative diseases, Christoph Wiessner, Merck-Serono, CH

5.2.3.3.2.1. State of the art, identification of 3Rs bottlenecks

The biggest problem with neurodegenerative diseases, like Alzheimer's Disease, Parkinson's Disease and Huntington's Disease, is that the etiology and the mechanisms of the degenerative processes are still largely unknown. Currently no treatments are available to prevent disease progression in patients contributing to lack of validated predictive models for neurodegenerative diseases. It has been impossible to develop clinically predictive *in vivo*, and certainly *in vitro*, models for neurodegenerative diseases, or to recommend standard *in vitro* or *in vivo* tests for efficacy profiling for potentially disease modifying drugs. This has resulted in the proliferation of many different testing methods and consequent use of animals.
Some suggestions and ideas which might help to overcome some bottlenecks in neurodegenerative disease animal models:
- Expert groups from industry and academia should work out recommendations for efficacy profiling in animal models for neurodegenerative diseases, with respect to study design, addressing topics such as: numbers of subjects, which models to use, what are the most relevant efficacy read-outs, and the possibility to use and develop target-based models including relevant pharmacodynamic markers.

- Improvement of the evaluation of translational, mechanism-based read-outs in animal efficacy testing, by using techniques such as genomics, proteomics, metabonomics and imaging is necessary. Thereby the relevance of individual animal studies could be increased. In addition comparison to corresponding clinical data from patients will help to better match animal and cell culture models to the human disease.
- Increased application of patient-derived IPS (Infused Pluripotent Stem cells) cell models should also be considered.

5.2.3.3.2.2. Implementation, further development of the 3Rs and recommendations

- In view of reduction, the screening paradigms should be carefully considered to avoid unpredictive models, such as the use of toxin models evaluating the neuroprotective potential of compounds for in Parkinson's Disease (Note: toxin models are still very important in the evaluation of improved symptomatic treatments for PD as well as the development of treatments for L-DOPA induced dyskinesias [Lane and Dunnett, 2008]).
- Development of imaging techniques should be supported.
- In terms of replacement and reduction, more studies on the effects of disease-relevant mutations on cell lines and lower animals, such as *C. elegans*, need to be carried out to show their real possibility to be used as a screening platform. Careful comparison with whole-animal study and available clinical data (such as genomics, proteomics, metabonomics, disease-relevant signalling pathways, specificity of compound effects, that is, do compounds hit the same target and mechanism in *C. elegans* versus mammalian system, etc.) will be required to validate these models.
- Over the past 2 decades animal models for Alzheimer's Disease and other neurodegenerative diseases are based more often on transgenic mouse lines expressing disease-relevant genes. However, mechanical and pharmacological lesioning technique are still being used (cf, Lindner et al., 2008). In terms of refinement, the development of more clinically relevant biomarkers for endpoints in long-term studies could be promoted such as, for example, functional outcome instead of time-to-death in ALS SOD1 transgenic models or in Huntington's Disease mouse models. Ideally such functional outcome measures should be closely related to the (pathological) function of the disease-relevant gene/protein in order and allow translation to clinical read-outs later.
- The use of inducible pluripotent stem cells or differentiation of embryonic stem cells might result in a better understanding of disease processes that can hopefully be "back-translated" to *in vivo* models and improve the quality and predictability of efficacy results of potentially neuroprotective compounds.
- *In silico* models for target relevance (such organ-specific signalling pathways, involvement in disease relevant pathways, interaction with other targets, etc.), drug properties (prediction of brain penetration, metabolism, transport), brain structures and connectivity clearly need further development, which should help increase predictability of results from animal data.

Predictability of *in vivo* and *in vitro* models of blood brain barrier penetration and brain exposure of drugs is a topic that is being actively discussed by academics and within European projects. There were questions raised whether available in vitro models are also

used to their full potential within industry. For in vitro systems assessing the affinity of new compounds for efflux systems (P-gp encoded by the multidrug-resistance gene MDR1) there is broad consensus that they provide critical and predictive information regarding potential brain exposure of compounds. With respect to attempts to model the blood-brain barrier in vitro (such as co-culturing of immortalized endothelial cells with astrocytes on filters, etc.) there is concern about the predictability of models available presently that limits their generalized use within industry, especially when new and hitherto unknown cemical space is explored. It is widely accepted that increasing the knowledge on how the brain controls in-and outflux of molecules using in vitro systems will help to improve CNS drug development with the potential to reduce the number of certain *in vivo* experiments such as PK studies.
Microdialysis in the brain is also an emerging refinement technique, but it is very expensive and is only occasionally used.

5.2.3.3.3. Inflammatory diseases, Päivi Alajuuma, Santen, FI & Janet Dawson, Novartis, CH

5.2.3.3.3.1. State of the art, identification of 3Rs bottlenecks:

Glaucoma and corneal defects:
Nowadays, *in vitro* studies are applied at a preliminary stage in drug discovery and *in vivo* studies are used to check for eye irritation, Currently there is no good *in vitro* eye disease model to study pharmaceuticals. The *in vitro* bovine cornea test is only used for chemicals, and is not refined enough for pharmaceutical drugs. *In vitro* models should be developed for testing for e.g. redness or slight corneal defects.
For aqueous humour outflow (in glaucoma research) there is a good *in vitro* model, using bovine and porcine eyes, but this is only usable for early stages. Rats, rabbits and monkeys are still used. It will probably be hard to totally replace them.

Acute inflammation models:
Acute inflammation models, are mainly established in rats. They are in use for drug candidates which are developed for treatment of pain, fever, swelling. One of the most frequently used model of pain is the “Randall-Selitto hyperalgesia” model, in which the pro-inflammatory substance is injected into the plantar surface of the rat’s foot and then the pressure eliciting a response (vocalization or withdrawal) is measured. Fever is induced through the injection of bacterial endotoxins in the form of LPS (lipopolysaccharide). The effects determined with this model are comparable in rats and in humans and results therefore highly translatable. Edema s modeled through the injection of carragenan into the paws of rats and the foot swelling is measured. All these protocols cause low pain to the rats and they do not cause long-term damage.
It was mentioned during the discussion that a number of groups use an acute synovitis model in dogs. This model, has very little variability associated with it and has been

accepted by FDA and EMEA for several registration packages of veterinary medicines. It has been used successfully as a preclinical evaluation for human development.

Mechanistic models:
Natural or artificial cavities are used in mechanistic models. Small Teflon chambers can be implanted in the backs of rats. Injection of pro-inflammatory substances can induce inflammation in these areas. In these chambers, mediators and cell influx can be easily measured. These models can be adapted and designed for particular questions. The tissue chamber models allows the reduction of the number of animals used through repeated sampling in the same animal at different time points.

Arthritis models:
Arthritis models are time-intensive as chronic inflammation is induced in animals. A rigorous pain assessment is performed that monitors the progression of arthritis and the efficacy of new drug treatments. This allows the experimenter to terminate an study in the case of lack of efficacy. Although the induction of arthritis is different between rats and humans, the resulting pathology is quite similar.
In general compounds are first tested *in vitro* and only if they have an effect at cellular levels (e.g. inhibition of TNFalpha or IL-1beta production/release, etc), which are known to correspond to an analgesic or antiphlogistic activity. Successful compounds are then tested *in vivo* for efficacy.

5.2.3.3.3.2. Implementation, further development of the 3Rs and recommendations:

Glaucoma and corneal defects:
In vitro models for corneal opacity, using bovine eyes, are already available. Other models of ophthalmic disorders with different endpoints are needed.
Cell cultures and cultured corneal tissue are available, but it is still a question whether this 3D model is actually usable to the full extent.
In regard to allergic conjunctival models (for late phase allergic reaction) rat and mouse models are available. In the future, a comparison should be done to determine whether the mouse is a more sensitive species than the rat.
There is a great intraspecies variability with respect to the allergic reaction response. The question is whether these problems could be reduced by using different species or even a different strain of animals, and thereby, refining it.

Inflammation models:
One issue in this field is that inflammation models are inherently painful, since pain is a cardinal sign of inflammation. Analgesics usually can not be used since these can adversely affect the outcome of the model. Even opioids have been shown to have an effect on cell migration in inflammatory models. It is almost never possible to use any sort of pain relief in these models.
In the past, arthritis model studies extended over periods of months. These have now changed and now rarely extend past 3 weeks. This is because controls reach a plateau,

which makes further testing unnecessary. This is one way of refinement that is already put into practice to a major extent.
The regulatory authorities can not easily be convinced that 3R models are relevant and useful.

In terms of recommendations some proposals were brought up:
- The tissue chamber model is a refinement of the peritonitis or air pouch model, since animals can be used as their own control. This reduces the number of groups that has to be used. Multiple time-points can be sampled from the same animal, this also reduces the number of groups needed.
- Many pain models are not very predictive, so direct studies in man would be preferable. It is also much easier with humans to let them score their pain.
-When sufficient toxicity tests have been carried out for the compound under consideration, then single dose tests in healthy volunteers could be carried out.
- Arthritis experiments should be very well planned and co-ordinated so that as much data as possible is obtained from a single experiment, such as swelling data, MRI (Magnetic Resonance Imaging) data, microCT, blood sampling, synovial fluid and urine can be taken, histology of the joints, metabonomics and genomics data.
- Much of this information can be gathered post mortem, so that the number of animal interventions can be limited.
- Arthritis research probably can not be done fully *in vitro*, because for example in the joints many different cell types are present and the joint structure and mechanics of joint movement quite complex.

5.2.3.3.4. Oncology by Robert Cozens, Novartis, CH

5.2.3.3.4.1. State of the art, identification of 3Rs bottlenecks

The majority of models in oncology are human or rodent genetically engineered tumour xenografts in immunocompromised animals. In oncology almost exclusively mice and rats are used. The tumours can be provoked by subcutaneous (tumours grow under the skin) or orthotopic (tumour grows in organ of origin) injections of tumour cells. There are many types of tumours, about 900 tumour cell lines are available, of which many, but not all grow in nude mice (or other immunocompromised strains) and many of these will also grow in nude rats. They are all distinct, so there is enough potential to test a compound in a large number of models.
In the last years genetically engineered mouse models have gained importance. Their genomes have been manipulated to change or control the expression of specific genes that has the consequence that spontaneous tumours are formed. In some cases spontaneous tumours can be transplanted into recipient mice for evaluation of therapies.
In targeted therapy complementarity is high, but the question is whether the large number of models is interchangeable. Do we have to test everything in all models? In targeted therapy this may not be strictly necessary, if the target drives the tumour in multiple

tumour types, then the effect in one tumour should predict the effect in other tumour types, however it is known that this is too simplistic.
The choice of models is dependent on the presence and importance of the target to tumour growth. Ideally, one would like to have tumours for drug discovery and development that are dependent on a certain target for growth. The inhibition of that target will result in no or reduced growth of the tumour. There are compounds, which are called neo-cytotoxic, they will not only target the tumour cells, but also normal replicating cells, in this case the choice of model may be based on tumour type as essentially all replicating cells possess the target
Information exchange is a common problem in oncology, particularly between companies. It would also be good to have access to negative data. It cannot be expected that negative results at specific targets are reported but, for example, it would be useful to know whether a particular tumour grows in a specific mouse strain, so that the test does not have to be repeated. Other useful information would be to provide tolerability data from already marketed anticancer drugs.
Reproducibility is another issue, because it is known that cell lines, which are nominally the same, may have differences in growth and/or response when coming from different labs. Also in different strains of immunosuppressed animals, there might be differences in results or even in the same strain from different suppliers.
The most important role of the animal model is to identify drug candidates for further development. This part of research probably uses fewer animals than profiling compounds during drug optimisation and profiling development compounds.
Profiling of development compounds is needed to extend the indications such as a new combination of drugs. The use of combinations in cancer treatment to support clinical trials is the rule rather than the exception, but this increases the number of animals considerably.
It is difficult to make conclusions on the relevance of using these animal models in oncology, since there are many different types of cancers, which makes it difficult to conclude by a handful of models, how a compound will work in the clinical trials for a particular tumour type.

5.2.3.3.4.2. Implementation, further development of the 3Rs and recommendations

Genetically engineered models are being put forward as being more predictive of the effects seen in the clinic, but actually few data supporting this view are available. It should be investigated whether transgenic models are indeed more predictive.
For example, a mutated gene leading to breast cancer in humans can be introduced into the mouse. The tumours develop stochastically. In some animals tumours might develop immediately, in others it could take some months. In this latency period, other mutations might take place, as it is generally believed that to develop cancer more than one mutation is needed. So it is thought that the stochastic variety that develops is similar to that in the patients.
So far there are, however, little or no data available yet that would back up definitely that data in such models is more predictive of that obtained in clinical trials.

In primary tumour models, e.g. tumour tissue biopsies taken directly from patients and transplanted into animals are also thought to be more predictive of the clinical situation. In this field few or no prospective studies have been done to prove this one way or the other. Gathering data from earlier research in these models and comparing these with the outcome of clinical trials would result in useful information.
In terms of complementarity one important question is how many models are needed to convince regulatory authorities or clinical investigators that a new drug is worthwhile to go into clinical trials.
Cancer is a complex disease and many companies probably can not afford the time and the animals to do all the tests that some authorities expect to be done. To get further clarity on the expectation and requirements of regulatory authorities may reduced unnecessary experimentation.
PK/PD studies could replace efficacy studies. They are already often used in terms of compound optimisation and selecting compounds for efficacy studies. The question is whether PK/PD studies alone could be used to select the compounds for development or even to further go into clinical trials.
3D cell cultures *in vitro* have been suggested to be more predictive of the *in vivo* situation than standard cultures in efficacy studies, but very few studies have been done to specifically compare the data from such cultures with those of the actual efficacy studies *in vivo*.
Cancer and infectious diseases have many elements in common. In the past, *in vitro* PK models have been used for infectious diseases. It would be worthwhile to look at whether these could be also used in oncology.
A refinement strategy could be to apply PK/PD in surrogate tissues, which is sometimes done in clinical trials. Instead of looking at tumour bearing animals, tissues of normal animals, could be used since many of the pathways that are targeted are fully functional in healthy tissues.
Gene reporters, in tumour bearing or healthy animals, could also be used more often.
It was also emphasised that it would be very useful if non-publishable data could be in some way made available to the oncology community.
As soon as clinical data are obtained, which usually are PK data, the PK profile could be reproduced in an animal and the efficacy of the compound can be studied under those circumstances. This has not yet been done extensively, but would probably deliver interesting results regarding the best way to use animal models in oncology drug discovery and development.

5.2.3.3.5. Infectious disease, by Barry Wright, AstraZeneca, USA

5.2.3.3.5.1. State of the art, identification of 3Rs bottlenecks

There are relatively few viral animal models available for human viruses. As a rule, anti-viral pharmaceuticals are generally progressed directly to the clinic after safety studies.

The efficacy of animal tests for anti-bacterial activity is held to be predictive of the outcome in the clinical effects. The infectious agents used in research are the same in clinical trials as in the preclinical and *in vitro* tests.
In vitro studies are carried out and have to show a positive effect before compounds are progressed for *in vivo* testing. This limits the number of compounds which are actually tested in animals.
Surrogate endpoints can also be used, because many of the diseases are fairly acute in animals. When the bacterial infection has reached a certain level, the test can be stopped, since it is known that the animal will succumb to the infection. By this method the suffering and pain for the animals is reduced.
Bacterial disease models, especially for respiratory diseases, are not amenable to the use of analgesics proposed to reduce the pain associated with the models. Part of that is due to a lack of knowledge on how non-inflammatory analgesics impair the model. It is, however, known that anti-inflammatory analgesics do have an impact on the model.

Bottlenecks:
- *In vivo* animal models, as well as more complex *in vitro* models for anti-viral research is required. These would help establish the predictive validity of preclinical discovery with clinical outcome and lower the cost for this research.
- There is a difficulty in agreement with regulatory agencies regarding the predictive value of *in vitro* methods.
- Specific strains of bacteria are sometimes published that could be useful in a certain model. The strain, however, is not available and can thus not be used for own research. So it would be good if those strains could become publicly available.
- The cost of running clinical experiments is huge, leading to the need for predictive animal work, so that no unnecessary clinical experiments are conducted.
- The cost of animal models also drives refinement. So if *in vitro* tests could be proven to be predictive of the effects in the clinic, less animal testing could be done, so that more and/or better data might be available for the same cost.
- Timeline and cost pressures inhibit companies from expending internal time/resource to develop more predictive *in vitro* or organoculture models. Is there a greater role for academia in this arena?

5.2.3.3.5.2. Implementation, further development of the 3Rs and recommendations

Validation of microdosing in humans to eliminate some or all of the PK work in preclinical species could be an option, but the ethical question of validating a model in humans needs to be solved.
Imaging of infectious agents *in vivo* should be made more sensitive in order to allow monitoring of effects over time, at least for bacterial infections this could be an improvement.
Negative data, e.g. when a model is not predictive, should be disclosed. Since there would be issues with intellectual property associated with chemical structures, it might be an option to start with generic drugs.

Published negative data must fulfill similar quality standards as the positive ones, and that journals/scientific associations should be encouraged to provide a forum within which negative data can be presented and discussed.
Independent review of data to determine numbers of animal experiments proposed or done within a clinical class of compounds could also lead to the reduction or refinement of protocols.
Improved PK/PD models that are increasingly predictive are being developed at the moment.
Population pharmacokinetic models (PPK, *in silico*) are not often used, since their predictability is not certain yet. It is often applied to generics only, due to intellectual property problems. This should be overcome via coded data and neutral data analysis.
Regulatory acceptance of submission data in absence of animal efficacy data (prior to first time in man) is also something to be worked on. It is already ongoing in some viral diseases. Such an approach could be extended to other disease areas.
Academics could be involved to mine preclinical data and look for predictive outcomes for clinical trials. Academia might also be involved in developing new *in vitro* models that are predictive of clinical outcomes and that would be accepted by regulatory agencies and companies.
- More peer-review of protocols or methods would be useful in order to refine models.
- CRO's (Contract Research Organisations) may have more knowledge, because they have multiple companies as clients that are approaching similar type of models or strategies, but this knowledge is not transferred between companies (particularly if a new model is viewed as being a competitive advantage). Companies might thus be using models from 20 years ago, without knowing about more recent models.

5.2.3.3.6. Respiratory diseases by Alexandre Trifilieff, Novartis, CH

5.2.3.3.6.1. State of the art, identification of 3Rs bottlenecks

The focus was on asthma and chronic pulmonary disease and models for these diseases. In the research for these diseases the severity is usually moderate. The techniques used are usually terminal, making the opportunities for refinement rather limited.
The main models used are small rodent models (mouse and rat) and the morphology of rodent's lungs is different from humans (i.e., rodents do not have alveoli). This is a problem in lung function measurement or when trying to induce structural changes. This of course poses problems for the predictability of the models. However, since lung structure is not so important in the development of inflammation, the models are considered predictive when used for studying anti-inflammatory approaches.
There is a lack of standardisation across laboratories. If one takes asthma, as an example the classical ovalbumin model, every lab has a different protocol. So there is a need for harmonisation of techniques.
In the past lung function measurements were exclusively done in anesthetized animals, using terminal procedures. In the last ten years a new technique (whole body

plethysmograph) has been developed, which allows performing lung function tests this in conscious animals. This has the advantage of being able to perform longitudinal studies in the same animal and therefore has the potential of reducing the number animals. However there is controversy around this test and leading respiratory journals have decided not to publish work that use this technique. Because of this, the known and validated tests are used and the reductive potential of tests is in most cases not put to practice.
Much work is done with inhaled products. This, however, also poses a problem since rodents are nose only breathers. Aerosol treatments, as a way to mimic the clinical dosing, is often used in rodents. However, because the nose is highly vascularised, this might lead to increased systemic exposure. Therefore, intratracheal delivery is mainly used, by which the product is injected into the lung with a syringe.
However, in man, 80% of inhaled drugs will hit the back of the throat and end up in the stomach. This could contribute to systemic side effects, which are not detectable in rodents, when the intratracheal route is used.
Isolated organ experiments are not commonly used. This type of assay is certainly helpful for the refinement/reduction. One animal can give an answer for four different conditions, reducing the number of animals per experiment.
The publication of negative data could be very helpful in saving animals. At the moment the acceptance of negative data in journals is limited. The inability to evaluate the true efficacy of certain tests would help reduce the number of experiments in different laboratories and companies.

5.2.3.3.6.2. Implementation, further development of the 3Rs and recommendations

In the last 10 years, many non invasive techniques have been developed, such as MRI based techniques or lung function measurement in conscious animals. But within the pulmonary field very few groups are working with MRI, for example. This means that such techniques are not yet fully characterized. It would be recommended that industry and academia be encouraged to develop more of these non invasive techniques. This will not only reduce the number of animals, but also lessen the burden for the animals used.
There are two major congresses in the respiratory field, being those of the American and the European Respiratory Societies, but when looking at their agenda the 3Rs are not considered there. The concept of 3Rs should be introduced in these congresses.
There are some naturally occurring animal models, as asthma can develop spontaneously in cats, sheep and monkeys and COPD (Chronic Obstructive Pulmonary Disease) in horses. These diseases in animals, however, are not thought to be fully representative of the human disease. Furthermore they are also poorly characterized. It is concluded that the use of these naturally occurring models is quite limited and very few groups are working on these models.
Another recommendation is to increase the use of isolated organ studies. This can deliver a lot of information, especially for pharmacology, but it is not very often applied.
A general problem, applying to all fields, is that a lot of scientists seem to have a lack of knowledge of the proper use of statistics, and the application of power analyses to

determine the minimum number of subjects needed in an experiment. Numbers of 8 to 10 animals per group are typically selected on the basis of literature and/or preference. The number of subjects needed is dependent upon the robustness and properties of the techniques used and the inherent variability of the results.
Stem cells could be used to grow whole organs that can then be used for testing. Human stem cells would eliminate the problem of species differences and improve the predictive ability of the tests.
In silico approaches are also being developed to target anti-inflammatory pathways. These approaches are not fully established as the outcome is dependent on the quality of the input.

5.2.3.3.7. Metabolism, by Robert McArthur, McArthur and associates, CH

5.2.3.3.7.1. State of the art, identification of 3Rs bottlenecks

Metabolic diseases are complex diseases, characterized by both pathophysiological and behavioural disorders. The focus of this group is mainly on diabetes and obesity.
Compounds available for obesity, are not very effective for long-term weight reduction or have major side effects limiting their use. Consequently, the pharmaceutical industry in general is not very keen on developing anti-obesity drugs. In addition none of the marketed diabetes drugs although effective in lowering blood glucose, modifies the disease progression and effectively treats diabetic microvascular and macrovascular complications.
The models used for obesity diabetes or dyslipidemia can differ and depend on the specific target of interest. Genetic animals (spontaneous mutations) are widely used for target based drug discovery. However, models of induced obesity/diabetes are more relevant to the human phenotype.
The screening cascade and the numbers of animals used are as follows:
- Target identification and validation as well as early safety toxicology is mostly done *in vitro* (tissue, cell lines…).
- In general PK studies are the first *in vivo* studies performed. For these tests less than 10 animals are used per compound.
- Following a satisfactory *in vitro* and PK profile the first efficacy studies are performed, in mice or rats, if possible acutely or subchronically. About 6 to 8 subjects are used per group, and usually there are 4 to 5 groups. This makes about 30-40 animals used per compound at this stage.
- For longer chronic treatment the same is done, so this uses again about 30-40 animals.
In most studies a reference compound (usually a marketed drug) is included as a positive or negative reference. The duration of every study is completely dependent on the mechanism of action of a specific compound.
If all of the above mentioned tests still indicate the efficacy and safety of a drug candidate, either mechanism of action specific studies or studies with a second species could be performed in dogs, mini-pigs or non human primates (rarely).

In Switzerland, for example, the reuse of rodents is restricted to experiments with low severity degree (e.g. PK studies) but as a rule applied with dogs, mini-pigs and monkeys if the wash out period is long enough.
The important thing about metabolic research is that one can get multiple read outs (physiological, anatomical, behavioural, toxicity) from the same animal rather than performing parallel studies. In order to reduce the number of animals used, one should always try to get as much information as possible from the same animal. Thus, a very thorough experimental design that addresses all possible questions without compromising the study is very important.
Feeding is important in such animal models as the metabolic syndrome can be induced by modifying their diet and the behavioural mode of action, specificity and toxicity of potential anorexic drugs can be accurately assessed (Dourish et al., 2008). Multiple behavioural parameters such as feeding, drinking, locomotor activity, as well as physiological parameters such as heart rate, metabolism, EEG can be measured from the same animal over long periods of time.

5.2.3.3.7.2. Implementation, further development of the 3Rs and recommendations

Reducing the number of animals would probably be difficult, since rather low numbers are already used. However, incorporation of multiple assessments in an experimental design allows each animal to be its own control, and the data can be stored for later evaluation, such as the videotaping of experiments. This reduces the number of animals needed for particular experiments as well as avoids having to do sequential experiments with other animals in order to assess the effects of an experimental compound on other parameters.
A database on standard procedures, data on compound etc would be very valuable, so that information between companies could be more easily shared, even though this is quite unrealistic.
Experimental medicine as an info basis has a lot to contribute. Biomarkers could be used, but bio patterns would probably give more information. Therefore there should be more emphasis on metabolics. This would not help in reducing the number of animals in the first place, however it would feed back to the discovery of new compounds and mechanistic understanding.
In general, it is doubtful that animal models will become fully replaced, because in metabolic diseases, the effects derive from all tissues, which requires the complexity of an intact organism.
It is also doubtful that there will be real cures for metabolic diseases, since there are always side effects and long term chronic toxicity to be observed. In fact the drugs on the market aid to prevent and/or slow down the development of metabolic diseases. This means that patients have to take these drugs for several decades.

5.2.3.3.8. General plenum discussion, cross sectional issues:

Reduction of animal studies could be obtained in industry by an optimisation of the process of compound development towards *in vitro* testing, in order to obtain a better filtering of compounds before *in vivo* experiments are started.
Population pharmacokinetic models (PPK) are a powerful technique which can be applied for different diseases. Its utility is particularly valuable in dose setting in the absence of pre-existing animal models (or as a potential method to replace animal models). Much of the development work is already done for this particular application especially in the human oncology area.
Microdosing which have application across individual areas – it may be worth giving this point more prominence in the final Road Map within the START-UP project.

This report was edited by the participants of the meeting, by the corresponding working groups and speakers and coordinated by Peter Maier.

This internal report will be used as guidance for the following 3Rs workshops within the EU project START-UP. The dissemination level is defined as restricted to all participants of the meeting. This final version was approved by the managing team of START-UP at the *ecopa* board meeting in Rome, 25th of February 2009.

Literature cited

American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders. Fourth ed, ed. A.P. Association. 1994, Washington, D.C.: American Psychiatric Association.
Bishop, SJ. Neurocognitive mechanisms of anxiety: an integrative account. Trends Cogn Sci, 2007. 11(7): p. 307-16
Cooper, J.E., Prospects for Chapter V of ICD-11 and DSM-V. The British Journal of Psychiatry, 2003. 183(5): p. 379-381.
Cryan, JF and Holmes, A. Model organisms: The ascent of mouse: advances in modelling human depression and anxiety. Nat Rev Drug Discov, 2005. 4(9): p. 775-90.
Delgado, MR, Olsson, A and Phelps, EA. Extending animal models of fear conditioning to humans. Biol Psychol, 2006. 73(1): p. 39-48
Dourish, C.T., J.P.H. Wilding, and J.C.G. Halford, Anti-obesity drugs: From animal models to clinical efficacy, in Animal and Translational Models for CNS Drug Discovery: Reward Deficit Disorders, R.A. McArthur and F. Borsini, Editors. 2008, Academic Press: Elsevier: New York. p. 271-315
Hyman, SE. Can neuroscience be integrated into the DSM-V? Nat Rev Neurosci, 2007. 8(9): p. 725-32.
Lane, E and S Dunnett, Animal models of Parkinson's disease and L-dopa induced dyskinesia: how close are we to the clinic? Psychopharmacology (Berl), 2008. 199(3): p. 303-12.

Lindner MD. Clinical attrition due to biased preclinical assessments of potential efficacy. Pharmacology & Therapeutics 115 (2007) 148–175.
Lindner MD, McArthur RA, Deadwyler SA, Hampson RE, Tariot PN (2008) Development, optimisation and use of preclinical behavioral models to maximise the productivity of drug discovery for Alzheimer's Disease, in Animal and Translational Models for CNS Drug Discovery: Neurologic Disorders, R.A. McArthur and F. Borsini, Editors. 2008, Academic Press: Elsevier: New York. p. 93-157.
McArthur R and Borsini F. Animal models of depression in drug discovery: A historical perspective. Pharmacology, Biochemistry and Behavior 84 (2006) 436–452
Millan, MJ. The discovery and development of pharmacotherapy for psychiatric disorders: A critical survey of animal and translational models, and perspectives for their improvement, in Animal and Translational Models for CNS Drug Discovery: Psychiatric Disorders, R.A. McArthur and F. Borsini, Editors. 2008, Academic Press: Elsevier: New York. p. 2-57.
Miczek, KA. Challenges for translational psychopharmacology research – the need for conceptual principles, in Animal and Translational Models for CNS Drug Discovery: Psychiatric Disorders, R.A. McArthur and F. Borsini, Editors. 2008, Academic Press: Elsevier: New York. p. 97-115.
van Os, J and Tamminga, C. Deconstructing psychosis. Schizophr Bull, 2007. 33(4): p. 861-2.
World Health Organisation, The ICD-10 classification of mental and behavioral disorder: diagnostic criteria for research. 1992, Geneva: World Health Organisation.

Expert Meeting 3

17 - 18 October 2008, Pueblo Acantilado, Alicante, ES

5.3. Expert Meeting 3

5.3.1. Executive Summary

The third expert meeting was intended to cover new methodological and technological developments in areas, which are prone to innovative approaches in 3Rs methodology. This is especially true for the cosmetics area where alternatives had to be introduced because of legislative changes accompanied by the stringent legal testing and marketing bans situation in the EU, and therefore, might have an inroad also to the pharmaceutical area of research and marketing. A number of cell systems, new and existing models, formulations and biopharma-related issues were discussed and in particular toxicological and regulatory implications with new methodology/technology were taken up and are summarised here.

- Promising use of cell culture systems:

• HepaRG cells seem to offer a new standard in cell systems of liver, i.e. the target organ for many toxicity effects of chemical substances, and therefore, of major use for the safety evaluation of chemicals in tests for acute, chronic and genotoxicity. The human HepaRG cell line expresses a unique set of liver functions including the major CYPs involved in drug metabolism and the key nuclear receptors, at least over 4 weeks, which is considered as long-term in an *in vitro* system.

Primary hepatocytes cultures stabilised by epigenetic modifiers such as histone deacetylase inhibitors could demonstrate usefull applications e.g. distinction between non-carcinogens and carcinogens, and should be further followed-up. A combination of both methodologies could be beneficial and represent a new development.

• In dermato-cosmetics as in pharmaceuticals, new assays are being developed, such as human keratinocyte models. In these, interesting aspects of microRNAs can be studied e.g. miRNAs as potential drug targets in skin diseases such as psoriasis, and of gene-silencing (siRNA (small interfering RNA) technology) enabling deletion of individual genes in a complex *in vitro* human cell system (less expensive, less time involved) instead of working with knock-out animals. In addition, several delivery systems for siRNA into the cytoplasm of the target cells seem interesting for further follow-up e.g. ultrasound combined with microbubbles, and polymers use.

- More attention for formulations / Biopharma-related issues:

• Nanoparticles exhibit a specific area of concern in safety testing in cosmetics and pharmaceutical formulations. Models such as precision cut lung slices might serve as a suitable tool to check the effects of nanoparticles on the respiratory tract, but there is also ample opportunity making use of nanoparticles, nanospheres and nanocapsules for topical and oral drug delivery and safety testing is a particular issue in nanomedicine in general.

The cell lines employed for checking into cytotoxicity often are (too) simple with limited functionality and further model development seems highly desirable.
• Biopharmaceuticals pose new problems to research and safety testing. Since they are complex molecules, they cannot fully be characterized like low molecular weight drugs. New, sometimes whole body tests (e.g. whole body glucose tests) have to be developed from the start, with a less invasive *in vivo* approach than standard pharmacological models of the past.
• Cancer models are a special target for further improvements, since by the development of biologicals for this topic, the area is more intensely covered (e.g. in 3 D culture models of human skin carcinoma). Also, further follow-up of specific *in vitro* biomarkers, which are also detected and relevant in the clinical situation, is necessary.

- Improvements in toxicological testing:
• In immunotoxicity testing, assays such as peptide reactivity or tests based on dendritic cells (e.g. in the lymphocyte transformation test, LTT) or monocyte-like cells are being investigated. Sensitisation testing in the LCSA (loose-fit co-culture-based sensitisation assay) might evolve, validated by murine local lymph node assay (LLNA) as the standard to test against and providing dose-response information.
In addition, differentiation of skin sensitizers by cell signalling pathways in murine and preferably human skin explants or reconstituted human skin models is identified as a promising target for further model development.

- Further development of new methodologies and technologies:
• More alternative methodologies, less invasive and more refined are being presented, in particular *in vivo* molecular imaging offers the possibility to non-invasively monitor health and disease of small laboratory animals and the effects of the drugs administered. Even a combination of different methods can be used *in vivo* in a non-invasive way.
Predicting organ biodistribution and targeting or e.g. preclinical cardiac imaging are to be given attention in this field. Also new *in vitro* applications using molecular imaging techniques were brought to the attention and need further development and follow-up.
• Sampling, even single individual cells or tissue parts can be performed by less interfering methodology such as the LCM (laser capture microdissection).
Methodology, already introduced into alternative methods, such as the *in vitro* Comet assay or *in vitro* flow based assays in blood for drug monitoring or embryonic stem cell based tests for tumour-induced angiogenesis, are currently substantially being improved.

- In a final discussion with regulator input and views presented, it became clear that trans-sector, cross sector - cutting information cannot be valued high enough – the best example given was the tremendous output of this expert workshop.

5.3.2. Recommendations

The workshop provided evidence for bottlenecks in the 3R methodological and technological area and provided some new solutions to the problems identified.

1. Cell systems: liver as a major target of toxicity of pharmaceuticals has to be analysed with most accurate and more appropriate cell lines ; the proposal is to work further on HepaRG cells, most likely with appropriate stabilisers (epigenetic modifiers such as histone deacetylase inhibitors) to further improve the *in vitro* model. In addition, as much user input as possible should be ensured to clearly define the applicability domain.

2. Models: new *in vitro* models for cancer/skin diseases research e.g. skin cancer, psoriasis as well as flow chamber systems (blood-based) in diverse areas of drug application/indications should be further encouraged.

3. Methods: imaging techniques *in vitro / in vivo*, such as SPECT (Single Positron Emission Computed Tomography) or Micro PET (Positron Emission Tomography) scanning, alone or preferentially several techniques combined, should be further improved and implemented in as well *in vivo* as in *in vitro* studies.

4. Formulation / Biopharma: new approaches to drug delivery such as the use of nanoparticles (including their toxicity testing orally, in skin and lung) have to be followed, especially under the aspects of more and more biologics in the global pharma pipelines. Other areas of nanoformulated drugs such as target respectively organ directed have to be anticipated.

5. Toxicology: already existing immunotox-/sensitising models and techniques (LTT, LLNA, resp. LCSA) should further be delineated with special emphasis on endpoints such as MAP (Mitogen Activated Protein) -kinases, which seem to exhibit discriminatory potential for sensitising and irritative compounds.

6. Regulatory: Communication on new models across sectors, involving the responsible regulatory agencies and competent authorities in the EU, should further be enhanced.

5.3.3. Report of Expert Meeting 3

START-UP Expert Meeting 3
17 - 18 October 2008, Pueblo Acantilado, Alicante, ES

New Avenues to research and development including safety testing of pharmaceuticals and cosmetics

Program:

17 October

08:30-09:00 Welcome and introduction. Vera Rogiers, *ecopa*, BE

Topic 1: New approaches in regard to research and safety testing of bio-pharmaceuticals

Pharmaceuticals

09:00-09:40 Functional Hepa RG cells: a breakthrough in 3R-research? André Guillouzo, Université de Rennes, FR

09:40-09:50 Hepatocyte *in vitro* systems for gene expression studies: comparison of 3 different culture systems. Markus Schug, University of Dortmund, DE

09:50-10:00 TSA-stabilised hepatocytes in the study of genotoxic and non-genotoxic carcinogens. Tatyana Doktorova, Vrije Universiteit Brussel, BE

10:00-10:10 Cytotoxic and genotoxic affects of two heterocyclic amines in functional hepatoma HepaRG cells. Julie Dumont, Université de Rennes, FR

10:10-10:25 Discussion with speakers

10:25-10:50 Coffee break

Dermato-Cosmetics

10:50-11:30 Markers of proliferation and differentiation in human keratinocyte models. Yves Poumay, University of Namur, BE

11:30-11:50 miRNA as new regulators in skin research. Andor Pivarcsi, Karolinska Institute, SE

11:50-12:10 Use of gene silencing in skin models. Michael Mildner, University of Vienna, AU

Nanoparticles

12:10-12:20 Safety aspects of skin penetration of nanoparties in cosmetics / pharmaceuticals. Alexa Patzelt, Charité University, Berlin, DE

12:20-12:40 Nanoparticle induced pulmonary toxicity. Susanne Boehn, BASF Laboratory of Alternative Methods, Ludwigshafen, DE

12:40-12:50 Toxicological studies of poly (anhydride) nanoparticles for oral drug delivery. Patricia Ojer Ojer, University of Navarra, ES

12:50-13:10 In vitro techniques and testing the ecotoxicity of nanoparticles. Richard D. Handy, University of Plymouth, UK

13:10-13:25 Discussion with speakers

13:00-16:00 LUNCH & free time

Topic 2: Immunotoxicology of pharmaceuticals and dermato-cosmetics

16:00-16:40 Important issues in immunotoxicity testing of pharmaceuticals. Marc Pallardy, University Paris Sud, FR

16:40-16:50 Development of *in vitro* strategies for studying T-cell mediated imunological reactions to drug. Enrique Gómez, University of Málaga, ES

16:50-17:05 New ideas in basic research in allergic skin diseases. Reinhard Wanner, Institute of Molecular Biology and Bioinformatics, DE

17:05-17:20 Reactivity studies to complement the development of alternative methods for the prediction of skin sensitisation: the case of formaldehyde and formaldehyde releasers. Mustapha Kireche, Clinique Dermatologique CHU, FR

17:20-17:30 In vitro differentiation of skin sensitisers by cell signaling pathways. Lydia-Mareen Koeper, Bayer HealthCare AG, DE

17:30-17:40 Identification of cell responses induced by a sensitizer in a reconstructed human epidermis. Aurélie Frankart, University of Namur, BE

17:40-18:00 Discussion with speakers

18 October

Topic 1: New approaches in regard to research and safety testing of bio-pharmaceuticals and cosmetics
Biotech products

09:00-09:40 Biotech - What's in the pipeline? Daan Crommelin, University of Utrecht, NL

09:40-09:50 Prednisolone-treatment induces hepatic insulin resistance in C57Bl/6J mice measured by a newly developed whole body glucose test. Anke Laskewitz, University Medical Center Groningen, NL

09:50-10:10 A suitable delivery system is key in the concept of siRNA -medicated treatment. Roosmarijn Vandenbroucke, University of Ghent, BE

10:10-10:20 NN

10:20-10:35 Discussion with speakers

10:35-11:00 Coffee break

Topic 3: Research / Development and 3R-Alternatives

11:00-11:40 Imaging biomarkers *in vivo / in vitro*. Tony Lahoutte, Vrije University Brussel, BE

11:40-12:00 Non-contact laser capture microdissection: an alternative method in live cell sampling. Renate Bugemeister, Carl Zeiss Microimaging, Bernried, DE

12:00-12:20 COMICS: Developing high throughput comet assays for DNA damage and DNA repair. Amaya Azqueta Oscoz, University of Oslo, NO

12:20-12:45 Discussion with speakers

12:45-16:00 LUNCH & Free time

Topic 4: Research / Efficacy Testing / 3R-Alternatives

16:00-16:40 Use of flow-based assays to monitor drugs with antithrombotic, anticoagulant or anti-inflamatory potential. Johan W.M. Heemskerk, University of Maastricht, NL

16:40-16:50 Confrontation cultures of embryonic stem cell - derived embryoid bodies and multicellular tumour spheroids: a new in vitro model for the study of tumour-induced angiogenesis. Nada Milosevic, Justus Liebig University, DE

16:50-17:00 Cancer associated fibroblasts influence epidermical morphogenesis and invasion in reconstructed 3D-culture models of human carcinoma. Susan Commandeur, University of Leiden, NL

17:00-17:15 Discussion with speakers

17:15-17:45 Coffee break

Topic 5: Miscellaneous

17:45-17:55 Evaluation of toxicological effects of POPs using primary fish cell cultures. Liv SØfteland, The National Institute of Nutrition and Seafood Research, NO

17:55-18:05 Applying the free concentration concept to *in vitro* assays. Nynke Kramer, University of Utrecht, NL

18:05-18:15 A cell-based assay for the detection of a protein algal neurotoxin in biological matrices. Mirella Belloci, University of Modena, IT

18:15-19:00 From successful research to regulatory implementation: hurdles for alternative methods and the validation necessity. Sonja Beken, Federal Agency for Medicines and Health Products, BE
Discussion with speakers and Round Table

Scientific committee:
Bernward Garthoff, treasurer *ecopa*, Bayer, DE
Vera Rogiers, chair *ecopa*, VUB, BE
José Castell, co-chair *ecopa*, Hospital Universitario La Fe, ES

5.3.3.1. Introduction to the report

This third Expert Meeting is one of the 3 pre-meetings running prior to 3 Workshops of which each addresses one of the 3Rs, being Refinement, Reduction and Replacement. These Workshops constitute the core of the START-UP project.
This pre-meeting is in particular concerned with methodological and technological bottlenecks and new developments in the 3Rs area. The meeting was deliberately combined with the two-yearly eSI (*ecopa* Science Initiative) workshop. Indeed, this *ecopa* initiative aims at bringing young and promising scientists working in the field of 3Rs and coming from all over Europe in contact with well-established "senior" scientists. It is a unique opportunity for young scientists to present and discuss their own results and to be challenged by experienced experts in the field. A delegation of EPAA (European Partnership for Alternative Approaches to Animal Testing) was present in order to let the young people think about the applicability of the methods and technologies proposed.

5.3.3.2. Introduction to the Expert Meeting

A short introduction was given by the chair of *ecopa*, Vera Rogiers, about the nature and goals of the START-UP project and the way it was structured with a series of expert pre-meetings to gather all relevant information possible with respect to alternative methods. The most important ideas will then be "loaded" in the 3R Workshops and will, together with the input given by the local organisers (National Consensus Platforms, NCP), constitute the main discussion points of the 3R Workshops to be organised later on in 2009. Further instructions about the development of the meeting were given by José Castell and Bernward Garthoff as vice-chair and treasurer of *ecopa* respectively.

5.3.3.3. Summary of the presentations within the different sessions

5.3.3.3.1. Session I: New approaches in regard to research and safety testing of bio-pharmaceuticals respectively cosmetics - Pharmaceuticals

5.3.3.3.1.1. Functional Hepa RG cells: a breakthrough in 3R-research?, by André Guillouzo, Université de Rennes, FR

Several hundreds of chemicals are hepatotoxic, amongst which a substantial amount of newly developed drugs. A problem here is that about one third of the drugs are found to be hepatotoxic in clinical trials, where this was not the case in animal models.

Since the liver is a target organ for drugs and other chemicals and plays a major role in bioactivation, *in vitro* human liver models are needed to mimic the human situation, as it is clear that this is not always possible with animals.
For the evaluation of chemical toxicity there are a number of *in vitro* liver models, amongst which tissue slices and primary hepatocyte cultures, cell-based and reporter gene assays, hepatocytes derived from stem cells and liver cell lines. All of these models show some positive, but also some negative properties. For example, the availability of tissue slices and primary hepatocytes is limited, there is an important interdonor variability, early phenotypic changes occur and the life-span of both models is limited. Hepatocyte cell lines obtained either by oncogenic immortalisation or from liver tumours, have lost most, if not all, major cytochrome P450 (CYP) activities. These models are usable for metabolic studies, but not good enough for toxicity studies.
However, the HepaRG cell line derived from a human hepatocarcinoma is an exception and exhibits morphological and functional properties close to mature human hepatocytes. The HepaRG cell line has a subnormal, subdiploïde female karyotype.
Adding DMSO (Dimethylsulfoxide) to the culture is important, because it induces maximum morphological differentiation, without DMSO the activity remains low.
Stable long-term expression of several liver transcripts (such as: CYPs and membrane transporters, phase II enzymes such as GSTA1 (glutathione s-transferase), UGT1A1 (uridyl difosfoglucuronly transferase) and mEH (microsomal epoxide hydrase), antioxidant enzymes including catalase and MnSOD (Manganese Superoxide Dismutase) and nuclear receptors (CAR (constitutive androstane receptor**)**, PXR **(**pregnane X receptor) is kept when the cells are seeded at low and at high densities. The responsiveness to inducers is also well maintained.
HepaRG cells can be used for the evaluation of chemicals in tests for acute and chronic toxicity and genotoxicity. Experiments with Aflatoxin B1 show that human hepatocytes and HepaRG cells have a comparable sensitivity, where this is not the case with HepG2 cells.
It is possible to prepare recombinant HepaRG cell lines, which can be useful for testing for oxidative stress.
The number of genes expressed in primary human hepatocytes and HepaRG cells is comparable. The expression patterns of the genes also is similar.
To conclude it was mentioned that the human HepaRG cell line expresses a unique set of liver functions including the major CYPs involved in drug metabolism and the key nuclear receptors (CAR, PXR, AhR (Aryl Hydrocarbon Receptor), PPARs (Peroxisome Proliferator Activated Receptor)). This cell line possesses both the functional capacities of adult hepatocytes in primary culture and the indefinite growth capacity of hepatoma cells and it can be used for various applications including hepatitis B infection. It must, however, be kept in mind that these cells are transformed and derived from a single individual.

5.3.3.3.1.2. Hepatocyte in vitro gene expression systems: what is the reason for differences between in vivo and in vitro systems, by Markus Schug, University of Dortmund (IfaDo), DE

An answer was provided on the question whether the gene expression alterations observed *in vitro* are also relevant for the *in vivo* situation.
Methapyrilene was taken as an example. The most suitable culture system was determined and was found to be collagen sandwich cultures.
Concentrations of 0.002, 0.02, 0.39, 6.25 and 100 µM methapyrilene were used with an incubation of 24h. It was concluded that alterations in gene expression were observed at *in vivo* relevant concentrations.
Furthermore, *in vitro* observed alterations in gene expression were also relevant for the *in vivo* situation. It was suggested that the reason for discrepancies often seen for *in vivo* and *in vitro* data could be the different pharmacokinetics seen *in vitro* and *in vivo*.

5.3.3.3.1.3. Non-genotoxic hepatocarcinogens specifically target mitochondria and gap junctions, by Tatyana Doktorova, Vrije Universiteit Brussel, BE

A major shortcoming of current *in vitro* carcinogenicity testing is the lack of clear-cut biomarkers to discriminate between genotoxic and non-genotoxic carcinogens. In this preliminary study, it was investigated whether mitochondria and gap junctions could be used to distinguish between both types of carcinogens. Primary rat hepatocytes functionally stabilised by treatment with the histone deacetylase inhibitor Trichostatin A and non-treated cells were exposed to the prototypical genotoxic hepatocarcinogens 2-nitrofluorene and benzo[a]pyrene, and the non-genotoxic methapyrilene hydrochloride and sodium phenobarbital. Mitochondrial function was assessed by means of a MTT test. Gap junction integrity was evaluated by measuring the protein levels of connexin32. Both methapyrilene hydrochloride and sodium phenobarbital decreased mitochondrial activity in a dose-dependent way. There effects became evident at lower concentrations in Trichostatin A-treated cells compared to non-treated ones. By contrast, hepatocellular mitochondria were left unaffected in the presence of 2-nitrofluorene and benzo[a]pyrene. Methapyrilene hydrochloride downregulated connexin32 expression, whereas benzo[a]pyrene did not. The results of this preliminary study suggest that abrogation of mitochondrial homeostasis and interruption of gap junction formation can be, besides other markers, used to distinguish genotoxic hepatocarcinogens from their non-genotoxic counterparts. In addition, Trichostatin A-stabilised hepatocytes seem to represent a sensitive system for assessing the non-genotoxic properties of hepatocarcinogens.

5.3.3.3.1.4. Cytotoxic and genotoxic effects of two heterocyclic amines in functional human hepatoma HepaRG cells, by Julie Dumont, Université de Rennes, FR

Two heterocyclic amines (HCAs), mainly produced by cooking meat or fish, but also detected in cigarette smoke, are mutagenic chemicals and carcinogenic at multiple sites in experimental animals.
Their chronic hepatic toxicity and genotoxicity, individually and in combination, was studied in HepaRG cells. These retain their specific functional capacities for at least 4 weeks at confluence, including CYP1A2 activity which is involved in HCAs metabolism.
Both compounds triggered no cytotoxicity in HepaRG cells. This was measured by intracellular ATP (adenosine triphosphate) content and neutral red incorporation measurements. No morphological changes were observed, nor increased caspase-3 activity.
Genotoxicity was tested via a Comet assay. Only one of the HCAs was positive and induced DNA damage. The genotoxicity detected could not be explained by an increase in cytotoxicity or in apoptosis.

5.3.3.3.2. Session II: New approaches in regard to research and safety testing of bio-pharmaceuticals respectively cosmetics - Dermato-Cosmetics

5.3.3.3.2.1. Markers of proliferation and differentiation in human keratinocyte models, by Yves Poumay, University of Namur, BE

Human epidermal keratinocytes were discussed with respect to their utilization as monolayer cultures and their utilization to reconstruct an epidermis with all typical basal, spinous, granular and cornified layers at the air-liquid interface, with special emphasis on the expression of proliferation and differentiation markers.
Monolayer cultures are preferably analysed in autocrine culture conditions where no peptide (e.g. EGF) is added to the culture medium. These conditions allow detection of early (keratin 10) as well as later (involucrin) markers of epidermal differentiation.
The epidermal differentiation of these keratinocyte monolayers depends on a number of factors, such as the choice of the house-keeping genes, the effect of cell density, the effect of cholesterol depletion and the presence of retinoids. The conclusion is that monolayers are easy to produce, can proliferate up to confluence without addition of growth factors, house-keeping genes must be tested in experimental conditions and detection of markers needs to be done under relevant conditions. Monolayers are adequate *in vitro* models for the study of alterations (stress, pharmacology) of epidermal differentiation.
When Reconstructed Human Epidermis (RHE) is kept under serum-free, collagen-free conditions, it is a simple model, which is adequate for easy interpretation of data regarding cell release. This model can be used for tissue response, for example in order to evaluate the release of interleukins, such as IL-1a and IL-8, during its response to irritants or sensitizers. Data from keratinocyte monolayers can be transposed to RHE: the role of

signalling intermediates and the activation of particular (stress-response) signalling pathways are under investigation.
For the time being, the RHE model is more difficult and more expensive to produce than the monolayer, it requires addition of growth factors to the culture medium. But it is available for studies of markers of proliferation and differentiation, cellular release and signalling.
In the future this model will, however, be produced by different manufacturers that will bring more information about the model and allow refinements in the analysis of the tissue.

5.3.3.3.2.2. MicroRNAs: new regulators in skin research, Andor Pivarcsi, Karolinska Institute, SE

MicroRNAs, discovered in 1993, are short, non-protein-coding RNAs, of approximately 22 nucleotides, which regulate the expression of protein-coding genes. They inhibit gene expression at the post-transcriptional level and thus their function and importance can be compared to those of transcription factors. They regulate basic biological processes such as apoptosis, morphogenesis, proliferation, metabolism…
MiRNA is encoded in the genome, transcribed from the DNA as pri-microRNA, then processed into pre-microRNA in the cell nucleus. In the cytoplasm it is processed into mature microRNA, this forms complexes with the 3 ´ UTR (untranslated region) of target mRNA and suppresses their translation or induces their degradation.
The emergence of miRNAs will not make the understanding of regulatory networks easier, because each miRNA regulates more than one gene, while at the same time each gene is regulated by more than one miRNA. Many miRNA targets are transcription factors, while miRNA expression is also regulated by transcription factors. Therefore the potential regulatory effect of miRNAs is enormous.
Abnormal miRNA expression can cause diseases, such as cancer and developmental and metabolic diseases.
Their importance for skin e.g. their role in psoriasis was investigated.
Genome-wide analysis of miRNA expression using an array with LNA (locked nucleic acid**)** probes showed that a set of microRNAs is expressed in human skin and that healthy skin, atopic eczema lesions and psoriatic lesions display distinct microRNA expression profiles.
It was shown that miR-203 is a keratinocyte-specific miRNA and that it is overexpressed in psoriasis. The up-regulation of miR-203 in psoriasis is concurrent with the down-regulation of its target, SOCS-3, which may result in an increased inflammatory response.
Due to their role in skin diseases, miRNAs are potential drug targets.

5.3.3.3.2.3. Use of gene silencing in skin models, Michael Mildner, Medical University Vienna, AT

Gene silencing in organotypic skin models is interesting, because it enables the study of gene deletions in a complex *in vitro* system of human cells and strongly reduces the necessity of animal experiments in dermatological research.
In this model system primary human keratinocytes are grown on a support consisting of collagen and fibroblasts at the air liquid interface for several days. During this time keratinocytes start to form a multilayered, well differentiated epithelium with a stratum corneum- like surface.
To investigate the reliability of short interfering RNA (siRNA) mediated gene knock down in this model, the phenotype of 2 genes (VEGF, matriptase-1) strongly expressed in keratinocytes, was compared to that of the corresponding knock out animals. It was demonstrate that the use of siRNA technology in this model is able to reproduce functional changes comparable to those seen in corresponding knock out animals. The genes remained silenced up to 8 days.
This method was also applied to DNase 1-like 2 (DNase 1L2), which is of particular interest for keratinocyte differentiation and for which no animal knock out data are available. The results showed that DNAse 1L2 is specifically expressed in the epidermis and that this expression correlates with keratinocyte differentiation. Knockdown of DNase1L2 results in parakeratosis. A specific knockdown of DNAse1L2 by si RNA technology in skin epidermis culture results in retention of nuclei in corneocytes
This model enables the study of deletion of individual genes in a complex system of human cells and it is less expensive and much less time consuming than animal models
The model could be improved by generating organotypic skin cultures containing other cell-types of the skin.

5.3.3.3.3. Session III: New approaches in regard to research and safety testing of bio-pharmaceuticals respectively cosmetics - Nanoparticles

5.3.3.3.3.1. Safety aspects of skin penetration of nanoparticles in topically applied cosmetics / pharmaceuticals, by Alexa Patzelt, Charité University, DE

The main question in this research is whether nanoparticles remain on the skin or whether they can pass the skin barrier and gain access to systemic compartments.
There are three penetration pathways: intra- and intercellular and follicular.
By using the method of tape stripping, it was found that titanium dioxide does not penetrate into the living tissue via the intercellular pathway.
However the results showed that nanoparticles of a certain size (most probably between 400 and 700nm) can penetrate into the hair follicle and remain there for some days. This makes nanoparticles promising for drug delivery systems for non-particulate active agents.

Most probably nanoparticles can also penetrate into damaged skin. Until now, there is no evidence that nanoparticles can penetrate through intact skin. Nevertheless additional investigations have to be performed for nanoparticles smaller than 10nm.

5.3.3.3.3.2. Nanoparticle induced pulmonary toxicity, by Susanne Boehn, BASF, DE

The production and use of engineered nanoparticles has dramatically increased due to their various applications such as in medicine, information technology, and chemistry. They are especially interesting because of their high surface to volume ratio.
A lot of research to possible toxicological effects still has to be done and exposure via inhalation is of the highest concern. Since there is no *in vitro* model available yet, *in vivo* inhalation studies are required for the investigation of nanoparticle induced effects on the respiratory tract. These experiments are resource intensive and, especially for nanoparticulate matter, technically challenging. Therefore research is being done to develop, establish, and validate precision cut lung slices (PCLs) as an *in vitro* model or alternative method for lung toxicity of nanoparticles. PCLs may be employed for grouping approaches in which a limited number of nanoparticles belonging to a group with specific physico-chemical properties are tested *in vivo* and the toxic properties for others are extrapolated from the in vitro methods. The aim of the study was to investigate the effects of different nanoparticles for their toxic potential. The effects of titanium dioxide, cobalt nanoparticle aggregates, and cobalt ferrites at different sizes with and without serum protein stabilisation were studied. The effects on mitochondrial metabolic activity (WST-1 assay), membrane integrity (LDH (Lactate Dehydrogenase) assay), total protein content (BCA assay) and IL1α production (ELISA (Enzyme Linked ImmunoSorbent Assay)) were assessed. The possibility to extrapolate from physiochemical properties of nanoparticulate material to toxic effects was assessed. Furthermore, DNA single and double strand breaks and alkali-labile DNA modifications induced by titanium dioxide were investigated using the Comet assay. This way it was demonstrated that PCLs as *in vitro* system are a suitable tool to investigate the effects of nanoparticles on the respiratory tract.

5.3.3.3.3.3. Toxicological studies of poly (anhydride) nanoparticles for oral drug delivery, by Patricia Ojer Ojer, University of Navarra, ES

“Nanomedicine” is a term which has been applied to diverse medical fields such as drug delivery systems. Nanoparticles can be divided into nanospheres and nanocapsules. The nanospheres contain the drug uniformly dispersed in a matrix and the nanoparticles contain the drug encapsulated inside. Using nanoparticles it is possible to decrease the administration dose and the drug degradation because of the protection that the nanoparticles give. It is thus possible to decrease adverse reactions and to increase efficacy. All these reasons make it possible to apply nanoparticles in oral drug delivery.

The main problem with this type of technology is that little is known about their potential toxic effects because there are no specific guidelines for assessing the toxicity of these new nanomaterials.
The polymer chosen to prepare the nanoparticles was Gantrez® AN, which is a copolymer of methyl vinyl ether and maleic anhydride. To evaluate the toxicological effects of Gantrez nanoparticles, different types of nanoparticles were prepared: conventional nanoparticles that increase the bioavailability of the presystemic drug metabolism, pegylated nanoparticles that decrease the interaction of nanoparticles with components of the lumen and cyclodextrin nanoparticles that increase the carrying capacity of the nanoparticles in the case of lipophilic drugs.
The physico-chemical characteristics (size, surface charge, shape and stability) of the prepared nanoparticles were determined. The nanoparticle formulations that have been studied displayed a size of approximately 190 nm, a negative surface charge due to the carboxylic groups of Gantrez® AN and a spherical morphology with a homogeneous size distribution.
In order to evaluate the toxicity of Gantrez® AN nanoparticles, *in vitro* assays were performed using the Hep G2 cell line derived from human liver adenocarcinoma. Cytotoxicity was first evaluated by visual inspection of the cells, under the microscope and by using the MTS (Soluble Tetrazolium Salt) assay. The nanoparticles were incubated for 30 minutes, 3, 24, 48 and 72 hours at different concentrations (0.0625, 0.125, 0.25, 0.5, 1 and 2 mg/mL). Then the MTS assay was performed and the survival percentage of treated cells with respect to not treated cells was calculated. The nanoparticles were non cytotoxic at the concentrations tested. In contrast, at 48 and 72 hours significant cytotoxic effects were observed for 1 and 2 mg/mL nanoparticles concentrations.

5.3.3.3.4. Session IV: Immunotoxicology of Pharmaceuticals / Dermato-cosmetics

5.3.3.3.4.1. Important issues in immunotoxicity testing of pharmaceuticals, by Marc Pallardy, INSERM, FR

Immune responses can be classified roughly as innate and acquired responses.
Immunotoxic effects are best divided into four categories: immunosuppression, immunostimulation, hypersensitivity and autoimmunity, as each category is associated with relatively specific and clinically distinct adverse events.
Due to new legislation, there is a need for more *in vitro* testing, therefore tests are being developed for *in vitro* testing of immunosuppressive molecules, based on the functional evaluation of the different components of the immune system.
In the case of chemical sensitizers detection several important results have been obtained. Based on the assumption that one of the features of a chemical-allergen is chemical reactivity, allowing binding to proteins, a peptide reactivity assay has been developed. Recent publications on this test show an 82% concordance for the detection of chemical sensitizers. Cell-based assays are also under extensive development, these tests are based

on dendritic cell or monocytic-like cell models. Dendritic cells present in peripheral organs such as the skin or the lung are capable to take up and to process allergens. Upon antigen capture, stimulation by Toll-like receptor (TLR) agonists, but also in response to inflammatory cytokines, the dendritic cells undergo a maturation process leading to the upregulation of co-stimulatory molecules (CD86, CD80, CD40), MHC (Major Histocompatibility Complex) class II molecules, the CD83 protein and cytokine production. Thus, in addition of antigen processing, dendritic cells need to receive signals through the so-called "danger signals" mainly composed of pathogen-associated molecular patterns (PAMPs) recognised by TLRs. A current hypothesis is that considering similarities between immunity to simple chemicals and that to infectious agents, it is reasonable to speculate that the chemical sensitizer itself stimulates dendritic cell maturation and that chemical sensitizers could be perceived by dendritic cells as "danger signals" with common signalling pathways leading to maturation and migration. All the cell-based assays currently under development are exploiting the above concept and prevalidation studies are currently ongoing.

5.3.3.3.4.2. Development of in vitro strategies for studying T-cell mediated immunological reactions to drugs, by Enrique Gomez, University of Malaga, ES

Part of all adverse drug reactions are produced by an altered immunologic response to the drug and are named allergic drug reactions with an immunological basis. They can be IgE mediated or T cell mediated. Immediate reactions, those occurring in less than 1 hour after the intake drug, are mediated by IgE. Non-immediate reactions, occurring more than 1 hour after the intake of the drug, are T cell mediated.
In T-cell mediated reactions, the hapten is recognized by the Th1 Response, this starts changes in the maturation state of dendritic cells. These cells migrate to the lymph node and present, the hapten, now in a mature state, to naïve T-Lymphocytes. The contact between mature dendritic and T-cells, together with adequate co-stimulatory signals, provoke a release of IL12 from T-Lymphocyte and a proliferation process of the specific T cell clones. These clones can become either Memory T cells or cytotoxic cells can develop the effector function. They recruit other effector cells, such as macrophages, to develop the effector action.
In vitro testing has some advantages to *in vivo* testing, for example that it bears no risk for the patients. However, the main problem of the *in vitro* tests is the low sensitivity in comparison to *in vivo* tests.
Non immediate reactions can be monitored and different genetic markers or proteins can be measured. Likewise one can also determine the different sub-populations. These data give important information about the cellular trafficking occurring during the reactions. To evaluate the non immediate reactions, common techniques can be used such as: immunohistochemistry, real time PCR (Polymerase Chain Reaction), flow cytometry and cell culture.
Another alternative method to evaluate non immediate reactions is the lymphocyte transformation or proliferation test. This method is based on the lymphocyte capacity to proliferate in the presence of the specific hapten or antigen.

In the last years, flow cytometric methodology has been introduced in LTT (Lymphocyte Transformation Test). As such information can be obtained, not only on the specific proliferation, but also on the proliferating cell sub-populations.
The LTT has demonstrated to be useful in various diseases and with many different drugs, but it can be improved with the use of dendritic cells as antigen presenting cells. The application of flow cytometric methods will help to identify the cell subpopulation involved in the drug recognition.

5.3.3.3.4.3. In vitro quantitation of sensitising potential, Reinhard Wanner, Charité University, DE

Allergic reaction development is controlled by dendritic cells, which are professional antigen-presenting cells. For the induction of an immune response, they sense potentially harmful environmental substances. Contact with these “danger signals” is then translated and integrated into signalling pathways which induce maturation. Antigens are sampled and processed for presentation to T-Lymphocytes. The function of dendritic cells is communication with their cellular environment. Keratinocytes may provide danger signal-induced cytokines or chemokines. An *in vitro* assay to predict sensitising potential should therefore contain both, keratinocytes and dendritic cells.
In vitro 3D models of skin, including keratinocytes, fibroblasts and melanocytes are available. The problem, however, is that dendritic cells do not steadily integrate into 3D-cultures. In the *in vivo* situation dendritic cells are also not fixed steadily to the keratinocytes. A more natural model exists of a loose-fit co culture of keratinocytes and dendritic cells.
The application of exogenous cytokines induced concomitant keratinocyte activation and differentiation of monocytes to dendritic cells. The coculture developed to an allergen-sensitive system. Sensitisation could be determined by analysing the expression of the dendritic cell maturation marker CD86. Estimation of the concentration required to cause a half-maximal increase in CD86-expression allowed quantitative risk assessment (QRA) for chemicals. The presence of keratinocytes increased the vitality and sensitivity to allergens of the cocultured dendritic cell-related cells as compared to solely cultured dendritic cells. This allowed testing at concentration ranges without general toxicity, allowing discrimination between allergic and irritant potential. The new assay was entitled LCSA (loose-fit coculture-based sensitisation assay).
The murine local lymph node assay (LLNA) is a validated sensitisation assay. LLNA-results can be used as a standard during development and validation of a new assay. Various substances, grouped by the LLNA into potency categories, were tested. EC3-values of the LLNA have good correlations with concentrations of half maximal CD86-upregulation in the LCSA. LCSA and LLNA achieve an analogous grouping of allergens into categories like weak, moderate and strong.
The LCSA provides dose-response information, thereby allowing prediction of the relative ability of a substance to induce sensitisation. This new model may have the capacity to replace the currently used animal-based assay for sensitisation.

5.3.3.3.4.4. Reactivity studies to complement the development of alternative methods for the prediction of skin sensitisation: the case of formaldehyde and formaldehyde releasers, by Mustapha Kireche, Université Louis Pasteur, FR

Allergic diseases are gaining importance because more and more chemicals are used. Since animal testing must be reduced, due to European recent legislations, there is a need for alternative methods to predict the allergenic potential of chemicals. Studies on the reactivity of allergenic compounds with nucleophilic amino acids on proteins, which is the key step to form the antigenic structure that will start the immunological process characterizing skin contact allergy, are essential for the development of new alternative methods. A French network aims at developing an *in vitro* method for the evaluation of the allergenic potential of chemicals.
In contact dermatitis there are two associated immunological steps: (i) the sensitisation phase leads to the formation of a hapten-protein complex, antigen formation and to the production of specific T cells. (ii) The elicitation phase also leads to the formation of a hapten-protein complex, the activation of the T cells and to an allergic reaction.
Formaldehyde is often used as a biocide, in textiles, painting, cosmetic products... However, it is a strong skin sensitizer and it has also been classified as carcinogenic. Formaldehyde releasers are used to control/decrease the free formaldehyde concentration in the product, thus reducing the toxicity of the preservative. It has, however, been observed, however, that persons non‐allergic to formaldehyde itself can be sensitised to formaldehyde releasers.
The reactivity of formaldehyde releasers towards nucleophilic amino acids was studied in order to detect whether adducts were formed by reaction of the released formaldehyde or by reaction of the releaser itself. Therefore the ^{13}C formaldehyde releasers were labelled at the reactive sites and their reactivity toward nucleophilic *N*‐acetylated amino acids was followed and analysed by ^{13}C mono and two‐dimensional nuclear magnetic resonance (NMR). The results obtained were compared with the reactivity of single formaldehyde with the amino acids.
The results showed that different adducts corresponding to direct reactions of the formaldehyde releasers on the amino acids were obtained, either by reaction of the intact molecule or by reaction of an original breakdown product of the molecule.
These results could be at the origin of the explanation of why some people are allergic to formaldehyde releasers and not to the released formaldehyde.

5.3.3.3.4.5. In vitro differentiation of skin sensitizers by cell signalling pathways, by Lydia-Mareen Koeper (Bayer HealthCare AG, DE

The focus of this study was to investigate whether analyses of cell signalling pathways can provide a methodology for the detection of sensitising compounds *in vitro*. For this purpose a differentiation between non-specific immune reactions (skin irritation) and skin sensitisation was of major importance.

The models used were murine and human skin explants as well as reconstituted skin models (EST-1000: Epidermal Skin Test, consists of keratinocytes; AST-2000: Advanced Skin Test, full-thickness skin model, consists of keratinocytes and fibroblasts). These models were exposed to different concentrations of sensitising compounds, such as Oxazolone and DNFB (2,4-dinitro-1-fluorobenzene), or to irritant compounds, such as SDS (sodium dodecyl sulphate) and TritonX-100. Both irritant and sensitising compounds have irritant properties leading to a dose-dependent decrease in viability. The lowest observed effect level (LOEL), defined as the concentration resulting in a decrease in viability of about 10% after 24h of exposure, was determined for each compound and skin model by using the MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphemiltetrazoliumbromide) viability assay. Each skin model was then exposed to the appropriate concentrations of the LOEL for 1h or 3h. Phosphorylation of MAP-kinases (p38, ERK1/2 and JNK1/2), STAT1 and PLCγ were determined by cytometric bead array (CBA).
In the murine and human skin explants, all three MAP-kinases (Mitogen Activated Protein) were exclusively activated after exposure to sensitising compounds. For the reconstituted skin models, phosphorylations of p38 and JNK1/2 were obtained after stimulation with allergens, whereas treatments with irritant compounds lead to ERK1/2 activation.
The conclusion of this study is that MAP-kinase activation was shown to provide a promising *in vitro* tool for the discrimination between sensitising and irritant compounds. Skin explants seem to have the best capability for identifying sensitizers since complex interactions leading to an activation of all three MAP-kinases can be measured. The disadvantages of skin explants are high inter- and intra-individual variabilities and, for human explants, a limited availability. Especially the reconstituted skin model EST-1000 showed high induction levels of phospho-p38. The inductions were comparable to those found in skin explants and specific for an exposure to sensitising compounds. With respect to availability, variability and simplicity in handling, the EST-1000 turned out to be the model of choice for further analyses of compounds.

5.3.3.3.4.6. Identification of cell responses induced by a sensitizer in a reconstructed human epidermis, by Aurélie Frankart, University of Namur, BE

With preliminary results obtained with one sensitizer (DNFB: 2, 4-dinitro-1-fluorobenzene) and one irritant (BC: benzalkonium chloride), it is confirmed that a discrimination between irritants and sensitizers is possible using a profile of IL-1α and IL-8 release from Reconstructed Human Epidermis (RHE). Indeed, acquired data show that DNFB induced only an IL-8 release, whereas BC induced release of the two analysed cytokines. In this study several cellular mechanisms, possibly responsible for the differences in IL-1 α and IL-8 release, were investigated.
In RHE, H_2O_2 in contact with the basal layer induces the release of IL-8 in extracellular medium. The hypothesis then is that H_2O_2 could be an intermediate in the response of RHE towards DNFB. To study this RHE were co-treated with DNFB/ H_2O_2 and an

antioxidant agent N-acetylcysteine (NAC). It was shown that NAC reduces the IL-8 release by RHE treated with H_2O_2, so it seems indeed to be an intermediate.
H_2O_2 is also an important intracellular messenger, which regulates protein dephosphorylation of tyrosine residues
To understand which mechanisms could be involved in cytokines release by RHE treated with DNFB, cell signalling pathways were analysed on RHE treated with sensitizer DNFB. A time-dependent phosphorylation of EGFR (Epidermal Growth Factor Receptor) and of ERK ½ MAPK was obtained after stimulation with DNFB, suggesting their possible involvement in the IL-8 release produced by this compound.
RHE was incubated with inhibitors of the kinase activities of EGFR and ERK (Extracellular Signal-Regulated Kinases) ½ MAPK (Mitogen Activated Protein Kinase). Both inhibitors reduce IL-8 release in the DNFB treated RHE.
The results of this study are preliminary results and the experiments still have to be repeated for better statistical results. Studies involving more sensitizers and with irritants, to compare with sensitizers, still have to be performed. However, analysis of MAP-kinase activation seems to provide a promising tool for the identification of sensitising compounds *in vitro*.

5.3.3.3.5. Session V: New approaches in regard to research and safety testing of bio-pharmaceuticals resp. cosmetics - Biotech products

5.3.3.3.5.1. About Pipelines and Shifting paradigms: biopharmaceuticals versus low molecular weight drugs, by Daan Crommelin, TI Pharma, NL

Biopharmaceuticals are pharmaceutical products consisting of (glyco)proteins. The first biotech therapy to earn FDA approval was recombinant human in 1982. Nowadays, a substantial number of the FDA approved drugs are biopharmaceuticals.
Later, endogenous biopharmaceuticals were brought on the market, such as interferons, interleukins, different growth factors and monoclonal antibodies. The clinical and economic importance of the group of biopharmaceuticals is illustrated by the fact that over 60 billion Euros are annually spent on biopharmaceuticals.
Biopharmaceuticals deserve attention as they have a number of special characteristics. Their activity depends on their complicated geometry based on secondary, tertiary and quaternary structures. The way biopharmaceuticals are handled is very important, since the conformational structure of a protein is easily disturbed. A correct structure is important in order to optimise the therapeutic effect and to minimize adverse reactions, among which immune responses.
These structures can not be fully defined with our present set of analytical techniques and approaches for potency testing, because they often closely resemble endogenous proteins. This means that in safety and clinical tests, much attention has to be given to species specific responses (cf. Tegenero case), correction for base levels, selection of dosing schedules and the possibility of immunogenicity. Immunogenicity can be tested by doing

sequence analysis, reactivity tests with antibodies and animal studies on conventional animals, non-human primates and transgenic immune tolerant mice.
At present there are five expression technologies for protein production. They can be derived from bacteria (*Escherichia coli*), yeast, mammalian cells, transgenic animals and transgenic plants.
Nearly all biopharmaceuticals are injected, but with the growing importance of these drugs, the need for alternative methods of administration increased as well. Alternative routes were studied for a long time, but at present injection is still the method mostly used.
In conclusion, biopharmaceuticals are complex molecules that can not be characterized fully in terms of their structure like low molecular weight drugs. The performance of biopharmaceuticals relies on strict production protocols and close monitoring of their activity in the clinical situation.

5.3.3.3.5.2. Prednisolone-treatment induces hepatic insulin resistance in C57Bl/6J mice, measured by a newly developed whole body glucose test, by Anke Laskewitz, University medical center Groningen, NL

Glucocorticoids are hormones produced by the adrenal gland under control of the hypothalamic-pituiary-adrenal axis. They play an endogenous role in salt and water metabolism, blood pressure, glucose metabolism and the immune system. Synthetic glucocorticoids, such as prednisolone, are widely used as immunosuppressive drugs in chronic inflammatory diseases. However, chronic use of glucocorticoids leads to severe side effects, such as hypertension, skin thinning, impaired immune response, weight gain, osteoporosis, and insulin resistance. The goal of this study is to identify the underlying mechanisms of glucocorticoid-induced insulin resistance. The organs which are most probably involved with insulin resistance are the muscles, brains, liver, pancreas and adipose tissue. Here the focus was on the liver and peripheral tissues. To this end glucocorticoid-induced insulin resistance was measured in C57Bl/6J mice fed by a newly developed whole body glucose test.
The mice were kept under a normal chow diet and treated with vehicle or prednisolone for 7 days. As a control, mice were fed a high fat diet for 6 weeks. Insulin resistance was assessed by a whole body glucose test (WBGT), at day 0, day 35 and day 42. Glucoselevels in the blood were measured every ten minutes for 1.5h. Label distribution was measured by GC-MS from which endogenous glucose production (EGP) and metabolic clearance rate (MCR) were calculated. After the last WBGT mice were sacrificed and characterized for general metabolic parameters.
The WBGT showed an increased endogenous glucose production upon prednisolone treatment, showing prednisolone–induced hepatic insulin resistance in mice. A high fat diet reduced metabolic clearance rate, which points at peripheral insulin resistance. This shows that this test can be used to discriminate between hepatic and peripheral insulin resistance.
These results show that the WBGT is a good method to measure insulin resistance in a less invasive manner than the standard hyperinsulinemic euglycemic clamp. Other

advantages of the WBGT are that it is repeatable, cheap and it doesn't require surgery. There are also some disadvantages, because it gives less information about insulin levels and the glucose and insulin levels are not adjustable

5.3.3.3.5.3. A suitable delivery system is key in the concept of siRNA-mediated treatment, by Roosmarijn Vandenbroucke, University of Ghent, BE

RNA interference (RNAi), a naturally occurring process of sequence-specific post-transcriptional gene silencing, is an important biological process for modulating gene expression. In gene therapy, it can be necessary to cause RNAi in target cells, this can be done by the delivery of chemically synthesized siRNAs, resulting in a sequence-specific, robust silencing of the targeted gene.
The major problem is that cells do not easily take up siRNAs. Therefore, it is important to have good methods, which can deliver intact siRNA into the cytoplasm of the target cells, by overcoming cellular barriers. At the moment viral and non-viral (polymers, lipids…) methods are used, however they both have their disadvantages (such as safety and immunogenic response in the case of viral and low transfection efficiency with non-viral methods).
There are two possibilities which give good chances. The first method is to use ultrasound combined with microbubbles to 'shoot' the siRNA complexes into the cytoplasm of the cells and the second is the use of polymers which induce endosomal release of the siRNA after endocytosis.
To avoid endosomal uptake and subsequent breakdown, ultrasound can be used. It is believed that ultrasound, especially when combined with microbubbles, causes small (100 to a few 100 nm large) transient pores in the cell membrane which allows large molecules to enter the cytoplasm. Microbubbles packed with a high amount of siRNA containing nanoparticles can induce higher gene silencing than free siRNA nanoparticles. The advantage of this system is also that the delivery is targeted.
Another method is to use Poly β amino esters (PbAEs). These are biodegradable polyamines that are synthesized by Michael addition of either primary amines or bis (secondary amines) to diacrylate esters. Two different PbAE:siRNA complexes were tested. They both induced efficient gene silencing in hepatoma cells and primary hepatocytes without causing significant cytotoxicity. An important result was that PbAE2:siRNA induced gene silencing was maintained for at least 5 days in fast dividing cancer cells. This supports the hypothesis that the endosomal release mechanism of PbAE:siRNA complexes is based on an increase of osmotic pressure in the endosomal vesicles after polymer hydrolysis.

5.3.3.3.6. Session VI: Research / Development / 3R-alternatives

5.3.3.3.6.1. Molecular imaging in vitro and in vivo, by Tony Lahoutte, Vrije Universiteit Brussel, BE

Molecular imaging is the visualization, characterization and measurement of biological processes at the molecular and cellular level in living systems.
The point of this research is that it is possible to do imaging of animals as if they were patients. Imaging does, however, not work if the animals are stressed, which is the same in the human situation.
The question posed (anatomy, physiology, cell or molecular imaging) determines the technique to be used, such as CT (computer tomography), PET (positron emission tomography), fMRI (functional magnetic resonance imaging), SPECT (single positron emission computed tomography)…
In the selection of the most suitable 'in vivo' imaging biomarker for a given molecular target there are two important properties that need to addressed: affinity of the radiolabeled molecule for its target and background biodistribution of the radiolabeled molecule. Both issues are equally important for the imaging process.
Dynamic imaging, using luciferase, is a technique that can be used for cell tracking purposes and in oncology. For example in cancer, the light is proportional to the size of the tumour.
Optical imaging gives 2D images, but the problem with this technique is that the reproducibility is not very high.
Fluorescence is another 2D method, this however has the disadvantage that it has limited reproducibility and the quantification of the results is difficult.
Thanks to new software and the advancement of the techniques, results which are really in 3D view can now be obtained, by MicroSPECT/CT and MicroPET/CT scans.
Another method is nanobody imaging, which uses molecular imaging probes for disease related cell surface biomarkers. With this method, antibodies have to be produced, then they are injected in the animal, which is, after 1h, anesthetized to make the image. This way you can quantitatively measure the amount of the drug reaching the target. This can be used for organ biodistribution and targeting, pharmaco-kinetics, intra-individual comparison and serial intra-individual monitoring.
A last method that was discussed is preclinical cardiac imaging. This can be used for serial measurement of myocardial infarct size and myocardial function. So, in an infarct, this method can be used to quantify the size of the damage to the heart and the recovery.
To conclude, imaging methods are non-invasive and allow repetitive measurements, disease related parameters can be measured at early stages, intra-individual comparison reduces the variability of the measurements. Thanks to this last point, a lower number of animals is needed for obtaining statistical relevant results.

5.3.3.3.6.2. Non-contact laser capture microdissection: an alternative method in live cell sampling, by Renate Burgemeister, Carl Zeiss MicroImaging, DE

Laser microdissection is an interesting technique with which it is possible to get a high-resolution control of the sample composition. Single individual cells, tissue-parts… can be selected or rejected for preparation. This is based upon a microscope with a built-in laser, connected to a computer program. The computer mouse can be used select a part of the sample that one wants to cut out.
This technique is also called Non-contact Laser Capture Microdissection (LCM), because there is no manual contact with the samples, which prevents contamination.
Zeiss developed the PALM MicroBeam system, which combines this laser technology with robotic tools for precise microdissection. Most striking is the patented pressure catapulting feature, which makes that the cut out sample is catapulted upwards, into the collection tube. As such, no impairment occurs to the recovery of DNA, RNA or proteins.
An important innovation is the laser driven isolation of live cells out of a cell culture. Individual or small groups of cultured cells or stem cells can be used for direct molecular analysis or re-cultivation. This makes it easier to isolate cell clones and separate different cell types by morphology or fluorescent labelling. With the focused laser it is possible to make small holes into cells and nuclear cell walls. This enables the injection of for example drugs or genetic material into the cell, without having to use viral vectors or having to treat the cells with chemicals.

5.3.3.3.6.3. COMICS: developing high throughput comet assays for DNA damage and DNA repair, by Amaya Azqueta Oscoz, University of Oslo, NO

Comics is an EC strategic targeted project (STREP) of the PF6, aimed at turning the comet assay into a high throughput assay suitable for use in screening chemicals for potential genotoxic and cytotoxic effects. It relates to the EC policy on REACH and the philosophy of the 3Rs.
The comet assay is a qualitative method for measuring single- and double-strand breaks in DNA. Cells embedded in agarose are lysed with detergent and a high concentration of NaCl, leaving the histone-depleted DNA as ‘nucleoids’, and electrophoresed. This results in comet-like structures, stained with DNA-binding dyes, which can be examined with fluorescence microscopy. The relative tail intensity reflects the level of DNA damage. The comet assay was adapted to measure DNA repair capacity in cell extracts, to be used in biomonitoring studies.
The applicability of the comet assay to the screening of chemicals for genotoxicity is limited because the number of samples that can be run at the same time is limited by the size of the electrophoresis tank. Even with computer-based image analysis, scoring is labour-intensive and time-consuming. Medium- or high-throughput versions of the comet assay and an automated scoring system are necessary for monitoring damage in a large number of samples.
One of the tasks of the University of Oslo is to develop an *in vitro* repair assay based on the comet assay to detect different pathways, and to compare this method with DNA

repair chips. This *in vitro* repair assay estimates the repair activity of cell extracts by measuring their incision activity.

5.3.3.3.7. Session VII: Research / Efficacy testing / 3R-Alternatives

5.3.3.3.7.1. Use of flow-based assays to monitor drugs with antithrombotic, anticoagulant or anti-inflammatory potential, by Johan Van Heemskerk, Maastricht University Medical Centre, NL

Cardiovascular diseases are the most common causes of death in the Western world. Thrombosis, through increased coagulation of the blood, and atherosclerosis, through inflammatory responses caused by circulating leukocytes, are two of the main diseases.
Prevention of these diseases exists in antithrombotic therapies and antiplatelet and anticoagulant medication. The treatment is often incomplete or, if too strong, there is risk of bleeding.
Validated laboratory assays are only used for testing platelet or coagulation function only. There are no integrative or bedside assays and no tests incorporating blood flow. Therefore dose and type of medication is very important in the treatment of these diseases.
Research to thrombus formation *in vivo* is difficult, because it is not possible in humans, so it has to be done in animals. Some of the models are the haemostasis model (mechanical puncture of rabbit mesenteric arteriole), the arterial thrombosis model (free radical-induced damage of mouse mesenteric arteriole) and the atherothrombosis model in mouse carotid artery (ultrasound-induced rupture of carotid plaque in apoE-/- mice).
Flow chambers are now also used for *ex vivo* measurement of thrombus formation. Also here there are multiple choices in design, substrate and measurement. One of these methods is the pulse-free flow of whole blood over collagen type I.
Next to immunoglobulin-type collagen receptor GPVI on platelets, there are a lot of other factors and proteins that are contributing to thrombus formation under flow conditions. It is known that human protein kinase C is required for thrombus formation.
To conclude there is a need for physiologically relevant integrative assays (bedside assays) to quantify the prothrombotic activity of blood. Flow chambers with collagen coating are valuable for measuring the multifactorial process of thrombus formation. This markedly resembles thrombus formation in vivo. Such flow chambers have a high potency for monitoring the efficacy of old and new antithrombotic drugs. They make it possible to test, among others, drug interactions.
The advantages in terms of alternatives are that blood from genetically modified animals can be tested *in vitro* instead of *in vivo*. Human blood instead of animal blood can be used for drug testing. However, there still are limitations in analysis power, therefore higher capacity test systems need to be developed.

5.3.3.3.7.2. Confrontation cultures of embryonic stem cell-derived embryoid bodies and multicellular tumour spheroids: a new in vitro model for the study of tumour-induced angiogenesis, by Nada Milosevic, University of Giessen, DE

Vascularization is a prerequisite for tumour growth and metastasis. Therefore anti-angiogenic therapy is one of the most promising strategies to defend against cancer. As a consequence, a lot of anti-angiogenic agents are currently tested in animal experiments.
Presented here is a new *in vitro* (hESC) model for anti-angiogenic screening based on co-culturing of embryonic stem cell-derived embryoid bodies and multicellular tumour spheroids. Embryonic bodies and tumour spheroids were brought in close contact by the hanging drop technique. After 24h, the two parts grew together and were plated on petriperm dishes for additional time needed. Tumour-induced angiogenesis was analysed by CLSM (confocal laser scanning microscopy), FACS (fluorescence activated cell sorting) and PCR (polymerase chain reaction).
The co-culture can be used for human tissue testing (such as hESC), exclusion of additional inflammatory processes from the animal and exclusion of metabolic activity effects from the animal.
This model can replace a number of *in vivo* models, amongst which the rabbit corneal assay, the mouse cheek pouch assay, and the mouse and rabbit cranial and skin window.
It was observed that blood vessels produced in co-culture are denser than in embryoid bodies grown alone and have directed growth of vessels towards the tumour.
The anti-angiogenic effect of three substances (thalidomide and the tyrosine kinase inhibitors SU5614 and ZM323881) are shown by immunocytochemistry and flow cytometry. The results showed that the number of tumour cells did not decrease, but that the degree of vascularization was reduced after treatment.
Inflammatory cells, invading the tumour spheroid, create a pro-angiogenic environment. However, one of the results showed that tyrosine kinase inhibitors also have anti-inflammatory properties.
Co-cultures have a great potential to reduce the number of animals needed for drug screening, since they have a good resemblance to host-tumour interactions. A possibility for further research is to study the role of axon guidance proteins in tumour induced angiogenesis.

5.3.3.3.7.3. Cancer associated fibroblasts influence epidermal morphogenesis and invasion in reconstructed 3D culture models of human skin carcinoma, by Suzan Commandeur, Leiden University Medical Center, NL

The aim of this project is the realization of an in vitro human skin cancer model for the development and testing of therapeutics. An explant technique is used to culture skin cancer biopsies of squamous cell carcinoma (SCC) on a fibroblast populated collagen layer. These biopsies are acquired from immune-suppressed individuals.
In these skin cultures, features specific of SCC can be seen, such as formation of keratin pearls, hyperproliferation and invasion.

To make the model more realistic, cancer-associated fibroblasts (CAFs) were added. These CAFs directed the invasion of the epidermal compartment into the dermal compartment.
These observations indicate that the dermal microenvironment plays a crucial role in tumour development and that this process can be extensively studied and modulated in vitro in order to increase resemblance of these skin cancer models to the in vivo situation.

5.3.3.3.8. Session VIII: Miscellaneous

5.3.3.3.8.1. Evaluation of toxicological effects of Persistent Organic Pollutants (POP) using primary fish cell cultures, by Liv Søfteland, National Institute of Nutrition and Seafood Research, NO

Since large quantities of fish are used every year for research, development of *in vitro* models is important. Since various fish species respond differently to toxic chemicals, species-specific systems must be developed, especially for farmed fish, such as salmon and cod.
The model presented here, is a factorial design applied for multiple-endpoint toxicity evaluation in salmon hepatocytes.
Organic pollutants (PCDDs (polychlorinated dibenzodioxins or dioxins), PCDFs (polychlorinated dibenzofurans) and PCBs (polychlorinated biphenyls)) are of great concern for food safety, relatively high concentrations are measured in farmed fish.
The toxic responses of dioxins and dioxin-like PCBs are induced through aryl hydrocarbon receptor (AhR)**.** Binding of a ligand to AhR results in transcriptional up-regulation of the CYP1A gene, which is used as a biomarker for these chemicals, and the AhR gene battery. The toxic responses of non-dioxin-like PCBs are thought to be induced through other mechanisms.
The toxic equivalency (TEQ) concept is used in risk assessment of mixtures with dioxin and dioxin-like compounds. This method relies on some assumptions, such as identical dose-response curves and that the combined effects of chemicals are additive. But these assumption are usually not achieved in complex mixtures. TEQ can therefore underestimate toxicity. Thus, an alternative method, able to evaluate the occurrence of combined effects, is needed. The aim of this exposure study was to use a multi-endpoint strategy to evaluate the combined effects of chemicals in a simple mixture with the non-dioxin PCB 138 and the potent AhR agonists 2,3,7,8'-TCDF (TCDF or tetrafuran) and 1,2,3,7,8'-PeCDD (PCDD or pentadioxin).
For the factorial design, primary hepatocytes were isolated from four salmons, the cells were exposed for 24h. qPCR (quantitative polymerase chain reaction) was used to quantify the transcriptional levels of eight genes, which were taken from phase 1 and II enzymes, cellular stress, apoptosis and control genes.
In this study, it was shown that only CYP1A and UDPGT (uridyldiphosphoglucuronyl transferase) were good biomarkers for PCDD, TCDF and PCB138. Using primary fish cell cultures and multivariate data analysis of qPCR data are shown to be a useful tool in

toxicological studies. A multi-endpoint strategy can enhance the quality of risk evaluation of chemical compounds. But TEQ is a useful concept in risk assessment of mixtures with dioxin and dioxin-like compounds, until a better evaluation strategy is developed

5.3.3.3.8.2. Free Concentrations in In Vitro Cytotoxicity Assays, by Nynke Kramer, University of Utrecht, NL

Only the free or unbound concentration is considered available for uptake by the target tissue or organism to cause an effect. Understanding the differences in free concentrations *in vitro* and *in vivo* may well be critical for *in vitro - in vivo* extrapolations. Currently, most *in vitro* studies base their results on the nominal concentrations, this may be incorrect, when the compound can bind to serum protein or evaporate.
The aim of the study presented here was to assess what system components and what physicochemical properties determine the free concentration and cytotoxicity of polycyclic aromatic hydrocarbons (PAHs) to Balb/c 3T3 fibroblasts in a neutral red uptake assay.
Solid phase microextraction (SPME) was used to measure the partition coefficients of PAHs to serum, well plate plastic, cells and headspace. These partition coefficients were used to model the free concentration of the compound *in vitro* and correlated with the compound's Henry's Law Constant (HLC) and octanol-water partition coefficient (K_{OW}). The estimated free concentration was compared to the free concentration measured in the assay using SPME, compared to the proportion of compound measured in each assay component, and linked the effect concentrations (EC_{50}). Results indicate that the free concentration of volatile and hydrophobic compounds are significantly reduced in a typical *in vitro* set-up as they readily evaporate and bind to matrices such as serum protein and plastic. This reduction in free concentration is accompanied by an increase in the EC_{50} and can be modelled using the partition coefficients of the compound to assay components, which in turn may be estimated by the K_{OW} and HLC.
Currently different tests and different laboratories use different *in vitro* set-ups with different serum levels in their culture media. If cell assay procedures are standardised, by using free concentrations, this may lead to a better understanding of the *in vivo* situation.

5.3.3.3.8.3. A cell-based assay for the detection of a potent algal neurotoxin in biological matrices, by Mirella Bellocci, University of Modena, IT

In the last years, the occurrence of toxic harmful blooms has increased in the Mediterranean Sea. The genus *Ostreopsis* produces palytoxin, which causes severe neurological symptoms and can be lethal. According to the current EU legislation all seafood must be analysed for the detection of algal biotoxins. The mouse bioassay remains the reference method in the EU for these analyses.
The goal of this study was to develop functional assays in biological matrices, preferably using cultured cells, to obtain toxicologically-oriented methods that can effectively detect

and quantify the overall burden of biologically active toxins in unknown samples and to decrease the number of animals used for these analyses.
A cytolytic assay that could detect palytoxin and its congeners in biological matrices by the use of an established cell line grown as monolayer was developed. MCF-7 cells were used and cytolysis was measured by the release of cytosolic lactate dehydrogenase (LDH) in the buffer added to treated cells (culture supernatant). A dose-dependent increase in LDH activity in culture supernatants was detected when MCF-7 cells were exposed to palytoxin and its analogue ostreocin D. The cytolytic response induced by palytoxin and ostreocin D was specific for this group of compounds, acting on Na^+, K^+-ATPase, as it was prevented when cells were pre-incubated with ouabain. The specificity of this assay for palytoxin and its congeners was confirmed by the finding that cytolysis was not detected when MCF-7 cells were exposed to unrelated toxins. Using extracts from biological materials after spiking with the palytoxin standard, a good correlation was found between palytoxin levels measured by our cytolytic assay, and the expected values. Palytoxin was detected in naturally contaminated materials in this assay. This assay is a viable alternative to animal-based methods for the determination of palytoxin and its congeners in contaminated materials.

5.3.3.3.8.4. Regulatory Acceptance and Use of in vitro Methods for Non-Clinical Testing of Human Medicinal Products, by Sonja Beken, Federal Agency for Medicines and Health Products, BE

Developing new medical products is a process which requires a lot of time and labour and bears a high risk of early termination. Therefore, strict regulations are needed to ensure the quality and safety of the new products.
For non-clinical studies recommendations are made by various institutions, for example in Europe by EMEA (European Medicines Agency) and Eudralex (EU legislation in the pharmaceutical sector), in the USA by the FDA (Food and Drug Administration) and in Japan by the Japanese Ministry of Health and Welfare. But regulatory acceptance of *in vitro* methods passes mostly via the ICH (International Conference on Harmonisation).
The CPMP (Committee for Proprietary Medicinal Products) Position Paper on Replacement of Animal Studies by *In Vitro* Models (CPMP/SWP/728/95 - adopted 1997) gives recommendations on a number of topics, such as the feasibility of replacing *in vivo* animal studies, the procedure for validating *in vitro* tests, the procedure for incorporating *in vitro* tests into the regulatory requirements and areas for which the acceptance of *in vitro* tests can be considered.
The criteria of acceptance of *in vitro* methods depends on the goal of the test. For example early toxicity, compound screening tests, can go by in-house validation by companies, there is no regulatory involvement. Whereas for exploratory, mechanistic studies for regulatory decision-making, they have to be based upon demonstrated “scientific validity”. Pivotal (guideline-driven) studies can go via different routes of validation: if they are historically introduced *in vitro* models, no formal validation is needed. Transition from exploratory, mechanistic screening models to pivotal studies is

based on accumulated experiences (for which databases are reviewed). Replacement of established animal studies by *in vitro* models requires formal validation.
Examples of *in vitro* methods accepted as pivotal non-clinical studies include formally validated tests such as the 3T3 NRU (3T3 Neutral Red Uptake) phototoxicity test. The use of this test is recommended by the EU, but not by the ICH yet. Another example is the *in vitro* micronucleus test, which was recently validated by ECVAM and is now taken up in the new ICH S2 Guideline. This test is to be used in the standard testing battery for genotoxicity. Also for safety pharmacology and hepatotoxicity testing examples were given of regulatory accepted *in vitro* tests, although they are not formally validated yet.
For the future it is important that there the communication with other sectors (such as chemicals, cosmetics…) is increased, although with the help of EPAA (European Partnership for Alternative Approaches to Animal Testing) this has already increased. It is also important to identify areas where new *in vitro* methods might be appropriate.

5.3.3.4. Discussion

Q: *In vitro* methods have a lot of variables, more than *in vivo* testing, so what is the motivation to develop *in vitro* methods?

A: Beken: This question can not be answered from a regulatory view only. But if an *in vitro* method gives the same endpoints as animal tests, then it is obligated, according to Directive 86/609, to use the *in vitro* method. Therefore, it is important that good *in vitro* methods are developed.

A: Boehn: An *in vitro* test, is a good test when it can be used for regulatory purposes.

A: Rogiers: One reason is also that the ethical views are changing, and people want to have to rely less on animal use. Animal welfare groups are also getting stronger and having a greater impact.
The development of alternative methods is far more difficult than thought before. Therefore, it is also important not to oversell new methods. Having good *in vitro* tests is more important than pressure. To this end, it is important that regulators trust researchers. Liability is another issue for which it is important that the industry has good *in vitro* tests.

A: Garthoff: From the industry, there is the question why add on *in vitro* tests have to be done, if the *in vivo* test is still used. The reason for this is that *in vitro* tests have to be performed in Europe, but for the rest of the world the *in vivo* tests still have to be done.
Industry is interested in *in vitro* tests, not only because of ethical views, but also because animal testing is costly. It even went unnoticed that in some fields *in vitro* tests were already performed. Cosmetics and safety testing is only a part of the fields where animals are used. For example, in development, the number of animals used has decreased a lot already over the past 20 years.

Trust is a very important issue. *In vitro* tests can be used in industry without being validated, but if a test is validated, that is a way of creating trust in that test.
Discussion between government and industry is often difficult, which is also a problem to come to consensus about certain issues.

Q: In Europe, Japan and the USA, there are strong ethical issues and regulatory guidelines. This seems to be less the case in upcoming countries like China and India. They seem to have less problems with *in vivo* testing. Is there a chance that Europe will loose competitiveness due to this? Or is the industry moving away to those countries?

A: Rogiers: For cosmetics, this is clear. From March 11, 2009 there is a cosmetic ingredient testing ban. Alternative methods have to be used to test cosmetics and their ingredients, with some exceptions, such as developmental toxicity. From 2013, also the existing marketing ban plays an important role, meaning that products and ingredients tested in animals may not be sold in Europe, even if they are tested outside of Europe. This means that companies will probably move out of Europe and only sell old products in Europe, because they can not test new ingredients for this market. The use of alternative methods is now a political field. The problem is that science can not be pushed and decisions are not always scientifically based.

A: Garthoff: The fact is that industry is already moving out of Europe, due to cost issues. However, discussion with regulators is important and for this it is necessary to have experts from the field and from the companies. Europe has a good knowledge base, so Europe will continue to play a role in this.

A: Beken: EMEA has already been approach by, amongst others, China and has sent people from their regulatory agencies to those in Europe, to learn about legislation and the implementation of alternative methods.

Q: Rogiers to all speakers: Can the subject of your research be used in a practical way? Are there possibilities to use it in safety, efficacy testing…?

A: Pivarcsi: The research to miRNA in skin is basic research, but some biomarkers may come out of it and of course the field moves quickly forward.

A: Søfteland: Evaluation of toxicological effects of POPs using primary fish cell cultures is the best way to study this subject. This can be used in legislation.

A: Laskewitz: Prednisole treatment can partly replace hypoglycemia

A: Kramer: Applying the free concentration concept to *in vitro* assays is more a principle now. It can be used as a checklist.

A: Belocci: A cell-based assay for the detection of a potent algal neurotoxin in biological matrices can be used for regulatory testing for toxins in seafood.

A: Wanner: New ideas in basic research in allergic skin diseases is about sensitisation. The new method is transferrable from one lab to another.

A: Commandeur: The research to cancer associated fibroblasts is basic science, but can also be used in efficacy testing.

A: Milosevic: Confrontation cultures of embryonic stem cell - derived embryoid bodies and multicellular tumour spheroids can be used in efficacy testing and for drug screening.

A: Schug: The comparison of hepatocyte *in vitro* systems for gene expression studies is useful, because carcinogenicity studies cost a lot of animals.

Q: Rogiers: It is very important to know the applicability domain, can the tests be used for chemicals or for pharmaceuticals?

A: Ojer: The study of nanoparticles is very important for regulators.

A: Oscoz: The main output of the COMICS project is to develop a high throughput comet assay.

Beken: The Comet assay is very controversial, because it gives a lot of false positives.

A: Koeper: *In vitro* differentiation of skin sensitizers by cell signaling pathways can be used to screen for sensitizers.

A: Frankart: Identification of cell responses induced by a sensitizer in a reconstructed human epidermis is also research to find methods to screen for sensitizers and to make a difference between sensitizers and irritants.

Beken: Different people are working on methods for sensitizers, but in the end regulators want one good method.

Vandenbroucke: People working on gene delivery and carriers, it is important to start with *in vitro* work.

Kireche: *In silico* methods for the prediction of contact dermatitis are developed to know more about the chemistry, which is important to understand the results of later tests.

Rogiers: Chemistry is indeed important. The number one reason why cosmetic ingredient dossiers are rejected at the SCCP level (Scientific Committee on Consumer Products) is that the physico-chemical properties and the chemistry are not done correctly.

Workshop Refinement

26 - 27 February 2009, Istituto Superiore di Sanità, Rome, IT

5.4. Workshop Refinement

5.4.1. Executive Summary

The first workshop on Refinement opened a highly necessary dialogue between the different parties involved in the use of experimental animals in the drug development process. It also offered different points for discussion and as summarised below.

Refinement was defined as “any approach which avoids or minimizes the actual potential of pain, distress and other adverse effects suffered at any time during the life of the animals involved and which enhances their wellbeing as far as possible”. Refinement does not only deal with the wellbeing of the animal during experiments, but the whole life of the animal is taken into account, starting from birth, until death. This means that also breeders and housing are factors to which Refinement can be applied. Breeders and animal care takers need to be educated about Refinement initiatives.

- Refinement as moral education:
From a philosophical point of view, Refinement is not only a technical practice, but a moral relation as well, because it involves responsibility and care, hence its implementation needs both scientific and philosophical analysis. It must be seen as a moral relation between human and non-human agents. Training concerning refinement could therefore be seen as a sort of moral education.

- Importance of well-trained and well-organised competent authorities:
Refinement can be reached through education, advice, control and support on aims, technical, operational, pre- and post-operational functions by official, institutional and public review. The competent authority must authorize personnel who carry out procedures on animals, including supervising or designing procedures and projects, supervising those who take care of animals, performing euthanasia. They must have followed education and training and have demonstrated their competence. Member States should establish National Animal Welfare and Ethics Committees to give advice to the competent authorities and to permanent ethical review bodies of establishments. The network of these national committees should play a role in the communication about and exchange of best practices.

- Opportunities for Refinement:
Some refinement opportunities were presented.
As far as “procedures”, i.e. direct inhumanity (Russell and Burch), are concerned, existing procedures clearly could be improved:

- use of surgical techniques taking more animal welfare into account;
- application of better methods for anaesthesia and analgesia;

- identification of more humane endpoints;

For “housing and environment”, i.e. contingent inhumanity, also opportunities for improvement exist:

- need for group housing of social animals;
- creation of possibilities for natural behaviour and for exercising a certain degree of control of the environment. The latter can be physically enriched through the use of objects and diminishing noise and controlling and quality, without disturbing standardisation of tests.

- New ways of interacting with animals: positive reinforcement training:

Positive interaction with humans is fundamental. An example is Positive Reinforcement Training (PRT), which enhances voluntary co-operation of the animals with husbandry, veterinarian and research procedures. Its success depends on a lot of factors including species, gender, age, history of the animal and individual factors. As it is a training technique, which is positive for both animals and trainers, breaking routine and decreasing the risk of injuries, it merits further encouragement. After the initial investment of time, later onwards it substantially saves time during daily work. It might also positively affect the quality of the data collected, as stress in animals is lower.

Pharmaceutical and chemical companies clearly showed interest in applying refinement techniques, also implementing the European ETS 123 convention on husbandry and using the new methods such as PRT, micro-surgical procedures and imaging techniques. All staff working with animals should be adequately trained and facilities for animals should be well equipped. Progress is being made, but a lot of work still has to be done. New techniques and methods are being developed and are already in use, but it is often not certain how to implement them. Implementation of new refinement techniques costs time and money, but could also have a positive effect on the public acceptance of the pharmaceutical industry and the use of experimental animals in the drug development process.

- Importance of vaccines:

Vaccines are not always considered, while there still is a lot of work to be done in that field. For this reason, the European Directorate for the Quality of Medicines (EDQM), started 16 projects on 3R methods of which 8 are focused on refinement in vaccines.

- Special concern for animals used as models to study human diseases:

• In the use of non-human primates for the study of Parkinson’s disease, refinement methodologies can be used covering different aspects and will result in an improvement of the quality of the animal life and the research quality. Positive training, through positive reinforcement, can be applied to procedures such as injections, removal of individuals from a cage, environmental enrichments, multiple recordings of neuronal activity in one session, use of telemetry. All phases of animals life should be improved: pre-, during and post-treatment.

• For Multiple Sclerosis, various refining methods/models exist. First of all, it is important to gain more pertinent information at the lowest possible cost of animals well-being. The animal model is carefully chosen taking into account the current state of

knowledge of species and protocols. Correct planning of protocols helps to avoid disturbances in early end-points. Preference should be given to non-invasive procedures, such as imaging and telemetry. If necessary anaesthesia/analgesia should be used, humane endpoints should be established beforehand. Non-invasive pharmacokinetic / pharmacodynamic techniques could be included.
When transgenic animals are involved, the general principles of bioethics apply, but special attention should be paid to unexpected effects.

- Barriers in the progression of Refinement:
Change is not always welcomed, since it comes with a certain degree of uncertainty for success, and habits of years have to be changed. New methods are often not easily accepted by regulators, who require validation of these methods. Researchers are not always aware of existing alternatives and even if they are, there is not always enough time and money to investigate and validate these alternatives.
Some of the barriers mentioned, however, also might be opportunities for change since attitude changes in researchers may result in more innovative thinking.

- The need for global harmonisation:
Global harmonisation is difficult and slow, but it is highly important to be pursued. Animal Welfare Acts exist in all EU countries, however, substantial differences exist with respect to their execution. European harmonisation should be further extended. The revision of Directive 86/609 will help in this process.

- Does Refinement lead to better science?
Whether refinement leads to better science was a crucial point of discussion. It was concluded that cases exist where it is indeed better e.g. environmental enrichment providing healthier animals. In other cases, too healthy animals might not respond to certain treatments.
Sometimes, there might only be benefits for the animals while the quality of the data might stay the same. Therefore, it is necessary to always perform cost/benefit analysis for any particular experimental procedure.

5.4.2. Recommendations

1. General Views: Replacement, Reduction and Refinement should not be considered as separate issues only, but they should be viewed as an entity on various levels, in education, research, and testing. In legislation, the interplay between the 3Rs should be recognised, and awareness of converging and diverging arguments and effects should result in practical and ethical guidelines for what to do in situations of conflict. Education and training in different aspects must be promoted among all stakeholders in the drug development process.

2. Planning and Performing Animal Experiments: Applying refinement methods for the most painful experiments as e.g. for cancer research, is urgently needed. Advanced methods, e.g. non-invasive imaging methods allow repetitive measurements and register the disease-related parameters at an early stage (non-painful stage) of cancer development. Early screening of drug candidates in drug development makes it possible to avoid unnecessary animal studies and reduce research costs.

3. Animal Breeding and Maintenance: It is necessary to promote the positive welfare of experimental animals rather than just focus on the minimisation of suffering. The notion of animal welfare as simply the ability of an animal to cope with the environment must be abandoned, in favour of a richer one that includes active improving of the degree of welfare.
   Positive training techniques should be enforced, in order to reduce the stress imposed on the animals participating in experimental procedures.
   Refining animal conditions entails the understanding of their needs. Ethological studies can provide such an understanding. As a contribution to animal welfare science, the possibilities of an "ethology of laboratory animals" must be investigated to fully understand the needs of species bred for generations in captivity, such as those involved in laboratory practice. For refinement to be successful, appropriate objective measures of welfare states need to be developed and validated. Refinement must include even the life of the animal outside the experimental procedures, with a focus on housing conditions and enriched environment, in order to promote a satisfactory level of welfare.

4. Global Harmonisation: The dissemination and promotion of refinement techniques should be a matter of the highest concern in EU countries.
   The most advanced practices (experimental techniques and animal housing) should be required in all the research sites of the pharmaceutical industry in different countries, as well as CROs and academic research partners.
   After the new directive in EU will come in force, pharmaceutical industry should verify and update the realisation of refinement in its functions with experimental animals. Similar principles should be followed in all countries inside Europe and measures should be undertaken to extend them also globally. The implementation of 3Rs in the spirit of the new European directive should be ensured also in research sites outside the European Union. Awareness of the importance of animal welfare should be emphasised.

5.4.3. Report of the Workshop Refinement

START-UP Workshop Refinement
26 – 27 February 2009, Istituto Superiore di Sanità, Rome, IT

Program:

26 February

14:00 – 14:30	**Welcome and introduction to START-UP** Annalaura Stammati, ISS, IT Vera Rogiers, *ecopa*, BE
14:30 – 15:10	**Introduction**
14:30 – 14:40	Conclusions on refinement from Animalsee EU Project. Flavia Zucco, Istituto Neurobiologia e Medicina Molecolare, IT
14:40 – 14:50	Refinement opportunities - an overview. Peter Thornton, Animals Scientific Procedures Inspectorate, UK
14:50 – 15:00	Point of view of a philosopher. Simone Pollo, Universita di Roma, IT
15:00 – 15:10	Discussion
15:10 – 15:40	**ROUND TABLE 1: Housing and legal aspects**
15:10 – 15:20	Refinement in the proposal for a directive of the European parliament and the council on the protection of animals used for scientific purposes. Kai Pelkonen, Ministry of Agriculture and Forestry, FI
15:20 – 15:30	Environmental enrichment and standardisation, by Simone Macrì, ISS, IT
15:30 – 15:40	Focus on Refinement in a pharmaceutical company is good for business, by Jan Lund Ottesen, Novo Nordisk, DK
15:40 – 16:20	Discussion
16:20 – 16:40	Coffee break
16:40 – 17:35	**ROUND TABLE 2: Recent methodologies in refinement**
16:40 – 16:50	New imaging techniques, by Vicky Caveliers, VUB, BE
16:50 – 17:00	Operant conditioning in lab primates, implementation of PRT in experimental and husbandry procedures. Fanélie Wanert, Université de Strasbourg, FR
17:00 – 17:10	Stress reduction in working with beagle dogs; a win-win situation. Pieter Verbost, Schering Plough, NL
17:10 – 17:20	Animal research in a global pharmaceutical and chemical company. Pierre Coërs, Solvay, BE
17:20 – 18:00	Discussion

27 February

09:00 – 09:30	**ROUND TABLE 4: Refinement in drug development process**
09:00 – 09:10	Refinement in the ADME studies, by Olavi Pelkonen, University of

	Oulu, FI
09:10 – 09:20	EDQM activities for Refinement of animal experiments in the field of quality control of vaccines, by Karl-Heinz Buchheit, EDQM, FR
09:20 – 09:30	Refinement in the development of biopharmaceuticals, by Gisbert Sponer, Bioassay GmbH, DE
09:30 – 10:10	Discussion
10:10 – 10:50	**ROUND TABLE 3: Animal models in human diseases**
10:10 – 10:20	Refinement in non-human studies of Parkinson's disease, by Augusto Vitale, Istituto Superiore di Sanita, IT
10:20 – 10:30	Application of Refinement in transgenic mice, by Igor Branchi, Istituto Superiore di Sanita, IT
10:30 – 10:40	Refinement in Multiple Sclerosis modelling, by Valeria Muzio, Merck Serono, IT
10:40 – 10:50	Refinement in animal epilepsy models, by Ralph Clinckers, VUB, BE
10:50 – 11:30	Discussion
11:30 – 12:00	Coffee Break
12:00 – 13:00	**Final discussion and conclusions**

Scientific committee:
Franca Fassio, member of NCP IPAM, Merck Serono, IT
Bernward Garthoff, treasurer *ecopa*, Bayer, DE
Peter Maier, Board Member of *ecopa*, Forschung 3R, CH
Vera Rogiers, chair of *ecopa*, VUB, BE
Annalaura Stammati, member of NCP IPAM, ISS, IT
Maciej Stepnick, member of NCP Polcopa, Nofer Institute of Occupational Medicine, PL
Hanna Tahti, member of NCP Fincopa, University of Tampere, FI
Augusto Vitale, member of NCP IPAM, ISS, IT

5.4.3.1. Introduction to the workshop

This workshop is the first one in a series of three, each of these is concerned with one of the Rs in the 3Rs concept of Russell and Burch, namely “Refinement”, “Reduction” and “Replacement”. This workshop is dealing with the first R, namely the sensitive issue of Refinement. Since it is important to understand the terminology used and the ethical and philosophical implications of Refinement in the use of experimental animals, much attention was given to the introduction. Annalaura Stammati (member of the NCP IPAM, IT) organised with her team the meeting and took care of a clear introduction.
Vera Rogiers, chair of *ecopa*, presented the structure and aims of the START-UP project and explained the role of the Expert Meetings and their input in this first Workshop.

5.4.3.1.1. Introduction lectures

5.4.3.1.1.1. Conclusions on Refinement from Anim.Al.See. EU project., by Flavia Zucco, Istituto Neurobiologia e Medicina Molecolare, IT

The Anim.Al.See project had as a goal to update the view on 3Rs models, both from a philosophical and scientific point of view. Among others, new definitions of the 3Rs were made.
The definition of Refinement: There are different interpretations of refinement and these should be harmonised for better communication. The proposed definition from the Anim.Al.See EU project and used here is: Any approach which avoids or minimises the actual or potential pain, distress and other adverse effects suffered at any time during the life of the animals involved and which enhances their wellbeing as far as possible. This definition applies to the whole life of the animal.

5.4.3.1.1.2. Refinement opportunities - an overview, by Peter Thornton, Animals Scientific Procedures Inspectorate, UK

Russell and Burch published 'Principles of Humane Experimental Technique' in 1959. Since then much progress has been made, however, there still is room for improvement. This lecture takes a look at some of the barriers in the progression of Refinement.
Russell and Burch divided Refinement into direct inhumanity, which contained the procedures, and contingent inhumanity, i.e. husbandry and housing.
Some Refinement opportunities were presented. First of all some of the procedures could be changed. Other surgical techniques can be used, as well as better anaesthesia and analgesia. Another possibility is the usage of a less sentient species. Finally, endpoints can be changed to more humane endpoints. For housing and environment there also are opportunities. For example group housing of social animals is a possibility, scientists in the UK have already tried it with success with rats. Enrichment and environment (noise,

air quality…) can be improved. Another important point is to give the animals enough space and the opportunity to perform natural behaviour. It is also important that the human-animal interaction goes smooth, to give the animals as little stress as possible.
Furthermore some possible barriers were discussed. First of all there is the problem that change is not always welcomed, since it comes with a certain degree of uncertainty on its success, and habits have to be changed. New methods are also not always easily accepted by regulators, who require validation of these methods. Then there is the problem that researchers are not always aware of alternatives. There is not always enough time and money to research and validate alternative methods. Housing and husbandry refinement might have an influence on data. It is not always clear to researchers who is responsible for initiating or implementing refinement in the organisation. Commitment is not always shared between the individual researcher and the organisation he or she is working for. Finally, regulation might also be a barrier, if amendments are required.
However, some of these barriers, might also be opportunities for change. If for example an establishment changes its view on Refinement, this may cause a change in the attitude of the individual researchers, so that they can be more innovative and risk new ways. Training and education increase the awareness of refinement and its advantages. A positive culture will also induce positive peer pressure. Since costs for doing animal tests are increasing, it might be good to look into refined procedures, as they in the longer term might cost less, since they may lead to better experimental design. Increasing costs can also lead to the use of a less sentient species, which is cheaper in housing etc.
To conclude, it is important to identify the barriers to and opportunities for refinement. Barriers may be turned into opportunities and for the remaining barriers a strategy for their removal must be developed.

5.4.3.1.1.3. Refinement: a philosopher's point of view, by Simone Pollo, Universita di Roma, IT

The practice of the 3Rs is not just science, since it involves ethics and morals as well. Therefore, the philosophy should be taken into the laboratory.
Refinement is not only a technical practice, but a moral relation as well. It involves responsibility and care.
For these reasons, the implementation of Refinement needs both scientific and philosophical analysis. Some issues concern both science and ethics, these are the responsibility of the research and animal welfare.
The responsibility in the practices of Refinement works on different levels. First of all, the researcher has a responsibility towards the animals, secondly towards the scientific procedure and lastly there is the responsibility towards the public, which trusts the researchers take good care of the animals they work with. These levels of responsibility must be balanced.
Animal welfare is a moral concept, and not just a description of an animal's state. It is normative, since it wants animals to be in a certain state of health and well-being, and entails philosophical concepts as well.

When comparing human and animal welfare, it becomes clear that in the case of animal welfare there is a search for objectivity, whereas in the case of human welfare there is a search for subjectivity.
To conclude, Refinement must be seen as a moral relation between human and nonhuman agents. Refinement procedures are the outcome of scientific and moral judgments. Training to Refinement could therefore be seen as a sort of moral education.

5.4.3.1.1.4. Discussion

A comment from the public was that although Refinement is meant to make the burden on the animals due to testing less heavy, it might happen that just due to Refinement even more animals are used than before. Peter Thornton answered that by Refining experiments the variation is reduced, which generally leads to a reduction of the numbers of animals for testing procedures. However, it can indeed happen that it is the other way around and that due to decreased variation, more animals have to be used in a certain test. What is more important than the numbers of animals used, is the burden of suffering they have during their lives, which is anyway less when doing Refinement.
Flavia Zucco told us that it would be good to have a list of recommendations on how animals should be best treated and which needs they have. This list should of course be made for every species on which tests are performed. This is only possible when the ethology of the animals is known. However, anthropomorphisms must be avoided. It is important that the animals welfare from birth to death is taken into account. The personnel that handles animals should be trained, so that the animals endure the least possible stress upon handling. It is

5.4.3.2. Round Tables

5.4.3.2.1. Round Table 1: Housing and legal aspects

5.4.3.2.1.1. Refinement in the proposal for a directive of the European parliament and of the council on the protection of animals used for scientific purposes, by Kai Pelkonen, Ministry of Agriculture and Forestry, FI

In the Directive, the 3Rs must be implemented through a strict hierarchy of the requirement to use alternative methods. Therefore the method must be chosen which is able to provide the most adequate results and causes a minimum of pain and distress.
According to the Directive, Refinement means to employ the methods which reduce the animals pain and distress to a minimum. Care, treatment and housing must be improved. The life-time experience of animals must be taken into consideration as well. Refinement is considered as an interim, since the ultimate goal is Replacement of all animal experiments.

Refinement can be reached through education, advice, control and support on aims, technical, operational, pre- and postoperational functions by official, institutional and public review.
People who carry out procedures on animals, including killing, supervising or designing procedures and projects, or supervising of those who take care of animals, must be authorized by the competent authority. They must have followed education and training and have demonstrated their competence. The authorization must be renewed every 5 years, after demonstrating the competences again. Minimum requirements concerning education, training etc must be published by the Member States.
The permanent ethical review body of the establishment gives advice on ethics, the application of refinement, gives information on technical and scientific developments. It must also establish and review internal operational processes in relation to the welfare of animals housed or used.
Member States should establish National Animal Welfare and Ethics Committees to give advice to the competent authorities and to permanent ethical review bodies of establishments. The network of these national committees should play a role in the communication about and exchange of best practices.
The permanent ethical review body of the establishment must annually review all projects which last longer than 1 year, with a focus on the use of humane killing methods and whether new developments are used. They must also examine whether the project authorisation needs to be submitted for amendment or renewal.
All establishments must be authorized, which is only possible after inspection and they have all the needed equipment and sufficient trained staff. This last point, however, is presented very weakly in the proposal. The authorization will specify a person responsible for compliance with the provisions of the Directive.
Each establishment should have the appropriate veterinary care available, a staff member responsible for the care and welfare of animals. Member States should carry out at least two inspections annually.
The project authorization can only be acquired after an ethical evaluation, which verifies that a project meets criteria, such as being scientifically justified or required by law, the purposes of the project justify the use of animals and that the project is designed as to be as humane towards the animals as possible.
Non-technical project summaries of authorized projects and updates, including a demonstration of compliance with the requirement refinement, shall be published by the Member States.
The EC and Member States must contribute to the development and validation of alternative approaches and encourage research in this field. Each Member State must designate a National Reference Laboratory for the validation of alternative methods.

5.4.3.2.1.2. Environmental enrichment and standardisation, by Simone Macrì, ISS, IT

Laboratory ‘standard’ housing conditions may vary considerably between different facilities. While in several laboratories rats and mice are kept in barren cages lined with bedding, in some others, they are given access to larger cages provided with toys, shelter

and various forms of enrichments. Instead of comparing the superiority of one over the other housing approach, this lecture covered the question whether the adoption of a unique standard (be the latter impoverished or enriched) constitutes an appropriate methodology capable of providing valid experimental results (externally valid and reproducible). The arguments discussed in this presentation relate to the studies in which possibility of generalization of results is involved. A normal population is generally assumed to show a Gaussian distribution for a given trait. Similarly, parametric statistical analyses – adopted to extrapolate experimental findings obtained in experimental subjects to a larger population – assume data to be normally distributed. Yet, experimental data obtained with a small sample size under a unique standard housing condition are unlikely to be normally distributed and to represent a general population that naturally occupies several different environments.
Another important point is reproducibility: when comparing two control groups, the exact same median will never be acquired, since there is always an amount of 'experimental noise'. Traditionally it was thought that standardisation and removal of environmental confound would reduce this variability and that environmental enrichment would increase the experimental noise. However, a study done by Crabbe, Wahlsten & Dudek (1999) showed that even when all possible variables were standardised, there still were differences in the results between labs. Results of another study (Wolfer et al., 2004) showed that enrichment did not cause more experimental noise. Thus, enriched housing and barren cages appear to grant a similar level of reproducibility. Such reproducibility, however does not guarantee an adequate level of external validity, considered as a measure of the possibility to extrapolate a given finding to the general population.
The working hypothesis of this presentation is that, independently of the prevailing view as to the dispute barren cages vs. enriched environments, the adoption of a unitary control group fails to represent the general population. Therefore, future studies may predispose the adoption of several control groups housed in systematically variable housing conditions. This effort would be aimed at mimicking natural environmental variation. Systematic variation should be an integral part of behavioural neuroscience studies.

5.4.3.2.1.3. Focus on Refinement in a pharmaceutical company is good for business, by Jan Lund Ottesen, Novo Nordisk, DK

Novo Nordisk is a large pharmaceutical company with more than 27.000 employees in 81 countries. Some of its main focuses are diabetes care, haemostasis management, growth hormone therapy and hormone replacement therapy. It is one of the largest users of experimental animals of Denmark. For the last ten years, about 50.000 animals were purchased annually.
In housing laboratory animals attention is mainly given to environmental enrichment and the possibility for the animals to show natural behavioural patterns.
An important milestone was the organisation of a series of workshops in 2000. International animal welfare experts were invited to discuss the most important needs of animals. The goal was to use this information to develop new types for housing the experimental animals.

Mice and rats are kept in Type IV cages of 1800cm², with a height of 25cm (mice) or 32cm (rats). Rats also have a shelf of 400cm², creating extra space. The cages are filled with an aspen bedding, paper based nesting material and a gnawing stick. The animals receive corn and maize twice a week and at least once a week the feeding is given in the bedding, so that they must seek for it.
Guinea pigs and rabbits are kept in a pen system of 30.000cm² with aspen bedding, hiding places and a gnawing stick. Next to normal food, they get cucumbers, apples, carrots etc twice a week. They also get hay, which serves both as food and nesting material.
Dogs are kept in pen systems of 3,5m² with free access to an outdoor area of 2m². The individual pens can be connected in a flexible way. There is a platform and a ramp which provides resting and observation opportunities. The dogs can play on regular periods in a large outdoor enclosure. The dogs are kept under an age-divided socialisation programme.
Goats are kept in a large outdoor enclosure with free access to shelter places. For food, hay is provided ad libitum.
Pigs and minipigs are kept in pens of which the size depends on the size of the pigs and is between 3 and 5,6m². They are covered with wood shavings and straw. They get carrots, apples… three to five times a week, once a week the food is hidden in the bedding. In the pens there are chains and other objects for the pigs to bite on. On regular times they also get exercise out of the pens. They, however, are not allowed to churn in the mud, because they have to be clean for the tests.
These examples can set new standards for housing and care of animals. These are initiatives taken to fulfil the basic needs of the animals. This, however, does not mean that this is necessarily the best practice. They also show how a company can combine the needs of a commercial business with the need to give enough attention to animal welfare.
Generally it can be said that having good animal welfare as a company contributes to a broader acceptance of experimental animal use by the public. A good example of this is that Novo Nordisk has recently been elected as Animal Friend of the Year.
To conclude this might not lead to better pharmaceuticals being developed, but it raises the awareness of ethics in business.

5.4.3.2.1.4. Discussion

Assessors and regulators must also be trained and educated. Inspectors of companies must be specialists in the matter to do their job well.

Recently it was stated that countries must implement reference labs, one question from the public was how small countries should manage this.
There are different levels on which this has to be implemented in different countries. First of all, there already must be an established laboratory, before it can be chosen as the reference lab. Finland already has such an established lab. Susanna Louhimies had earlier on already mentioned that countries that can not afford a reference lab, can be referred to other countries.

Another comment was that Refinement can actually stop studies, very healthy mice can stop experiments, due to their high immunity. For certain tests, anxiety has to be relatively high and mice, living in an enriched environment, can be unresponsive to anxielitics. The conclusion is that enriched environments is not a gold standard for all experiments.

5.4.3.2.2. Round Table 2: Recent methodologies in refinement

5.4.3.2.2.1. New imaging techniques, by Vicky Caveliers, VUB, BE

Molecular imaging is the visualization, characterization and measurement of biological processes at the molecular and cellular levels in living systems.
With these techniques it is possible to do imaging of the animals in the same way as is done with human patients.
Different techniques can be used, such as CT, PET, MRI, SPECT. The required information (anatomy, physiology, cell or molecular imaging) determines which technique has to be used. In the selection of the most suitable *in vivo* imaging biomarker for a given molecular target there are two important properties that need to addressed: affinity of the radiolabeled molecule for its target and background biodistribution of the radiolabeled molecule. Both issues are equally important for the imaging process.
Bioluminescence, using luciferase, is a dynamic imaging technique. This is a 2D technique that can be used in cell tracking and oncology. The interesting fact is that the emitted light is proportional to the size of the tumour. The tumour can be detected immediately after injection while it is still very small and almost invisible.
Fluorescence Imaging is a quantitative method with which 3D images can be acquired, when they are combined with MicroSPECT/CT and MicroPET/CT scans.
Another method is Nanobody Imaging, which uses molecular imaging probes for disease related cell surface biomarkers. The first step is to produce marked antibodies, which are injected in the animal. After 1h, the animal is anesthetized to make the image. This way you can measure quantitatively how much of the drug reaches the target. Nanobody Imaging can be used for organ biodistribution and targeting, *in vivo* pharmaco-kinetics, intra-individual comparison and serial intra-individual monitoring.
Advances in nuclear medicine technology and the ability to radiolabel a wide variety of compounds for in vivo use in humans, have created an interesting technology for assessing pharmacokinetic properties of drugs in the process of drug development. This can be accomplished by isotopically labelling a potential new drug and tracing its pharmacokinetics and metabolism with imaging and by studying the effect of the unlabelled drug using established radiopharmaceuticals, like alterations on blood flow…
Due to the non-invasive nature of these approaches, physiological, pharmacological and biochemical information can be obtained in human subjects and accordingly aid in the evaluation of new drug therapies.

To conclude, imaging methods are non-invasive and allow repetitive measurements and disease related parameters can be measured at early stages. Early screening of drug candidates in drug development makes it possible to avoid unnecessary animal studies and reduce research costs. Intra-individual comparison reduces the variability of the measurements. Thanks to this last point, a lower number of animals is needed for obtaining statistically relevant results.
One question was how the animals were immobilized during the imaging. This is done with anaesthesia, which can indeed have an influence on the data. To avoid this, comparisons are made and the best method is used.

5.4.3.2.2.2. Operant conditioning in lab primates, implementation of positive reinforcement training (PRT) in experimental and husbandry procedures, EUPRIM-NET NA3 Activity, by Fanélie Wanert, Université de Strasbourg, FR

EUPRIM-NET (European Primate Network) is an FP6 project, which brings together 8 main public European primate centres that combine research and breeding. It consists of 4 network activities contributing to the 3R concept. One of these activities is implementation of training methods to increase the level of animal cooperation in experimental procedures.
Nonhuman primates are easily stressed when being handled, which can have a negative impact on animal welfare and the data. Positive reinforcement training (PRT) is a method to make animals used to handling, by rewarding them. This improves animal welfare, reliability of data and safety for both animals and personnel.
The present activity will ensure that positive reinforcement training of primates becomes a standard technique across Europe.
The objectives of this action are to organise workshops on PRT with the 8 primate centres and to start an education program and develop written instructions and a video on the PRT for personnel working with primates. These instructions will then be published and presented to other institutions holding primates.
PRT is part of behavioural management of primates, which also includes enrichment of cages and socialization. Behavioural management improves welfare by providing mental stimulation. PRT enhances voluntary cooperation of the animals with husbandry, veterinarian and research procedures.
PRT is done by operant conditioning, so that the frequency of a behaviour increases or decreases according to its consequences. There are different tools to communicate with the animal. You need a primary (such as food) and secondary reinforcer (a learned signal, such as a clicker) and a target (something the animal can touch, such as a shoe horn). Natural behaviours can be reinforced, a behaviour can also be divided in tiny responses and cues, such as discriminative signals and stimulus controls are also given.
An animal is reinforced when it shows the desired behaviour. In PRT timing is very important, hence the bridge is very useful.
PRT can be used for target training (sitting, entering in capture devices…), cooperate during handling (no aggression, show hands, arm, feet…, take temperature…), avoid

unnecessary handling (while weighing animals, urination…) and tolerate painful events (such as injections).
The success of PRT depends on a lot of factors, such as the species, sex, age, the history of the animal and individual factors. The experiments performed with this method may be limited by previous harmful experiences, painful procedures and timing schedules. There are also time and cost limits, because you need qualified people who dedicate time to this every day and you need an adapted caging design.
Labs should be encouraged to start working with PRT, because this is positive for both the animals and the trainers, it breaks routine and decreases the risk of injuries. It is also rewarding, because after the initial investment of time, it saves time with the daily work. It reduces the animal's stress response, so it might also increase the quality of the data.

5.4.3.2.2.3. Stress reduction in working with beagle dogs; a win-win situation, by Pieter Verbost, Schering Plough, NL

Schering Plough has recently changed its way of dealing with Beagle dogs. The basic methods did not change, but the procedures and caring were changed.
The dogs were kept in standardised environments, which, as was discovered recently, makes them more stressed. This can negatively influence the results. Therefore, as well as for ethical reasons, procedures were changed, so that the dogs could show more of their natural behaviour. It also was known that the care takers do have a better feeling about their work if the animals are feeling good as well.
Dorthe Odefey (of Dyreadfærdscenteret), who has expertise in working with dogs, was asked to give instructions to the technicians about how to deal with the dogs in the most positive way. She helped to let the care takers understand the signals of the dogs. Understanding this leads to a better intercourse with the animals. Reinforcement is also used, also with clickers.
These new methods are first tested in a small group of animals and the results are very promising. The care takers experience is that the animals behaviour is different now. From the dogs perspective, they are more willing to cooperate, they hardly bark anymore and they are physically stronger. From the human point of view, there are, however, some concerns. Since these new methods require an investment of time, this might become a problem when it has to be done with a larger number of dogs. This might be a point of discussion with companies management, however, here legislation can help. If these methods are in the legislation, then the management has to allow it.
One of the important points is to approach the dogs in a uniform way. This way procedures are predictive for the dogs, which gives them the least stress. Of course it is not possible to make the handling and procedures completely stress free, but it is important to minimize this as much as possible.
The ultimate goal for Schering Plough is to have more trained staff and provide the dogs with outdoor pens for daily exercise.

5.4.3.2.2.4. Animal research in a global pharmaceutical and chemical company, by Pierre Coërs, Solvay, BE

Solvay is a global group of pharmaceutical and chemical companies.
The "Laboratory Animal Policy" is based on the 3R principles. They are applied in a hierarchical order. Replacement is done wherever possible, if this is not possible Reduction is applied. Refinement is then used wherever the two former options are unusable.
The health of the animals is monitored daily by the care takers, every 3 months they are taken out and investigated for infections. Environmental enrichment is applied to the cages, the animals are also housed in groups when it is possible. The European ETS 123 convention on husbandry is implemented. Techniques used include PRT and the use of micro-surgical procedures. The technicians and care takers are also trained for this purpose. Before procedures are started, the animals are given pre-emptive analgesia, whenever possible. For experiments the appropriate species are always chosen, lower order species are preferred. New techniques, such as NMR, are used, as well as mini-pumps to avoid daily treatment. Humane endpoints and killing methods are used. The aim of Solvay is to refine all aspects of animal use.

Next some of the challenges and limitations were presented.
Solvay has three pharmaceutical research sites in 3 different countries, the site in France was added only recently, but in the other two the work was organised independently. The result is that there is a different cultural view on animal welfare. Still, there is no global coordination yet.
30% of the animal work is done by CROs (Contract Research Organisation) and universities. At these places, there is no control regarding animal welfare.
Some other limitations are the fact that there is a need for technical improvement and more space on one of the sites. Sometimes there also are restrictions due to the scientific objectives.
The solutions include the implementation of a Global Laboratory Animal Science Officer and Animal Welfare Officers at the sites, who must have regular meetings. A general document on animal care and use standards will be drafted.
Plans are made to improve the sites, due to the existing differences between the site, the plans will be adjusted to each site individually.
Staff will be educated and trained on a continuous basis.
The SACUC (Solvay Animal Care and Use Committee) defines general standards for Solvay's laboratory animal practices and for any third-parties performing animal tests for Solvay. An overview of all activities involving animals testing in the different sectors is maintained and areas of special interest or concern are critically reviewed. Finally, SACUC informs the management about animal care and use when necessary.

5.4.3.2.2.5. Discussion

The first comment on all four presentations given, was that the work being done was all very exciting, but that all required a lot of investment, thus the question was how much this is a barrier for implementing these techniques.
Pierre Coërs told that in his company the management has changed the way it looked at animal testing. Thanks to this change in view changes could be implemented. To this end it is important to have good communication between the animal caretakers and the management.
Fanélie Wanert admitted that PRT is difficult to implement in all companies. One of the barriers is that you have to start the training before testing. With monkeys, the training takes at least about four to six weeks, although it depends on the tests to be performed. CROs, for example, don't have enough time to do this. The amount of animals not responding to the training depends on the age on which the training is started, the earlier, the better, therefore it is best if the breeders already start training the animals.
Vicky Cavéliers told that it is important to show new techniques to management and the outside world, to get them to spread the message to companies that new techniques can be useful. Making images is easy, but quantifying these images is difficult. Since scanners are very expensive, work is also done as a service lab for small companies. Of course, these techniques are also important for basic research.
Pieter Verbost commented that one of the main barriers is time, because testing whether new methods work often requires quite some time.

Another comment was that there actually are few papers on the scientific successes of Refinement. The reason is that nobody is willing to do a study to compare experiments performed with and without Refinement techniques.
Due to this, it can not be said with certainty that Refinement leads to better pharmaceutical results. So, the question one must dare to ask, is, if Refinement does not lead to better results, then why does it still have to be done? This might indeed be true, however at the moment it is a good reinforcement to perform the Refinement.

As said, the implementation of new Refinement techniques costs time and money, but they can make the public acceptance of the pharmaceutical industry bigger, which is an important argument for the management of companies.

One problem is that in 86/609 the reuse of animals is denied, there even was one proposal to avoid any reuse of animals.

Asian companies and CRO's have other views on animal welfare, which is according to their national laws, but this sometimes makes it difficult for European companies to work together with them. If, for example, people from Chinese companies come to European laboratories, they want the same methods, for the quality of the data, and not so much for the animal welfare. China and India are not fully GLP compliant, so the EU is not obliged to accept their data. Of course, the one who pays the CRO's defines the rules, so if a European company will work together with an Asian CRO, the management can ask the

CRO to adapt to the European animal welfare rules. This once again makes clear that all stakeholders (technicians, management…) must know the benefits of alternative methods and animal welfare.

Finally, the academic world deals with a money issue, which makes it difficult to apply Refinement in this area. However, 15% of the testing animals is used in academic research. In the UK, the academic world often works together with the industry, so that there is no problem to use expensive equipment and thus imply Refinement.

5.4.3.2.3. Round Table 3: Refinement in drug development process

5.4.3.2.3.1. Refinement in the ADME studies, by Olavi Pelkonen, University of Oulu, FI

The number of biochemicals on the market is increasing rapidly.
In doing PK studies, and especially in predicting essential ADME (Absorption, Distribution, Metabolism and Excretion) characteristics, there are some hurdles. The mechanistic processes can be reproduced and studied by performing *in vitro* studies, but the biological component must be human or human-derived. Animal studies can not reproduce or predict human PK (pharmacokinetics) characteristics without detailed mechanistic investigations, and even when these are available it still is difficult.
As a means to overcome these problems, Physiologically Based Modelling (PB-PK/PD) is used. With this method, *in vitro* data are integrated into a physiological model, which incorporates variable factors such as sex, age, disease… This way a kind of virtual experiments are carried out.
The ICH has guidelines for pharmaco- and toxicokinetics. In ICH M3 (Guidance on Non-Clinical Safety Studies for the Conduct of Human Clinical Trials and marketing authorization for Pharmaceuticals) there is a statement about PK studies: *In vitro* metabolic data for animals and humans, and exposure data in animals should be evaluated prior to initiating human clinical trials. Further information on absorption, distribution, metabolism and excretion in animals should be available prior to exposing large numbers of human subjects or treating for long duration (generally prior to Phase 3). These data can be used to compare human and animal metabolites and for determining if any additional testing is warranted.
Animal PK/TK (Pharmacokinetic/Toxicokinetic) studies are done for exposure ascertainment and concentration yardsticks, for interpretation of toxicology and safety pharmacology studies. The possibility to extrapolate this to the human situation is limited. Therefore, animal PK/TK studies are useful for, for example, veterinary drugs, but they can not be used for predicting human PK characteristics.
There are some very useful advances in methodology, such as new analytical possibilities for the identification and quantification of chemicals, *in vivo* non-invasive imaging techniques and the appliance of microdosing. Cell cultures are also being developed for better use. Examples of these are the enzyme- and transporter-competent human cell lines

and the 3D culture systems, which mimic tissues in a more realistic way. Finally, with the development of -omics techniques, better screening is available as well.
Due to these advances, animal experiments are not longer needed for predicting human PK characteristics. Most PK factors can be studied in human *in vitro* systems, based on cellular and molecular mechanisms. Animal studies still are performed, but mainly as 'proof of concept' studies.
In early drug development, a number of factors about the metabolism and the interaction with the drug have to be known. Metabolic stability or instability of the drug is a very important factor, just as the main metabolites and the routes by which the drug is metabolised. The enzymes which catalyse the drugs during the metabolism must also be known, as well as whether there are potential metabolic interactions. Finally, genetic deficiencies and species differences also play a role.
Some *in vitro* drug metabolism screening methods include the use of recombinant expressed enzymes, sub cellular fractions, cultured hepatocytes, liver slices and tumour-derived cell lines. Immortalized cells and genetically manipulated cellular systems are used as well.

5.4.3.2.3.2. EDQM activities for Refinement of animal experiments in the field of quality control of vaccines, by Karl-Heinz Buchheit, EDQM, FR

EDQM (European Directorate for the Quality of Medicines & HealthCare) is part of the Council of Europe, who published the "European Convention for the protection of vertebrate animals used for experimental and other scientific purposes" in 1986. This was the first international legal text in this field. The work of EDQM is based on this convention.
EDQM is involved in ensuring the quality of medicines, including vaccines, for human and veterinary use. To this end, the European Pharmacopoeia (Ph. Eur.) is published. EDQM also runs the Biological Standardisation Programme (BSP), which gives information about reference standards and methods for quality control of biologicals. The 3R methods are taken up in this programme. EDQM is also part of the Official Control Authority Batch Release (OCABR), who is responsible for control on human and veterinary vaccines and plasma derivatives. Attention is mostly given to Reduction and Replacement and less to Refinement.
In Ph. Eur. it is stated that reduction of animal use and the use of alternative test methods is encouraged, if they gives satisfactory results compared with the official methods.
For both human and veterinary vaccines, Reduction is mostly discussed, followed by Replacement and then Refinement.
The general monograph recommends the use of humane endpoints wherever possible. For human vaccines for diphtheria and tetanus serological assays are performed. For a number of animal vaccines, such as *Canine leptosprirosis*, *Clostridium novyi* and *C. perfringes*, serological batch potency assays are performed.
BSP gives information on the establishment of standards and the validation of alternative methods. It contributes to international harmonisation by common projects with WHO, FDA and the Japanese Authorities.

BSP has initiated and/or concluded 98 projects, of which 30 are on methods development. 16 projects are done on the 3R methods, of which 8 on Refinement (all on vaccines), 2 on Reduction and 6 on Replacement. 68 projects deal with the establishment of BRPs (including antisera standards for Refinement).
Lethal challenge tests are replaced by serological assays for tetanus, diphtheria and swine erysipelas vaccines. Assays for acellular and whole cell pertussis are being developed. The serological vaccines for tetanus and diphtheria allows the perform assays of 2 antigens in 1 experiment and the use of a one-dilution-assay. This way, there is a significant reduction in the number of animals used. The method is described in Ph. Eur.
There are some problems in introducing the 3Rs. EDQM did a survey among manufacturers and OMCLs (Official Medicines Control Laboratory) about which measures have been implemented and what are the reason for non-implementation. Responses were received from 26 OMCLs and 8 internationally active vaccine manufacturers.
For example in the case of using a serological potency assay instead of a lethal challenge assay for tetanus and diphtheria vaccine, this is only used by 1 manufacturer and 2 OMCLs for batch release. The reasons given are that costs are too high, the validation process is long and costly, personnel and training and housing and equipment are not adapted. It is also said that it is too time-consuming and that there is a lack of global harmonisation.

5.4.3.2.3.3. Refinement in the development of biopharmaceuticals, by Gisbert Sponer, Bioassay GmbH, DE

Biopharmaceuticals is a name for the group of recombined peptides and humanised antibodies. A lot of the humanised antibodies which are tested in preclinical studies aimed for the treatment of oncological diseases.
The burden of animals in cancer research might be larger than in other areas of research, due to the fact that in cancer research the animals are subdued to painful cancer growth, toxic side effects of test compounds and immunological problems.
To do refinement in a research project it is very important that all researchers, personnel and staff are aware of this goal and are adequately trained for it. Rating scales are available, but most companies are not aware about this and thus do not use them. When planning a project, it is important to weigh benefits against the potential discomfort caused to the animals. Another way to possibly refine research, is to start with pilot studies with a limited number of animals. This way, if a certain drug or method doesn't give any positive results at all, no further studies have to be done. The number of animals sacrificed is then only small. In the case that it does give positive results, this only costs a small number of animals extra.
Defining humane endpoints is another important point for refinement in research projects. As a start, it is important to avoid survival rate as an endpoint. The use of imaging techniques is a humane endpoint, this also gives more information and can be done repeatedly. Euthanasia is also a humane endpoint. Daily monitoring is another way to look after the animal's welfare.

AAALAC certification, which can be received if the rules of accreditation are fulfilled, can help in the effort to effectively implement refinement. These rules include, amongst others that all staff working with animals is adequately trained and that facilities for animals should permit a state of well-being.
Concerning the regulatory and legal aspects, it is obvious that Animal Welfare Acts exist in all countries of the EU, however, there are substantial differences between the member states concern the execution of Acts. European harmonisation is started, but should still be extended. The revision of Directive 86/609 will help in this process. It must be taken in mind, however, that over-regulation can have adverse effects.
Finally, global harmonisation is a topic that should be taken up by the ICH. As a matter of fact, this is now in the progress of being taken up by ICH and guidelines are being published.

5.4.3.2.3.4. Discussion

When European companies want to do work together with international partners, they must be involved from the start, early in the development. For example, in Japan a lot of attention is given to animal studies. Therefore, global acceptance is important.

The new incentives, of which Karl-Heinz Buchheit speaks, are for example reducing the costs of animal experimentation. For example, to do tests for Type II Variation, companies have to show clinical data, which are very expensive and time-consuming to get. So, incentives could help convince companies to perform Refinement.
If a company replaces a stressful experiment with, for example, an *in vitro* approach, marketing this product will be easier.

Classic animal PK studies are not very predictive for the human situation. For biopharmaceuticals for example, it is not very useful, since it are proteins which are made similar to endogenous proteins. However, animal toxicology studies still have to be done, and to interpret these, PK data must be available. One new technique which could help in the future, are *in silico* test methods, they also can process a lot tests in a very short time.

Another difficult point is global harmonisation, which is a very slow process, it runs about 10 years behind on the actual scientific developments.
Refinement and Reduction are often more easily accepted than Replacement. This is especially the case for old vaccines, which often are badly analysed mixtures. One of the reasons for this is that, although a lot can be learnt from *in vitro* testing, the mechanisms of a whole animal or human still provide very useful information, which can be acquired from *in vitro* testing alone.
Also, in general a kind of first confirmation in animals is required. Even though, it is not always very predictive, as was clear from the Tegenero case, which is why it is best to start with small doses. It is also very important to interpret the data correctly, as misinterpretation of data also has been a frequent cause of disasters.

One very important comment was that the regulators do not ask to do a tick box approach, but to think and be critical about which tests one is going to perform.

5.4.3.2.4. Round Table 4: Animal models in human diseases

5.4.3.2.4.1. Refinement in non-human studies of Parkinson's disease, by Augusto Vitale, Istituto Superiore di Sanita, IT

Parkinson's disease is a pathological and fatal condition. It is a widespread disease, therefore it is important to understand the underlying causes to find a possible cure.
To this end invasive animal experiments are performed. Although rodents are used for the study of Parkinson's disease, non-human primates are used as well, because they can give more information about the underlying mechanisms in relation to motor deficits. Among non-human primates, the common marmoset (*Callithrix jacchus*), who are a good model for different diseases, and *Macaca spp.* are used.
Because of the ethical concerns in working with non-human primates, the 3Rs must be implemented wherever possible.
Refinement methodologies can be used in different aspects of the use of monkeys for the study of Parkinson's disease. For example positive training, through positive reinforcement, can applied to procedures such as injections, removal of individuals from the cage. The use of restraining chairs gives stress to the animals, but they have still to be used in the case of precise neuronal measurements. One possibility for refinement is to sit the monkey in the chair in the presence of a companion, which gives the individual a more secure feeling. However, the companion can experience some stress as well, and therefore a cost-profit analysis must be made before the project is started.
Environmental enrichments are also important in providing some sort of comfort to the Parkinsonian monkey. However, because the ill monkey cannot move properly. It has to be assured that the enrichment do not require special motor skills to be properly exploited.
Performing multiple recordings of neuronal activity in one session can reduce the number of animals, and reduce the number of recording sessions required as well. A tight recording session schedule will reduce the stress caused to the experimental animals, which will not have to habituate to the procedure again every time is tested.
Finally, telemetry might be another good way of refinement, if it is possible to collect data without having to handle the animals, less stress is caused.
Refinement is the easiest of the 3Rs to implement in PD studies utilising non-human primates. All phases of the life of the monkeys must be improved: pre-, during and post-treatment. This effort will result in an improvement of the quality of the life of the animal and of the quality of the research.

5.4.3.2.4.1. Application of Refinement in transgenic mice, by Igor Branchi, Istituto Superiore di Sanita, IT

Since 1995 the number of genetically modified animals used, has increased a lot, but the total number of animals used in all areas of research has decreased.
Of all the genetic animals used, mice comprise 68%. The reason for this is that mice are the easiest species to modify.
However, there still are concerns about the potential risks in using genetically modified mice. In consequence, it is uncertain which the effects on the animal's welfare is. But it is important not to forget that in the first place genetically modified mice are still mice and thus the same welfare concerns as with normal mice should be taken care off.
So, even when working with transgenic mice, the 3Rs must be implemented. If it is possible to replace a certain test with an alternative method, it must be done, no matter whether the tests are performed on transgenic mice or not.
In some cases, such as with obese, nude or green fluorescence mice, further considerations must be made. But most of the genetically modified animals show no notable physiological and behavioural alterations. An example of this are the SCID (Severe Combined Immune Deficiency syndrome) mice.
Transgenic mice might have unexpected effects due to the genetic manipulation, alterations in pain threshold, sensitivity to psycho-social stress, which can lead to a high susceptibility to infections, and subtle metabolic can appear. It is important to know this, to rightly interpret the results. To this end, score sheets are available, to quantify how the animals react to opening of the cage and handling. It then can be deduced whether these reactions are normal or not. The question must then be posed what is normal for a mouse line, since each strain has its own characteristics anyway. It is probably most meaningful to compare the reactions of the transgenic animals to those of wild type animals.
There are several ways to improve the conditions for the mice. In housing, reducing noise and crowding are good ways to reduce stress. Background strains can be changed to improve the breeding performances and to lessen the effects of a severe phenotype. Breeding strategies can also be changed, for example by maintaining heterozygous breeding pairs. Care takers can be trained. Finally, humane end points are another way of refining the conditions for the animals.
To conclude it can be said that genetically manipulated animals represent a powerful way to understand the genetic bases of a phenotype or to dissect the complexities of a phenotype. General principles of bioethics can be applied to transgenic animals, however special attention should be paid to unexpected effects.
A comment from the public was that the first 12 days of a mouse's live can serve as a window period, because then the animals are still developing. Then unexpected behaviour can be published, which would be interesting information for everybody working with the same strain. The speaker then told about a strain of mice, who were unable to learn, so it was thought that they were knock-out for intelligence. Only after a while it was found out that they were blind. This is a good example of how important it is to know the unexpected side effects.

5.4.3.2.4.2. Refinement in Multiple Sclerosis modelling, by Valeria Muzio, Merck Serono, IT

Multiple sclerosis as a complex disease, which is known for a heterogeneity of clinical forms, with inter-individual variability in its progression. Various forms of CNS lesion and pathogenic mechanisms can be involved. From a certain point in its development, after reaching a certain threshold of axonal loss, the disease is irreversible.
Merck Serono's strategy for developing disease modifying therapies is performing a translational research from the clinic to the preclinical field. This approach incorporates data from genetic, genomic and proteomic analyses and imaging techniques from the different Multiple Sclerosis disease phenotypes. The hallmarks of these data are used to select relevant animal models. Taken together with molecular profiles of the compounds this leads to targeted pathways in *in vivo* animal models.
This disease runs in phases, starting with an inflammatory phase, followed by a demyelination phase and finally going into axonal loss and neurodegeneration.
Genomic approaches are used to validate the target. These biomarkers and surrogate markers are then used to apply clinical relevant read-outs on the relevant models. Preclinical imaging, MRI with clinical non invasive biomarkers is applied.
The exposure range obtain from the imaging techniques is used for a PK/PD analysis and the establishment of a safety range.
As a conclusion, the different methods of Refining the Multiple Sclerosis modelling were summarized. First of all, it is important to gain more pertinent information at the lowest possible cost of animals well-being. The animal model is carefully chosen taking account of the current state of knowledge, which concerns both the pertinence of the choice of species and the protocol itself. Planning the protocol correctly helps to avoid any disturbances likely to establish the earliest end-points possible on the basis of the experiment's objective. Preference must be given to non-invasive procedures, such as imaging and telemetry. Appropriate care must be given at all times and if necessary anaesthesia/analgesia must be used. Human endpoints must be established beforehand.

5.4.3.2.4.3. Refinement in animal epilepsy models, by Ralph Clinckers, VUB, BE

In epilepsy research there are several objectives, such as unravelling the pathogenesis of epilepsy, identifying new drug targets and drug screening. To this end, two models are used, the seizure and epilepsy models. Seizure models have been used for years and are still used in the study of epilepsy. They have been very helpful in understanding the basic mechanisms of epilepsy and for screening new antiepileptic drugs. Nowadays, antiepileptic drugs work on suppressing the seizures of patients. However, in a number of cases, the drugs do not work and the seizure persists, so there is a need for drugs which completely prevent epilepsy. For this purpose, animal epilepsy models are used.
A number of techniques is used to maximize the number of read outs obtained from an animal in an acute seizure or chronic epilepsy model, which thus helps to refining the research to this disease.

In vivo microdialysis works on the principle of kinetic dialysis. *In situ* administration tools are used for proconvulsants and drugs. Proconvulsants are used in the focal pilocarpine model for complex partial seizures as it generates an epileptic focus. Samples are taken i*n vivo*. This is done for endogenous compounds, such as neurotransmitters, metabolites and neuropeptides. This gives information about the mechanisms of action of antiepileptic drugs and the pathophysiology of epilepsy. It can also be used for exogenous compounds, such as antiepileptic drugs, which then gives information about the biodisposition of AEDs (Anti-Epileptic Drugs) in the brain, blood-brain barrier penetration and pharmacokinetics.
Pharmacokinetic/pharmacodynamic modeling is also used in epilepsy research. An integrated pharmacokinetic model using nonlinear mixed effects modelling is developed. PK parameters are derived from a model, for the overall population as well as the individuals. Parameters are estimated in terms of fixed and random effects. Covariates affecting PKs and intra- and interindividual differences are identified. It was discovered that seizures affect brain disposition of AEDs.
In seizure models EEG (Electro Encephalography) monitoring is performed. When PK/PD modelling is done, with the EEG parameters, this can be used for predicting the progression of the disease. For example, an increase in β-activity (11.5 - 30 Hz) of the EEG reflects facilitation of GABA-ergic (Gamma-aminobutyric acid) inhibition; i.e. biomarker for the effect of GABA-ergic ligands. Possible explanations for this are the loss of neurons or $GABA_A$ receptors.
Imaging can be used in combination with PK analysis. In this case, quantification of alterations in $GABA_A$ receptor density in epileptic rats using positron emission tomography (PET) is performed. Simultaneous population PK analysis of blood (LC-MS/MS) and brain (PET) concentration-time profiles of flumazenil are performed. This shows that both the $GABA_A$-receptor density, and possibly also the blood–brain barrier transport of flumazenil are altered after kindling.
Magnetic resonance imaging (MRI), to monitor histological artefacts and epilepsy-related brain damage, and Single Photon Emission Computed Tomography (SPECT) (+ microdialysis), monitor blood-brain barrier permeability changes and identify biomarkers, are done to evaluate the efficacy of anti-epileptogenic interventions.

5.4.3.2.4.4. Discussion

A general question was whether, since techniques to really measure pain, they are actually used. The general response was that they could be used, although they're quite costly and therefore most of the time not used. Monitoring can be used to see whether an animal is in pain or not. From the behaviour of animals it can also be known whether they are in pain or not. Ralph Clinckers told that at his lab the tests are stopped, if the suffering can't be stopped. In other tests, like for example the cases reported by Augusto Vitalo, it is known that certain procedures cause some pain, but there is (currently) no way to avoid this.

Genetically modified animals are already used. For example, the use of knock out mice is already routine when looking for new targets. However, humanised transgenic animals are not that popular yet, but they will most probably be used more in the future.

A remarkable comment was that science might have to enlarge its attitude, since, even without the most appropriate model, science can develop further.

5.4.3.3. Final Discussion

To start the final discussion, some points were summarized first.
One comment was that the way to view experiments is to cooperate with the animals, this way animal welfare gains importance. One example of cooperation with the animals is PRT.
Progress is being made, but a lot of work still has to be done. New techniques and methods are being developed and are already in used, but often it is not certain how to implement them. There also are barriers, like time and money.
It is important that everybody dealing with animals (breeders, animal caretakers, researchers…) all are educated on how to handle animals. The educational demands are also gaining importance.
Global harmonisation is, although difficult and slow, important to be pursued.
Permanent ethical committees will control whether animal welfare is taken into consideration when performing animal tests.
National Reference Labs will be started to validate alternative test methods.
Environmental enrichment does not disturb standardisation of tests.
Finally, Refinement can improve the marketing strength of companies, since there will be more public acceptability.

Vaccines are not always taken into consideration, while there still is a lot of work to be done in that field. One problem is that nobody seems to know who to turn to, to talk about this and make sure more work is done. This is a task that *ecopa*, for example, could take up, to spread information about where one has to go for new innovative ideas.
For biopharmaceuticals a lot is in the pipeline, however the techniques are not fully established and agreed upon. Here the question is also how these techniques should be looked upon in the light of the 3Rs.
Other techniques that are interesting to be used more in the future include the use of genetically modified animals and more human *in vitro* cell cultures.
On the other hand there are a lot of methods available, which are not used. For this problem, ideas must be found to make people use them. For companies, a reward in some form, for investing in new techniques and methods, might be interesting.

Concerning Global Harmonisation, it was mentioned that *ecopa*, amongst other groups, is active, but that there is no certainty whether the message spreads in the rest of the world.

The question is also whether the other parts of the world are interested in this? Does it make sense to demand Global Harmonisation if the rest of the world is not interested?

Another interesting discussion was held on the topic of replacing one species with another for a certain test. If for certain tests, monkeys are replaced with minipigs, this can not really be called Replacement, however, it can be Refinement, if the new species is indeed less sentient. There might be the risk that the data are of less quality, because the species is less related to the human. For example for cardiovascular research, this might not make a difference, since there is a large similarity between humans and pigs in this organ. In the view of animal welfare activists, it might not make much difference which species of animal is used. It is also not really clear whether a species feels less pain or not, since it is expressed differently by each species.

Whether Refinement always leads to better science is another point of discussion. As became clear from the presentations and former discussions, there will be cases where it indeed is better, for example thanks to environmental enrichment and thus healthier animals. In other cases, too healthy animals, might not respond to certain treatments.
In other cases, there might only be benefits for the animals if Refinement is done, but the quality of data might stay the same. In cases such as these, must the Refinement procedures then be followed or not?

It was argued that more attention might be given to morality in research, this is especially interesting for the marketing of products. Other than that, the main barriers within companies are time and money. There are also global differences in morals, linked to culture, religion… Therefore, there should actually be one consensus in the world, which goes beyond the actual moral discussions.
What might be more feasible is to find very good definitions on pain and distress, since there is little discussion on this, and connected to this, the use of anaesthetics.

Finally, it is necessary to always compare the risks and benefits, for both the companies and the animals, before performing certain tests. A question that comes to mind here is also, whether people are willing to use drugs that are not fully tested for safety if this way animals can be saved.

Workshop Reduction

3 – 4 July 2009, Hotel Grauer Bär, Innsbruck, AT

5.5. Workshop Reduction

5.5.1. Executive Summary

The second Workshop on Reduction was concerned with the question of how to reduce the number of animals during the whole drug development process. It was found that different possibilities exist at different stages of the process.

Reduction was defined as “any approach in scientific research, product testing or education that leads, directly or indirectly, to a decrease in the number of animals used while meeting the scientific requirements of the project”. This makes clear that Reduction can only be performed, when there is no danger that the quality of the data will decrease. Reduction can be reached, for example, by using stronger statistic analytic techniques or by getting more data out of a single animal.

Some important ideas, focused on Reduction, are summarised here.

- Reduction possibilities for new chemical entities (NCEs):
For NCEs, during early drug development *in vitro* and *in silico* methods are well applied. It is in particular in the more advanced development phase that animal reduction is possible. *In vitro* testing and alternative methods in general can help to reduce the number of animal studies by helping to decide on the right drug candidates. Important fields for reduction are genotoxicity (e.g. *in vitro* micronucleus), pharmacology (e.g. receptor binding studies), and toxicology (e.g. zebra fish eggs in reproductive and developmental toxicity; embryonal stem cell test (EST) in early embryotoxicity screening). Optimisation of *in vitro* tests in these fields have priority.

- Reduction possibilities for new biological entities (NBEs):
For NBEs, several parameters could be combined in one animal study, performed on a relevant species e.g. safety pharmacology, pharmacokinetics, local tolerance and immunogenicity. Furthermore, standardisation is of crucial importance (animal strain, environmental conditions, techniques…). Transgenic animals and humanised rodents are often used.

- Special case of vaccines:
High numbers of animals are used in the quality control of vaccines since in the EU every batch must be tested. Moving from this traditional quality control approach of safety and efficacy to a so-called Consistency Approach of releasing a whole batch when all critical steps during production have been monitored and confirmed product consistency, could save a lot of animals. Global harmonisation, however, remains critical.

- Better integration of in silico and in vitro data:
More user-friendly and comprehensive mechanistic platforms are necessary for integrating ADME models and databases to optimise clinical trial design, predict drug exposure and interactions in different populations. More attention is needed for identifying and determining the data and their standardisation to populate the models envisaged.

- Availability of negative results and data sharing:
The importance and availability of non-public data and negative results was discussed together with issues such as intellectual property, confidentiality and quality of these data.
It was considered that more efforts should be done in between companies to share information and that coding of substances by a neutral body (super party entity) that could guarantee confidentiality seems highly relevant. In addition, quality control of data is crucial together with the standardisation of protocols used.

- Non-invasive diagnostic methodologies:
Transfer of non-invasive diagnostic methodologies e.g. magnetic resonance imaging (MRI) from human medicine to laboratory animals allows long-term monitoring and could represent a powerful tool in animal reduction, in particular when different non-invasive techniques are combined (e.g. in case of respiratory disease models where experimental methods are very invasive).

- Need for further qualification of more biomarkers:
Non-invasive biomarkers are the only ones that can be translated into a clinical setting. Further detection and also qualification of suitable biomarkers for different species is necessary in order to allow translation from animal to man. Species such as dog are actually not well covered. This is a resource intensive work and the potential to reduce animal numbers only works on the long term.

- Better controlled production and use of transgenic animals:
In basic research, animals used often are transgenic. It was emphasised to leave the development and surveillance of transgenic animals to quality controlled and certified specialised facilities. Indeed, to produce transgenic animals in only a few well-structured centres with high qualified personnel could result in smaller numbers of animals needed to generate and breed transgenic animals.
Furthermore, the use of switch-able gene constructs may further reduce the number of animals in the maintenance of transgenic animal strains. Microbiological quality control is important in animal reduction. Indeed, clinically silent infections of experimental animals could remain undetected and be at the origin of modified study results. It was mentioned in particular happening at universities and research centres and less in pharmaceutical industry (e.g. parvovirus in mice). It was generally agreed that the use of microbiologically standardised animals is a basis concept in animal reduction.

5.5.2. Recommendations

1. Priority for test development/optimisation should be given to those *in vitro* methods that can be easily applied in the later stages of drug development to eliminate toxic substances and as such screen more efficiently for interesting new compounds e.g. zebra fish eggs and embryonic stem cells test (EST) for embryotoxicity.

2. Application of non-invasive diagnostic methodology, alone or in combination, should become common practice, in particular in those cases where invasive animal disease models are still in use.

3. Data sharing between research institutes/companies/academia should be substantiated by the creation of a pan-European neutral body, guaranteeing confidentiality, data quality and standardisation.

4. In vaccines production, monitoring of all critical stages should become the quality control rule rather than traditional animal testing of batches.

5. In the development of biologics, *in vivo* safety testing should only be performed when human relevant species have been identified. Human-based *in vitro* screening (pharmacology and safety) should be applied wherever feasible and appropriate.

6. Transgenic animals should be produced only for relevant purposes and in a limited number of well-controlled specialised facilities of high-quality standard.

5.5.3. Report of the Workshop Reduction

START-UP Workshop Reduction
3 – 4 July 2009, Hotel Grauer Bär, Innsbruck, AT

Program:

3 July

14:00 – 14:30	**Welcome and Introduction**
14:00 – 14:10	Welcome Address. Walter Pfaller, zet/Innsbruck Medical University, AT
14:10 – 14:20	*ecopa*: Realisations and Future Perspectives. Vera Rogiers, *ecopa*, BE
14:20 – 14:30	Conclusions on *Reduction* from the Anim.Al.See-Project. Flavia Zucco, IPAM, IT
14:30 – 17:15	**Session I: Bottlenecks in the Development of NCEs and NBEs (new chemical and biological entities)**
14:30 – 14:45	In vitro Prediction of Side Effects of Newly Developed Pharmaceuticals. Joachim Coenen, Merck, DE
14:45 – 15:00	Animal Reduction in the Investigation of Embryotoxicity Studies of Candidate Drugs. Femke van de Water, Schering-Plough, NL
15:00 – 15:15	"Reduction" in the Development of Biologicals. Werner Höllriegl, Baxter, AT
15:15 – 15:30	Reduction in Animal Testing in Vaccines for Human Use. Johan Descamps, GlaxoSmithKline Biologicals, BE
15:30 – 15:45	Discussion (Part I)
15:45 – 16:15	Coffee Break
	Session I (continued)
16:15 – 16:30	Early Screening Methodologies in Pharmaceutical Research. Katja Conrath, Galapagos, BE
16:30 – 16:45	Use of Modelling and Simulation in the Prediction of inter-individual Variation in Human Pharmacokinetics: Potentials for Reduction of Animal Studies. Zoe Barter, Simcyp, UK
16:45 – 17:00	Assessment of Current Guidelines and Facilities to Reduce Animal Testing in the Development of Pharmaceuticals. Jay Iyer, TI Pharma, NL
17:00 – 17:15	Discussion of Session I (Part II)
17:15 – 18:30	**Session II: Reduction Potential by Handling Statistics and Experimental Data**
17:15 – 17:30	Systematic Approach to Tune Number of Subjects to Scientific Hypothesis. Peter Vijn, Schering-Plough, NL
17:30 – 17:45	Data Mining Tools for the Reduction of Animals for Adverse Effect Testing of Pharmaceuticals. Christoph Helma, in silico toxicology, CH
17:45 – 18:30	Round Table: Importance and Availability of Non-Public Data and Negative Results

4 July

09:00 – 10:30	**Session III: Reduction via Study Design and Methods in Drug Research and Development**
09:00 – 09:15	Non-invasive Small Rodent Imaging: Specific Examples and Contributions to Reduction in Drug Research and Development. Nicolau Beckmann, Novartis, CH
09:15 – 09:30	The Principle of Reduction of Animal Testing in non-clinical Safety. Stefan Platz, Roche, CH
09:30 – 09:45	Reducing the Number of Animals Used in Antibiotic Discovery. J. Barry Wright, AstraZeneca, UK
09:45 – 10:00	Tissue-Specific, Non-Invasive Toxicity *Biomarkers*. Patrick Müller, Novartis, CH
10:00 – 10:30	Discussion of Session III
10:30 – 11:00	Coffee Break
11:00 – 12:00	**Session IV: Transgenic Animals in Pharmaceutical Research and Drug Development**
11:00 – 11:15	Reduction in Generating and Maintaining Transgenic Animals. Thomas Rülicke, University of Veterinary Medicine, Vienna, AT
11:15 – 11:30	Transgenic Animals in Pharmaceutical Industry: Contribution to Reduction. Rainer Nobiling, Heidelberg, DE
11:30 – 11:45	SPF-Housing in Relation to Reduction. Werner Nicklas, DKFZ, DE
11:45 – 12:00	Discussion of Session IV
12:00 – 12:30	**Plenary Discussion and Summary**

Scientific committee:
Janna de Boer, member of ZonMW, NL
Bernward Garthoff, treasurer *ecopa*, Bayer, DE
Gerhard Gstraunthaler, member of NCP ZET, Innsbruck Medical University, AT
Peter Maier, Board Member of *ecopa*, Forschung 3R, CH
Walter Pfaller, Board Member of *ecopa*, Innsbruck Medical University, AT
Vera Rogiers, chair of *ecopa*, VUB, BE

5.5.3.1. Introduction to the workshop

This workshop is the second one in a series of three, each of these is concerned with one of the R's in the 3Rs concept of Russell and Burch, namely "Refinement", "Reduction" and "Replacement". This workshop is dealing with the second R, namely the issue of Reduction. Walter Pfaller (member of the NCP ZET, AT) gave a welcome address and introduction to the general aim of and ideas behind this workshop.
Vera Rogiers, chair of *ecopa*, presented the structure and aims of the START-UP project and explained the role of the Expert Meetings and their input in this second Workshop.

5.5.3.1.1. Introduction lectures

5.5.3.1.1.1. Conclusions on Reduction from Anim.Al.See. EU project, by Flavia Zucco, Istituto Neurobiologia e Medicina Molecolare, IT

The Anim.Al.See project had the goal to update the view on 3Rs models, both from a philosophical and scientific point of view. Among others, new definitions of the 3Rs were made.
The definition for reduction proposed by this project is: "Reduction refers to any approach in scientific research, product testing or education that leads, directly or indirectly, to a decrease in the number of animals used while meeting the scientific requirements of the project."

5.5.3.2. Summary of the presentations within the different sessions

5.5.3.2.1. Session I: Bottlenecks in the Development of NCEs and NBEs (new chemical and biological entities)

5.5.3.2.1.1. In vitro Prediction of Side Effects of Newly Developed Pharmaceuticals, by Joachim Coenen, Merck Serono, DE

Legal requirements regarding the preclinical developmental process of new drugs include proof of pharmacological efficacy and safety *in vivo*, therefore pharmacology and toxicity testing must be performed. Apart from European guidelines, also international guidelines exist. Since drugs are developed on a global scale, the efforts of the ICH (International Conference on Harmonisation) are very important in reducing the number of animals, although there still is room for improvement. However, the M3 guideline requires the performance of animal toxicity studies in two mammalian species, one of which must be a non-rodent.

When looking at attrition rate and choosing the right candidate, a general pattern seems to return. In basic research, one starts up with about 10 to 30.000 entities and eventually ends up with one compound on the market. Saving animals is not possible at the level of early basic research, because this is mostly done already with using *in vitro* or *in silico* studies. So, most animals could be saved in the advanced developmental part.
Before the first studies are performed in humans, it is required to do animal studies and also during the clinical development stage, accompanying animal studies are performed.
The first phase, being screening and early toxicology, is mainly done using *in silico* toxicology. QSAR (Quantitative structure-activity relationship) is a technology, which is capable to help in deciding whether a compound represents the best candidate for development.
In the next phase, tests such as genotoxicity, safety pharmacology, and single and repeat dose toxicological studies come in. Most of the genotoxicity studies are quite old, such as the Ames and bacteria tests, they can be regarded as replacements, but they help also to reduce animal studies. Chromosome aberration and micro-nucleus tests can be done *in vitro*, but it is also required to perform one *in vivo* genotoxicity study. Thus, *in vitro* studies help to decide on the right candidate and therefore reduce animal tests.
Similarly, in safety pharmacology and in particular in cardiovascular studies, animal studies are performed, but can be further reduced. For example, by using simple receptor binding studies, electric field potential studies and automated electrophysiology. With these tests, for example, 100 candidate compounds can be reduced to 10. Then with more specific studies, such as papillary muscle refractory time, the number of compounds can eventually be reduced to one compound, which will then go into *in vivo* studies.
During toxicity studies, *in vivo* tests still have to be performed. Thanks to the aforementioned tests, however, these can start with the right dose and thus the number of animals is reduced.
Before clinical development starts, *in vivo* studies have to be performed, as it is required by the Health Authorities, but *in vitro* tests can accompany these studies, for example in the fields of developmental, reproductive and mechanistic studies. For reproductive toxicity studies, the fish egg assay is a well established *in vitro* model (drafted in OECD guidelines). It can also be used as a screening assay for developmental toxicity. Advantages are that a huge number of eggs are available at the same time and the turn-over is fast. However, proteratogenic compounds are not detectable. Merck developed a new test that includes a metabolic activation system so that teratogenicity can be demonstrated. So far, this has been successfully performed with four proteratogenic substances, being cyclophosphamide, ethanol, benzo[a]pyrene, and thalidomide. Of course there still are some challenges.
‘Omics will, in the future, help predict side-effects in preclinical studies. But at present they are mostly used as a mechanistic instrument.
In conclusion, the pharmaceutical industry already uses a lot of *in silico* and *in vitro* tests to predict side-effects and to select drug candidates. One of the bottlenecks still is harmonisation in guidelines.

5.5.3.2.1.2. Animal Reduction in the Investigation of Embryotoxicity Studies of Candidate Drugs, by Femke van de Water, Schering-Plough, NL

The Embryonic Stem Cell Test (EST) might be able to reduce animal experimentation. Regarding reproductive medicine, the main part of drug failure due to toxicity was caused by reprotoxicity. Therefore, it is interesting to develop an alternative test method, that could be introduced early in the screening. This may save the costs because it is expensive to delete drug candidates after relatively late performed *in vivo* studies and reduce animal testing.
There are different alternative test methods for embryotoxicity screening, such as the Rat Whole Embryo Culture, Rat Micromass Assay and Zebrafish Embryo Teratogenesis Assay. The EST does not use animals at all. In one embryonic stem cell line, the cells do not differentiate, if Leukaemia Inhibitory Factor is added. If this factor is withdrawn, the cells differentiate into beating cardiomyocytes.
The EST consists of two parts. The first part is the differentiation assay, which gives information on embryonal development. The second part consists of cytotoxicity assays, performed on embryonic stem cells providing information about general embryotoxicity and on a 3T3 cell line, giving information on maternal toxicity. In the differentiation assay, hanging drops of ES cells are exposed to different concentrations of test compounds, after three days embryoid bodies are formed. After ten days, these cells can be seen beating in multi-well plates. When exposed to different concentrations of the test compound, ID 50 values can be obtained. In the cytotoxicity assay, cells are seeded in a well plate and after several re-exposures, an MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphemiltetrazoliumbromide) test is performed on day ten. From both types of cells an IC 50 value can be determined . In the prediction model, the function with the highest value determines to which class the compound belongs. If function 1 contains the highest value, then the compound is likely to be non-embryotoxic. When function 3 is the highest, the compound is probably strongly embryotoxic.
The EST was scientifically validated by Genshow et al. 2004. 22 chemical compounds with well defined *in vivo* embryotoxic potential were tested in four different laboratories. No pharmaceutical compounds were included. Tested at Organon, the predictability was quite good for non-embryotoxic compounds. The precision however was not good for non- and strong embryotoxic compounds. All came out as weakly embryotoxic.
A lot of bottlenecks in the EST were identified. In the differentiation assay, only a single differentiation route was detected, the sensitivity of ES-D3 cells to different culture methods is unclear, the scoring method is subjective; degree of beating not taken into account, there is a possible bias by direct action of compound on cell contraction (e.g. diphenhydramine, A.K. Peters *et al.*, 2008) and there also are problems with the exposure schedule: it is unclear whether the effect measured is on proliferation or on differentiation. There are also some problems with the prediction model, it is not validated for receptor-mediated pharmaceutical compounds, the prediction of relatively poorly soluble compounds is incorrect; most strong embryotoxic compounds in the validation study were extremely cytotoxic/mutagenic and the non-embryotoxic compounds were not cytotoxic at all. The prediction model is based on µg/ml concentrations. Then there are also some other issues, such as translation to *in vivo* and

human situation, the fact that there is no discrimination between embryotoxicity and teratogenicity and a lack of well defined *in vivo* data.
To conclude, the current EST protocol needs further optimisation before it can be successfully implemented in pharmaceutical industry. Successful implementation may prevent drug attrition in a relatively late phase of drug development and hence reduce animal experimentation. In the near future, the main focus should be on extending the general knowledge on stem cell development, improvement of the current protocols, obtaining well-defined *in vivo* data and the development of an accurate prediction model for pharmaceutical compounds.

5.5.3.2.1.3. "Reduction" in the Development of Biologicals, by Werner Höllriegl, Baxter, AT

Focus here lies on plasma-derived recombinant therapeutic proteins.
In the past these compounds were derived from human plasma. Nowadays, a CHO (Chinese Hamster Ovary) cell line is used. It is possible with new production strategies to develop recombinant analogues of plasma proteins and modified therapeutic proteins.
There are analytical problems for these proteins in healthy volunteers, because they have the same characteristics as the native proteins. On the other hand, in dogs with canine haemophilia, therapeutic proteins for the human disease can be tested.
Before those investigations will be performed, extensive biochemical and *in vitro* characterisations have to be carried out. In the past, a considerable number of tests had to be done in dogs, but by now, more and more genetically modified mice are used.
The number of animals used in such investigations is influenced by the strategy, which should aim at including as many parameters as possible one single study. In well designed studies assessment of safety pharmacology, pharmacokinetics, local tolerance and immunogenicity can be done in one study. Another way is to try to reduce scattering in pharmacokinetic and efficacy studies, this can be reached by standardising mouse strains, experimental environmental conditions, and experimental techniques. A last option is to reduce the number of highly sensitive animals such as dogs by replacing them by appropriate transgenic mouse models.
An example of standardising procedures is with the Tail-Tip-Leading Model. Mice are anaesthetised, the tail cap is clipped off and the blood can be withdrawn. They are on a heating pad so that every animal has the same temperature, since this has an enormous effect on the diameter of blood vessels.
Another possibility in pharmacokinetic studies is to use automated blood-sampling machines, but this is not easy in mice. The catheter must continuously be perfused with isotonic saline in order to prevent clotting. This method is capable to reduce the number of mice, if used properly.
An important issue with therapeutic proteins is their immunogenic potential when derived from bacteria. The protein undergoes modification. If a patient is exposed for a period of time to such a protein. it is possible that the patient develops neutralizing antibodies against the protein with the consequence that it becomes biologically inactive. Animal models therefore can only be used to assess relative immunogenicity, because there are

immunological differences between animals and humans: for example, when a human protein is injected into mice, identified as a xenoprotein for this species. This problem may be solved by humanisation of the murine immune system. This can be done by translation of human genes into mice with the consequence that the human protein is no longer identified as a xenoprotein.

5.5.3.2.1.4. Reduction in Animal Testing in Vaccines for Human Use, by Johan Descamps, GlaxoSmithKline Biologicals, BE

Animals are often used in vaccines testing, in the basic research phase, candidate testing, preclinical phase and in the production to validate the process. Furthermore, regarding quality control of vaccines, each batch has to be tested. This is not only a requirement within Europe, but also when importing vaccines in, for example, from Korea. This means that vaccines can be tested on safety up to four times. So, in vaccines, most of the test animals are used for the quality control phase. This is completely different compared to classical pharmaceuticals, where most animals are used in investigations during research and developmental phases. Animal testing represents 50 – 70% of the total production cycle of a vaccine. The testing time frame *in vivo* would take up to 70% of the time. *In vivo* testing is thus a major limiting factor in the development of vaccines.
Besides ethical reasons, there are also other reasons to look for alternatives to *in vivo* testing. For example, the fact that *in vivo* tests have a high variability, additionally, *in vivo* testing is also very time consuming.
The 3Rs have been applied in this field, examples for this are the implementation of humane endpoints, the move to single dilution *in vivo* assays for diphtheria and tetanus, the *in vitro* potency assay for Inactivated Polio Vaccine, a move from monkey NVT (Neurovirulence Test) to transgenic mice in polio and finally, for all new vaccines it is tried to avoid *in vivo* tests for potency testing. Improved statistics can also help, although in some cases, the requirements of statistic rules can increase the number of animals. By increasing the sizes of the produced batches, less tests have to be done for the respective step. On the European level, the national control laboratories now have a procedure for reducing *in vivo* testing.
A new approach for routine commercial release is the Consistency Approach. Using this procedure each batch of a product can be regarded to have the same quality within the given limits of the specifications. Thus, all the clinical trial lots can be regarded to be safe, efficacious and stable. Quality control testing is aimed at confirming product consistency by monitoring all the critical steps during production rather than relying on control of some samples as the sole indicator of safety and efficacy. This procedure is being discussed, but is not always yet implemented. If the principle of moving from the traditional quality control approach of safety and efficacy to the batch release principle of consistency is generally accepted as a proof of quality, then *in vitro* tests are more powerful tools.
An example of this approach is the inactivated polio-vaccine, the *in vitro* assay for potency, called D-antigen assay. The D-antigen assay is able to detect a degradation of

the vaccine much sooner than the *in vivo* test, when exposed to high temperatures or for instance to thiomersal. This procedure has now been approved and has been introduced in the European Pharmacopoeia. The basic idea is that the same principle can be used for many other vaccines, for instance Diphtheria, Tetanus and acellular Pertussis vaccines.
The prerequisites for acceptance of *in vitro* testing are that *in vitro* tests do not measure the intended outcome of the vaccine i.e. the immune response but are rather a marker of consistency and that alternative tests have to guarantee the highest standards of quality. The main advantage of *in vitro* testing are that the methods are powerful and reproducible, allows quantitative figures, they can be used appropriately for multiple antigens, they inexpensive and saves animals.
However, there are also some limitations. *In vitro* tests will not be available in a short time due to divergence in regulatory authority requirements, the lack of a global harmonised environment, the ongoing development and validation, lack of collaboration between various users for test development and the currently limited scientific and technical knowledge. Also, when developing new vaccines or changing major factors, there is a request by regulatory authorities for *in vivo* testing.
A technical platform should be set up for discussion between regulators, control authorities, manufacturers and international organisations. The goal would be to combine efforts and technical know-how for a common objective targeted to development of tests applicable to several vaccines. Ideally, this discussion would be chaired by the control authorities. Funding is needed for developing alternative tests for classical vaccines. A timeframe should be defined in order to make progress in the next future

Discussion

Answering the question (of Jon Richmond), whether the lack of funding might have to do with the fact that ECVAM is more involved in cosmetics and chemicals at the moment, Mr. Descamps answered that vaccines and biologicals in general are suffering a bit under a low visibility at the moment.

Q: Vera Rogiers (VUB) asked each speakers which test he or she would change and where his or her priority would lay if there were no restrictions.

A: Descamps responded that he would focus on products that already are on the market, for which lot of animals for routine quality testing are being used, and to develop alternative *in vitro* tests for these.
Van de Water responded that the EST may be very interesting, but probably, it would not work sufficiently in the future, so her focus would be on the zebrafish assay.
Coenen and Höllriegel stated that industry is on the right track, efforts should be combined to find the right battery or right set of alternative methods that would eventually at least reduce the number of animals tested.

Q: It was asked to Mr. Höllriegl regarding biologicals whether the number of animals would decrease , kept constant or maybe even increase when humanised mice will be used.

A: The main goal is of course to provide the patients with better and safer medicine. The numbers will probably stay in the same range, but hopefully there will be possibilities to decrease the numbers.

A discussion was held on who had to take the leading function in the search for alternatives. Should industry search for alternatives, regardless of the demands of the regulatory authorities and push the authorities to accept these methods when they are finally developed? Or should the regulatory authorities push the industry to look for alternative methods? It became clear that in the past both situations have already taken place. No general consensus could be reached, however, except for the fact that both regulators and industry need to know each others needs and requirements.

5.5.3.2.1.5. Early Screening Methodologies in Pharmaceutical Research, by Katja Conrath, Galapagos, BE

The focus of this presentation is on the early stages of the drug discovery process.
A first step in the drug discovery process is to identify the human protein responsible for the disease. A generally recognised powerful tool for determining protein function in a specific biological or disease pathway is RNA interference (RNAi). Target-specific RNAi can be introduced into mammalian cells either by the transfection of synthetic small interfering RNAs (siRNAs) or by viral delivery of short hairpin RNAs (shRNAs) that are subsequently intracellularly processed into siRNAs targeting a specific mRNA thereby leading to gene silencing. However, within human primary cells, viral siRNA delivery is more efficient. It also provides long-term expression of the siRNA (example was given for 1 week post infection) which allows for experiments with longer read-out times. In addition, adenoviruses specifically have the advantage that they do not integrate into the genome (unlike retroviruses and lentiviruses) and therefore there is no risk of altered cell phenotype as a result of genome mutation.
Within Galapagos, a biologically-driven target discovery and validation strategy is used based on the use of human primary cells and arrayed recombinant adenovirus production using the PER.C.6 cell line. Since the PER.C.6 cells produce adenoviral vectors without any replication-competent or wild-type adenovirus contamination, they can be safely used for the production of adenoviral libraries in microtitre plate formats. The produced adenoviruses cause lysates of the PER.C6 cells and the supernatant is subsequently put in a high throughput screening using human primary cells. These cells allow the identification of human targets of interest in a more *in vivo*-like situation than is the case with animal or tumourgenic cell lines. As such, arrayed adenovirus platforms have been generated at Galapagos that can deliver either shRNA (SilenceSelect®) or full length cDNA (FLeXSelect®) to effectively knock-down or knock-in human genes in primary cells, respectively. The libraries focus on different small molecule tractable (druggable)

human targets. Galapagos mainly applies their adenoviral technology platforms for the identification and validation of targets in bone and joint diseases, osteoarthritis, rheumatoid arthritis and osteoporosis. However, the strategy has also proven to be successful for other therapeutic areas. The target discovery strategy followed by Galapagos was exemplified with a case study on osteoarthritis.
The next step after target identification and validation is to find chemical entities or proteins that bind to the respective protein target. The 'critical path' followed is mostly the same for all companies, i.e. interesting compounds are selected by screening compound libraries using biochemical and cellular target-based assays. The new 'hits' then go through a specificity screen, also to look for close homologues of the target. Next, *in vitro* ADME and disease relevant cellular assays are performed. In addition, within Galapagos, new hits are tested on rodent ortholog targets, either by biochemical assays using rodent recombinant proteins or functional cellular assays using transiently or stably transfected cell lines. As an example, Galapagos' early drug screening methodology was explained for kinase.
In conclusion, in order to be certain that potential new drugs will be effective in humans, Galapagos uses human primary cells for new target discovery in combination with their adenoviral technology for knock-down and knock-in of genes. As such, no animal models are needed for target validation. In addition, early during the discovery phase, new hits coming up are tested on rodent ortholog targets. In that way, the amount of compounds entering the preclinical studies and risk of absence of efficacy in disease models is significantly reduced.

5.5.3.2.1.6. Use of Modelling and Simulation in the Prediction of inter-individual Variation in Human Pharmacokinetics: Potentials for Reduction of Animal Studies, by Zoe Barter, Simcyp, UK

When using animals to predict human pharmacokinetics there are several problems, for example, the predictability of animal data to estimate human bioavailability is not good. Few animals or animal-based systems mimic the human function or response. For example, regarding cytochrome P450 enzyme activities there are big species differences leading to different pattern in the metabolism of test compounds.
Traditional allometric approaches, scaling metabolic data from animal models versus human *in vitro / in vivo* extrapolation, have been under debate. In a paper of Zugo et al., 2001, a range of different models were compared, the conclusions were that "the most cost-effective and accurate approaches, such as physiologically based direct scaling and empirical *in vitro-in vivo* correlation, are based on *in vitro* data alone", "inclusion of animal *in vivo* preclinical data did not significantly improve prediction accuracy" and "prediction accuracy of allometric scaling (using animals) was at the lower end of all methods compared". A study was performed to characterise some of the scaling factors needed to extrapolate from recombinant metabolic enzymes to human beings and to compare a number of allometric scaling models with human *in vitro/in vivo* extrapolation. It was also found that the human *in vitro* data provided better predictions in most cases than the allometric approach.

Simcyps's task is to develop and update a user- friendly, comprehensive, mechanistic platform for integration of ADME models and databases. There is a consortium with 19 pharmaceutical companies as members. Each year a software product is released, which is used by the industry to optimise clinical trial design and predicting drug exposure and interactions in different populations. There is also interaction with regulatory bodies and licenses are offered to academia. Regular workshops are held with key players in drug development and consultancy services are provided.

One of the main focuses of Simcyp is to integrate demographic, physiological and genetic information with *in vitro*, drug absorption, metabolism and transport data. The goal is to simulate and predict *in vivo* drug absorption and drug-drug interactions in virtual patient populations. To simulate population pharmacokinetics, systems data, drug data and trial design are integrated, these together give a mechanistic estimate of *in vitro/in vivo* extrapolation, and physiologically based pharmacokinetics. These parameters can be translated into variables of population pharmacokinetics and can help to identify co-variates. Systems data includes information on demographics, physiology and different ethnic groups. This is time consuming, but once done for a certain population, such information can be used for many different drugs. Drug data, such as physico-chemical properties, are always different of course. It is also important to be able to simulate different dosages, however, trial design is also very important. These two factors have to be adapted each time.

When establishing a population, in order to generate virtual patient populations, information on demography, with factors such as addiction, AIDS, cardiovascular diseases diabetes etc. are very important.

Some of the models used are mechanistic models for absorption, physiologically based pharmacokinetics, distribution models and models for hepatic metabolism. These models alone are not enough. The goal is to incorporate them into individual variables. For absorption, there are differences in pH, permeability, diameter of the small intestine… For distribution, the volumes of the various different organs will be affected by age, sex, weight and height etc. With respect to metabolism, each individual has a specific pattern of each enzyme activities which obviously has an impact on the metabolism of a drug.

It is also important to look to population differences. Because of ethnic differences in the expression of certain enzymes. Thus depending on the population in which to market a drug, clinical trials should be adapted. It might also be possible to eradicate compounds faster.

To conclude there are many limitations for the use of animals and there seem to be many advantages in the application of in-vitro systems. The bottleneck is identifying and determining the data to use within the model. Often one must rely on using literature studies to populate the models. Standardisation of protocols would help to use all data together.

5.5.3.2.1.7. Assessment of Current Guidelines and Facilities to Reduce Animal Testing in the Development of Pharmaceuticals, by Jay Iyer, TI Pharma, NL

TI Pharma's mission is to stimulate public and private collaborations between academia and the international pharmaceutical industry. To reach this, TI Pharma has the following objectives: to create excellence in basic research, cross-disciplinary research, within the framework of Priority Medicines to focus on improving the efficiency of the entire drug development process, to facilitate direct contact with and to get input from regulators and to educate and train the next generation of biomedical researchers.
One project of TI studies is the assessment of current guidelines to reduce animal testing in the development of pharmaceuticals. The reason for this is that studies on animals play an important role in the drug discovery and developmental process. These studies are performed according to the guidelines and are issued by the different regulatory bodies to establish the efficacy and safety of pharmaceuticals, but their predictive value is under discussion. There is no recent comprehensive study evaluating the predictive value of animal tests and no recent assessment of alternatives to replace, to reduce and to refine the use of animals within already marketed drugs. The potential benefit of this study is to identify and evaluate the redundancies and inefficiencies of the dossier system, simplify regulations and add a reduction to the use of animals in testing and reduce the cost and time taken to bring a drug to the market.
This project is financed by several ministries and regulatory associations of the Netherlands. The input of regulators is also ensured. The goal is to bring all parties together which are involved in animal testing.
The goals of this project are to assess the current practices and the guidelines and to discuss the consequences and implications for the regulatory guidelines, to assess the predictive value of animal studies based on databases of selected companies for classical chemical and biological entities. Another goal is to evaluate the predictive value of animal tests in the various development phases with respect to toxicity and pharmacokinetics and to evaluate emerging technologies which can replace animal studies. A last goal is to design further scenarios and actions in pharmaceutical discovery and development strategies.
Individual files and data sets from the public domain of already marketed drugs were merged with predictive value parameters of these databases. The output will be presented as recommendations, workshops and publications. An inventory is made of all the data available at the Dutch Medicines Evaluation Board for all the products in the market. National and international regulations and guidelines on these drugs are being screened and a study is made of animal studies used in batch release testing. Products introduced into the market within the last 5 years with safety interventions will be studied. Studies in primates in biotech products will also be studied. Finally, there is a student program dedicated purely to technological assessment of new methods for activity and toxicity of drugs.

Discussion

Q: Gianni dal Negro (GlaxoSmithKline) asked Ms. Barter which would be the minimum information they would need to make a model like the ones she discussed.

A: Some of the information needed includes molecular weight, and some information about the clearance of the compound. Ideally also some measure of the metabolism and elimination would be needed.

Q: Vera Rogiers asked Ms. Iyer about how they were able to overcome the confidentiality of clinical data, because of the involvement of patients.

A: The same regulations as those of the Dutch Medicines Evaluation Board were applied. The researchers were blinded in the study, so that they could not see actual patient data. All data were coded by the Dutch evaluation forum.

Q: Jon Richmond (Home Office) said to Ms. Barter that he agreed with the desirability of standardised protocols. But apparently this is very difficult to get this working. He also asked whether the database includes lifestyle information, such as physical activity, smoking, dietary preferences, social factors, and others

A: This information is not included at the moment. This is also a matter of gathering the data and link them all together. There has been a collaboration with pharmaceutical companies, which had their own smoking population. With these models you can modify the parameters yourself, so if you do have the information, you can use it. There is a kind of wish list, with factors, such as health, that would be likely to be included in the models. But basically it costs time and effort to put it all together and gather enough data.

Q: (speaker unidentifiable) asked Ms. Iyer whether the 3Rs were also part of the training programme.

A: Since the project agreement was signed in January 2009, this is not part of the programme. It might be interesting to discuss with Simcyp about possible collaboration, since their education and training symposium series is interesting for this.

5.5.3.2.2. Session II: Reduction Potential by Handling Statistics and Experimental Data

5.5.3.2.2.1. Systematic Approach to Tune Number of Subjects to Scientific Hypothesis, by Peter Vijn, Schering-Plough, NL

In this presentation a live connection with the corporate data base of Schering-Plough was made. Using this technique, a presentation was given of the data base which this company's research scientists use every day in animal experiments. This database is global, it can be interactively worked with. This way, there is the possibility of doing matter analysis and statistical methods, which can bring up interesting questions for science.
The example was an osteoporosis study, where a lot of *in vivo* studies have been done. A lot of parameters are measured for each individual animal, amongst others bone parameters, such as. structure of trabecular bone and mineral density. These factors can be only measured *in vivo* under anaesthesia or during sleeping. The computer shows besides others, positive controls and all drugs that were tested in this assay including statistics such as average response, differences between the negative and positive control. These plots are then shown to the senior management to make decisions.
There also is a button 'extend'. In this case, it pulls out 8.120 animals. This will show how all the parameters are distributed in these animals., e.g., as histograms. The program shows the effect of the drug, but also the biological noise.
Since all the basic parameters are known, also the mean value of the negative and positive control is known and can be displayed for a particular parameter. It is also known how much noise there is on the variation of the data.
An interesting option is that simulations are possible. For example, factors, such as the number of treatment groups, can also easily be changed. Based on the compounds, which are based on real-life data, the number of animals per testing group can be decreased. When using, for example, only one animal, the noise is very high, but when using about 25 animals in a group, the means get stable. The interesting part here is of course that it is interactive and that factors can easily be changed.
Automatic simulations can also be run, this is then driven by a routine that automatically varies the parameters. This results in a power graph, from the real-life data. In this case, it advises the scientist that he should use, if he wants with 80% power to separate the negative from the positive control, 3 animals per treaty group.
These simulations can be made as realistic as possible. So, with this database, new questions can be asked and it can be used to train scientists. Of course, scientists can also get training to use the database in the company.

Q: Karl-Peter Pfeiffer asked Mr Vijn whether the information about the trial is also in the database? So that scientists know what kind of data they are dealing with.

A: All information is available in the database. Even of pharmacokinetic studies, all individual concentration data points and the drugs that were administered to the animal or excreted by the animal are available in the database. This way, no files have to be transported, which makes it very efficient. But because of this efficiency, the

management asks to handle more projects, which again improves the performance. But on the other hand the more projects are running the more animals have to be used.

5.5.3.2.2.2. Data Mining Tools for the Reduction of Animals for Adverse Effect Testing of Pharmaceuticals, by Christoph Helma, in silico toxicology, CH

The limitation of databases is that you can only look at information that is already available in the database. But if you for example want to see if a new compound is toxic and you only have part of the structure, then you can look at compounds in the database which contain this part of the structure as well. Then, with data mining or prediction algorithms, a toxicity prediction can be made, using the information from the training data.
Data mining algorithms are generic, therefore they are applicable to all types of small molecules, be it pharmaceuticals, pesticides or industrial chemicals. They are also applicable to all types of endpoints, such as adverse effects, efficacy etc. Data mining algorithms need experimental training data, but this data can come from very different sources (e.g. *in vitro, in vivo,* clinical trials, epidemiology…). Data mining models have a limited applicability domain, which is mainly determined by the training data that an algorithm has seen.
Data mining can be used when not all processes are understood and not all parameters are known, as is the case with toxicology. The Lazar system can be used for this, it is publicly available on the internet. A chemical structure can be drawn and then different endpoints can be selected. The main goal of this program is to automate the process that the risk assessors would do. The first step is to see whether there already are experimental results for this. If not, one can look for similar structures in the database. If those are available they will be listed as neighbours. The program also shows whether these compounds are active or inactive. Based on this, predictions are created. Also a confidence number is given for each prediction, so if this number is high, it is relatively certain that the prediction is correct. If the confidence is low, it might be better to do additional tests. A lot of links are also given, which provide additional information, links to other databases…
This program works quite well, but it depends on the confidence in the predictions. As an example, rodent carcinogenicity is a rather hard to predict end point, while salmonella mutagenicity is easy to predict. In both cases, the accuracy decreases with confidence, which is an indication that the applicability domain is very important. A cut off can then be made. When looking at the accuracies, they are quite close to the replicability of the individual assays. For the salmonella assay the replicability is about 80%, for rodent carcinogenicity it is not exactly known, because there are not much replicated experiments. But it is to be expected that the replicability of an *in vivo* assay is much lower than for an *in vitro* assay.
OpenTox is an FP7 project, with the goal to collaborate on a common framework for predicting chemical toxicity. The whole framework is an open source, so that scientists can check what the programs are doing with their data and it also has the advantage that other people can extend the program if they will. The website (www.opentox.org) will be

made open for the public in the future. The main target audiences are toxicologists, model developers and algorithm developers. There is an initial focus on endpoints that are relevant for the reduction of animal experiments. People working on the database side are also supported.
One goal is to make simple end user interfaces, which can be easily used by toxicologists. Thus, there is an easy access to well validated and scientifically sound prediction models, which would result in an interpretable presentation of prediction results and rationales.
Data mining techniques could help reducing animal experiments. *In silico* screening for toxicity is already established for early candidate screening for adverse effects. Successes vary between companies, but in some companies very good reduction of attrition rates have been gained. The US FDA is already using *in silico* techniques to support their regulatory decisions in the pharmaceutical area. Targeted testing strategies can be used to fill information gaps. In some areas the *in silico* models are already very reliable and it is not necessary to do additional tests. In other areas, where there is little information, data mining can help to select compounds with specific testing to cover these empty spaces in an optimal way with at least experiments as possible. Direct utilization of human data (e.g. from clinical studies, adverse effect repositories) is possible. A test was done with liver toxicity and the predicted capabilities of the *in silico* system were much better than those of the in vivo assay.

5.5.3.2.2.3. Round Table: Importance and Availability of Non-Public Data and Negative Results, Moderator: Bernward Garthoff, ecopa, DE

For this round table all speakers were asked to come to the front, including Mr. Gianni dal Negro, who is representing EPAA.

Q: Bernward Garthoff told that in both previous meetings as in this one, the issue of a journal of negative results was mentioned. The question then is how to get these data? Because there are several issues, such as intellectual property, confidentiality and the quality of those negative data, these have to be solved.

A: Mr. dal Negro made clear that the expression “negative data” should be more defined. Do we mean negative in the sense that the expectations in terms of pharmacological efficacy were disappointing or do we intend toxicity related data because in this case, they should be considered true data, which would enable to avoid useless use of additional animals.
To avoid intellectual property issues someone suggested that there could be a provision of some generic information or background information on class of compounds for example, but very similar molecules with just small differences can make big differences. So, this solution would be far from ideal.

A: Statistician (unidentified): A journal of negative results would not be necessary if results are just published, independent of the fact whether they are negative or positive data.

A: Man (unidentified): In the USA part of his company some people keep negative results to themselves, they are not even in the database. In the EU, all results are kept in a database, so they can always be checked back. He expects that his company will never share data in the public domain.

A: Helma: To make it really useful, there should be a combination of a journal and a database, so that the data are searchable and can be used for data mining. However, discussions on sharing data have been going on for a decade and everybody seems to agree that it would be good, but it never gets pushed through. Probably the scientists agree with it, but the idea is stopped at the management level.

Q: Garthoff: In some companies data sharing is done already and there are also some examples of archives of toxicology. Then the question is, why is there no world-wide data sharing yet, if it proves to be possible in some companies?

A: Iyer: When the consortium was started, there also was resistance, due to confidentiality, which is why the Dutch regulatory board was brought in. This gives some leverage to get some data, but this can only be done through partnerships. In the first year, only data from already marketed drugs could be obtained, so no negative results yet were available. Now, the consortium started to ask people about compounds that passed safety and toxicity, but failed in efficacy testing, because these compounds could have relevance for other indications. In this way it can be seen whether there is an interest in expanding the use of the compound libraries.

Q: Garthoff: One possibility is to code substance information, but how is this done? Which information does the index have, which part do the other people have? How can this be done and still share important information for the data base?

A: unidentified: Coding could be a way to encourage companies to share information, because a code is signed to the molecule, which is thus blinded. The problem is that at a certain point there must be the disclosure of the molecular structure, because otherwise it is not possible to make structural relationships. Then there again is the problem of intellectual property. So, there must be a super-party entity that guaranteed that the confidentiality related to these molecules is secured. But this party should then be totally neutral world-wide, so it cannot be one of the stakeholders or even the authorities.

A: Iyer: one of the things they try to do is to pool data sets, biological entities are pooled together, which gives the provider of each individual entity a certain leverage, because the results are based on data, predicted values of this pool of animal tests, or this pool of indications are also available.

Q: unidentified: If meta-analysis is done, sometimes a bias can be found in a selected paper. The same can be done here and if a bias is found, publications for negative results

are missing, then imputation strategies can be performed to get more balance. When this is introduced and data are pooled, the question is how data are pooled.

A: Vijn: Pooling is done across valid comparable studies, this is possible when the protocols are very valid and stable, to guarantee that the data are legitimate. Only then, meta-analysis or data mining should be done.

Q: Garthoff: A lot has been said already about the harmonisation of guidelines and directives, which would help having to do less control experiments, such as batch control. Maybe data which are normally confidential and intellectual property could in some cases be used for guidelines? Could this be managed by the ICH?

A: dal Negro: Global harmonisation of guidelines could be a potential way for sharing of negative results. The ICH does not pay to achieve a constant support to revise guidelines, therefore revising this implies the revision of data by regulatory agencies and companies.

Q: Garthoff: Are appropriate data banks available or are there more needed?

A: Helma: It would be best to focus on one system and not scatter data over too many places, which makes it harder to search for it.

Q: Garthoff: Is some kind of internal database needed?

A: Vijn: There is a discussion in Schering-Plough on how open the database should be, even within the company. In the USA all data are protected, it is considered to be very important. It is impossible to really dig into them.

Q: unidentified, comment to Helma: It was said that data mining tools are useful for people who do not know much about statistics. This is dangerous, however, since these persons will not recognise when results are wrong. Therefore, in this case data mining is good for general hypotheses, but not for conclusions.

A: Helma: It was meant that the tools can be used by toxicologists, without necessarily having all statistical background. The developers have statistical background of course. They also inform the users when results are not statistically reliable. So, data mining is used together with statisticians to make final conclusions.

Q: Maier: The production of homogenous results for simulation and speculative analysis might also include the same protocols, which might help to change an assay or test.

A: Vijn: One of the objectives of this exercise is to consult statisticians and other experts in the field of pharmacokinetics/pharmacodynamics, and people dealing with clinical studies, to get them earlier on in the research process. Their knowledge must then be put into a pool of which scientists can later make optimal use. Most of the work done in research is rather basic, such as logarithmic transformations, drug tests, repeat measures

analysis. When this information is collected and validated tools are added to the database, then one has an Oracle database, a SARS module and a user interface. These three together make the system very powerful.

Q: unidentified, to Iyer: How many years could one go back with the analysis?

A: Iyer: It is yet unknown how far can be gone back in general numbers. It is different in each study. Last week there was a discussion with the evaluation board to see which sets of data could be stopped. Concerning dissemination of results, different layers, such as advisory boards, arbitration committees and others, must be contacted before results can be disseminated or published in a journal, which is the first goal. The second goal is to organise or attend workshops and disseminate results via that way. EMEA supports this project and since some of them are also in the Dutch Medicines Board, they will also have access to the results.

Remark: Van Gompel (Johnson & Johnson): There is an initiative with 6 or 7 pharmaceutical companies that donated data of 50 compounds. In second trials, each company donated ten compounds, then this goes further on. The donation and openness, however, is only between these companies and data can only be accessed if the same amount is shared. This project is about genotoxicity data. With IP related compounds there was a pool of about 5000 pharmaceutical compounds. Here coding was applied. Information could be extracted or probabilities could be calculated, but the chemical structures could not be seen. The data were also only accessible to the companies sharing the data.

Q: Richmond (Home Office): In the pharmaceutical sector, product or target market protection is present in an early stage. There is more risk that two companies will sequentially be pursuing the same molecule to an equal application, rather than simultaneously. “Protection databases” could be made, where there is some protection compensation for the organisation that generates the data and the cost is then for example for the company that seeks the information. Another point is that data comes in different qualities. Regulated submissions are often done to GOP (General Operating Procedure) and high standards, whereas this is less certain with academic partners. The danger is that the EC is trying to achieve the sharing of information by amending Dir. 86/609 and is not distinguishing between the commercial sector and academia.

A: unidentified: GOP does not stand for quality, it only tells about traceability.

Q: unidentified: Many of the guidelines have not been formally validated, because they are too complex. The OECD has started to analyse datasets that have been obtained with these guidelines. It gives promising results, in particular in the area of reproductive toxicity, where it was found that the second generation studies, required by regulators, is not contributing significantly to regulatory decisions. Leaving this second generation studies out would lead to a reduction of more than 1,200 animals per chemical.

A: Iyer: It would be good to expand the project after 2010 outside the Netherlands, so it would be interesting to combine results and methods. How are results disseminated? Certain guidelines are recommendations and are presented to the regulatory bodies, but are they also presented to the public?

Q: Garthoff: Are there chances in pharmacology, in efficacy measurements to exchange or share data? Without going into certain substance classes?

A: unidentified: The field of pharmacology is more self regulated, because the human system is the most important system. So, in pharmacology there is less use of animals, but in toxicology it is more difficult due to regulations and restrictions.

A: Garthoff: There are some activities in the ICH field to go as early as possible to human clinical data.

A: unidentified: The problem in sharing data from pharmacology is that they are not so well standardised. For example, with cardiovascular research, the investigation in anaesthetised or conscious animals is very different.

A: unidentified: In early research (for example in oncology), models are developed on an almost daily basis, so there is a shift away from standardised protocols. This is also not shared with other companies, because there is information even in sharing the protocols, because they are highly specific for one particular pathway.

A: unidentified: In TI Pharma, there are PhD students, data are shared with them, but they have to sign confidentiality agreements, but nothing gets published, not even in their thesis.

Q: Rogiers: In European projects, data is often taken from a lot of different sources, put in one big group and then used. This is mainly in other areas than pharmaceuticals, but the question is whether this is useful at all. The question is how collaborations, more or less as the one done in the Netherlands, could be realised on a more international scale with different companies, but at their benefit and having it paid, and of course, keeping protection in mind.

A: dal Negro: The EPAA can help in this work. But there are also other initiatives going on in Europe. One example is the sharing of negative, toxic results. A consortium exists that brings a number of private companies and academic institutions together. It is hosted by Agnum. This joint effort has the goal to fill in the gap still existing in predicting unexpected severe liver injury in humans. In cases like this toxic results are shared for specific purposes.

Q: unidentified: When pooling data, one should always be aware of the problem of heterogenicity. When data sets are contradictive, one has to find out what the reasons for this contradiction could be.

A: Mittlböck (Medical University of Vienna): Heterogenicity may be a problem, but one can also learn from it, as it offers the possibility to investigate why the results of one study differ from those of another one.

5.5.3.2.3. Session III: Reduction via Study Design and Methods in Drug Research and Development

5.5.3.2.3.1. Non-invasive Small Rodent Imaging: Specific Examples and Contributions to Reduction in Drug Research and Development, by Nicolau Beckmann, Novartis, CH

Imaging is used for diagnostic purposes in hospitals. Techniques have been developed since a long time, but recently they have also been developed for imaging small animals, to characterise animal models of disease. Magnetic resonance imaging (MRI) is based on the use of magnetic fields and radio frequency. This technique is used e.g. for the study of models and for respiratory diseases.
The most common methods used in experimental research are terminal or very invasive. An example is bronchoalveolar lavage. Fluid is taken after the animal is killed and invasion of inflammatory cells into the lumen, lung function, airways resistance are investigated. Also, pulmonary chambers can be implanted to collect fluid from the lung.
Research was done to see whether MRI could be used in this area. The idea was to develop a very simple technique in which measurements could be carried out without any preparation of the animal, except anaesthesia. The measurements are performed in spontaneously breathing animals. A technique was developed to eliminate the movements from respiration or from the beating heart. Although the lung is a black area, some inflammatory signals can be detected. In practice, a series of images is made through the chest, so that, after computation, volumetric data are available. This normally takes about 20 minutes.
Allergen-induced changes can be retrieved, when animals are actively sensitised to an antigen for three weeks, the allergen is instilled directly to the trachea, eliciting inflammation of the lung. When saline is instilled, nothing happens, but when ovalbumin is instilled, a quite pronounced signal is received after 24 hours. This signal usually lasts a couple of days. This could be correlated to perivascular oedema determined histologically or to the protein in the lung fluid, which showed that this signal is indeed related to the inflammation, elicited by the allergen.
The technique was tested by using pharmacologically active compounds. In the following example 12 animals were used. With a traditional method 60 animals would be needed. 24 hours after the installation of ovalbumin a compound (budesonide) was administered. On the MRI a reduction of the signal was detected very fast. The results of the MRI were compared to those of conventional lung: At this early point there were no changes in the fluid parameters determined by the conventional methods. But with the MRI technique, there was a good correlation between the finding in the imaging and the perivascular oedema determined histologically. This example shows that the MRI signal reflects better

what is happening at the level of the tissue. This technique can replace the conventional, terminal technique to analyse compounds and it can refine experiments since it is able to detect effects which were not detected with the conventional fluid lavage technique.
Lipopolysaccharide induced mucus secretion is another model where MRI can be used. The signal is related to mucus formation elicited by the lipopolysaccharide. Also here the number of animals can be reduced.
Oxygen has paramagnetic properties and can maybe be used as a natural contrast agent. The acquisition parameters where changed a bit, the images are less clear, but some signal from the lung is acquired. An experiment was done in which the content of oxygen through the mask were changed, which gave a difference in signal intensity.
This information was used to investigate whether bronchoconstrictive effects could be detected. A series of images was made, then a drug causing bronchoconstriction was injected intravenously, thereafter, an increase in signal intensity was observed. Then a bronchodilating substance was injected, which was followed by a decrease of signal intensity. This was compared with the invasive functional analysis of the airway resistance, which gave the same response. The change in intensity is thus a result of the bronchoconstriction. This information can thus be used to obtain functional information. The advantage is that this technique is non-invasive and the experiments can be repeated.
Another model is elastase-induced emphysema, where the alveoli are destroyed by the injection of elastase. After two and four weeks there was a reduction of the signal and a peak after eight weeks. If compared to the histological method of the same animal, the MRI technique provided exactly the same results. This is again an example how a reduction in the number of animals can be reached.
One goal is to try to adapt the image acquisition to the physiological situation of the animal and not vice versa. To get very good images, the acquisition time can be heightened, but then the animals need to be intubated and artificially ventilated. But it is probably better to go for spontaneously breathing animals, then the acquisition time can be much shorter.
By using this technique, the number of animals can be reduced between 60 – 90%. It also works as refinement, since it is non-invasive, thus, the same animal can be monitored over time, which is important in the study of chronic models of diseases. This also facilitates the study of compounds and the response to treatment. The animals are spontaneously breathing and are neither tracheotomized nor intubated.

5.5.3.2.3.2. The Principle of Reduction of Animal Testing in non-clinical Safety, by Stefan Platz, Roche, CH

There are several constraints on reduction, out of scientific reasons and regulatory requirements. First of all, *in vitro* systems cannot provide a reliable picture on a complete biological organism, because of cell interactions and functional aspects. There are almost no alternative methods which allow responsible safety and risk assessment on endpoints, covering subchronic and chronic toxicity, reproduction toxicity and carcinogenicity. Wherever feasible, alternatives are used, for example for phototoxicity testing, local

irritation, pyrogen testing etc. Global regulation requests a fixed setting of animal studies for risk assessment and marketing authorization.
In regulatory toxicity studies different phases exist. In phase 0/I (enabling entry-into-human studies) the tests needed are a 2- / 4-week toxicity study in rodent and non-rodent animal species, including toxicokinetics and recovery, a single-dose toxicology study in rodents, for the Material Safety Data Sheets, local tolerance studies, genotoxicity studies and safety pharmacology. In phase I/II (early clinical development), there are 3 month studies, 6 month rodent and nine month non-rodent studies, *in vivo* genotoxicity study, embryo-foetal toxicity studies in rats, pilot studies in rabbits and fertility studies in rats. In phase III (entry into 'life-cycle management') natal or post-natal development studies are run and carcinogenicity studies in two rodent species or one rodent species and an alternative.
The estimated number of animals in repeat-dose toxicity studies for full development is about 600 – 800 mice, 1000 – 1700 rats, 100 – 140 rabbits and 150 – 250 non-rodents. This will probably not change in the coming years. But in some parts of the developmental process there are possibilities. For example, the predictability of mouse carcinogenicity studies is low, but it still has to be done due to regulations. The classical approach is to use two rodent species (rat and mouse) and have three doses and one control, so about 400 – 500 animals per study. There are some alternatives, such as transgenic mice, who are not oversensitive, but more subject to false negatives.
There are opportunities for reducing the number of animals in early predictive toxicity, so that the right compounds are brought forward, and acute toxicity studies, where the M3 guidance allows MTD (Maximum Tolerant Dose) studies, these data can then be used on behalf of acute toxicity studies. Then there are alternatives in carcinogenicity and dose-range finding studies can be optimised. From the regulatory environment ICH guidelines can provide opportunities. Also, new approaches of earlier Entry-Into-Human can help, such as exploratory clinical studies (e-IND (Exploratory Investigational New Drug), microdosing procedures).
In discovery research, *in silico* approaches have made progress, and can be used in target selection, for example there are assessments to calculate the 'likely fit to target' and for molecular modelling. Also high throughput screenings are now available. Target validation and efficacy models are being used in *in vitro* pharmacology for structure-activity-relationship (SAR) and selectivity /specificity screening. In the lead development and optimisation phase, progress is also made in pharmacological profiling, evaluation of potency, selectivity and pharmacockinetic properties and metabolism (DMPK) and in early (predictive) safety studies.
In discovery, Clinical Candidate Selection (CSS) tests for early (predictive) safety studies thus include *in silico* tools, Ames, micronucleus and embryonic stem cell test. The hERG test is very well established for cardiotoxicity; phototoxicity and phsopholipidosis can both be used *in silico* as well as *in vitro*. Other areas, which are still under evaluation, are *in vitro* toxicogenomics and the primary cell cultures for organ toxicity.
Some examples for the DMPK parameters and for CSS in discovery include P450 interactions, time-dependent inhibition, reactive metabolites, these tests might help predicting idiosyncratic toxicity and the industry is looking into biomodels for this. Other tests include microsomal hepatocyte stability, CaCo-2-cell monolayer, with which

progress is made for permeability, protein binding and transporters, such as P-glycoprotein.
The ICH M3 guideline for the Timing of Pre-clinical Studies in Relation to Clinical Trials has created some possibilities. Separate acute toxicity studies can be eliminated and repeated dose toxicity studies now have exposure and dose limits which allows to establish valid study designs. New exploratory clinical studies section will reduce the use of animals needed to support clinical studies and offer refinement of toxicology study design. Local tolerance toxicity is recommended against stand alone designs. Reproductive toxicity studies are deferred into later stages of development, which means that they are eliminated for failed compounds. Pediatric recommendations will eliminate routine use of a second juvenile toxicity study and minimizes the cases where juvenile toxicity studies are needed. Recommended photocarcinogenicity studies generally are not of value for pharmaceutical development. Abuse liability is generally recommended against use of primates with preference for rodent with limited doses. Combinations are recommended to be limited to one species, usually rodent
On the ICH meeting of June 2009 in Yokohama, an update on S9 (nonclinical evaluation for anticancer pharmaceuticals) was done. Maintained and expanded specific accomplishments leading to a reduction of animal use are: a 3 month studies will be sufficient for marketing authorization, the need for fertility and pre- and postnatal development studies is eliminated, only one embryo-foetal development study is required, safety pharmacology assessments could be conducted within the general toxicology studies, the need for the non-rodent study for initiation of clinical trials with cytotoxic drugs is eliminated, this is reduced to one rodent study, recovery requirement is reduced to a single species prior to phase I (need for recovery period based on scientific justification) and there is no need for photosafety testing anymore.
The pharmaceutical industry is seriously involved and interested for a variety of reasons. There are alternative models in place, on which active research is done. Good progress on some aspects of ICH guidance documents is made. A disappointment from the industry is that the FDA requires carcinogenicity studies for biologics, which needs to be discussed again.

5.5.3.2.3.3. Reducing the Number of Animals Used in Antibiotic Discovery, by J. Barry Wright, AstraZeneca, UK

Reduction of the number of animals in antibiotic discovery is possible through new testing paradigms, modelling and new technology. Bottlenecks to reduction and benefits of reducing the number of animals required are also important items. The use of animals in antibiotic discovery is still needed, for example in efficacy studies and in safety studies before going into human. But costs and timelines need to be trimmed to increase efficiencies.
An advantage in this field is that the target is the same in animals and humans. Theoretically, when a drug works *in vitro*, it also works *in vivo*. Of course, there are also some complications, such as with pharmacokinetics, to get the right concentration of the drug to the right place, protein binding, physical properties and toxicology.

A paradigm shift has occurred: In the past a lethal mouse model was used to screen compounds. The bacteria were administered to the mice and then the drug was given. If the mice survived, then the next step could be taken; if the mice did not survive, one would move on to the next compound. Nowadays, a lot of screening is done *in vitro* before a compound goes to *in vivo* examination. *In vitro* screenings include plasma protein binding, compound solubility assays, physical property determinations and toxicological screens. Pharmacokinetic screening is also done, which is often done in animals that are re-used.

Phamacokinetic/pharmakodynamic (PK/PD) modelling is also becoming more sophisticated and more effective in predicting the chances of success in humans. Data collection is done through these models, which also involves animals work, but less than in the past. New tests are also becoming available, such as a variety of *in vitro* tests, where different PK parameters can be simulated and results be acquired on the effect of a compound on a particular bacterium. It is also possible to determine some of these PK-PD parameters upfront in an *in vitro* system. The advantage is that animal model studies are very well designed and defined in terms of concentration and other such factors.

There are also additional tests which allow to look at PK. Probes can be put in place and the concentration of the antibiotic at the actual site of infection can be looked at. Other new techniques allow to take microsamples and get much more data from a single animal, which can reduce the number of animals in PK/PD experiments.

Of course there are a number of bottlenecks in reduction in antibiotic discovery. A first consideration might be that the ‘easy’ steps have already been taken. New methods are being developed, but they often have their own problems. One of those is the non-specific binding of compounds to plastic, sometimes to the extent that the *in vitro* method is useless. Microsampling can sometimes cause problems with detection limits. Separation of these samples can also be a problem. A question which one has to ask oneself is whether the two different kinds of studies lead to the same results and answers.

Bioluminescence seems to be a technique that is useful in antibiotic discovery, but, except for very peripheral type infections, the detection limits are not broad enough, to see changes in bacterial viability within a deep tissue.

Organisational aspects are another bottleneck. First of all, there are regulatory concerns, but also at the corporate level, it has to be proven that new methods are equally or more predictive than the old methods. It takes time and data to convince both company management and regulators. However, at the moment the industry is looking for new ways of reducing the number of animals required, because the mindset has changed and the 3Rs are now accepted as good science. This can help increase the quality of the data, because, for example, full PK can be done in one animal, as opposed to one individual PK (e.g. plasma concentration) value being acquired from one animal in the past. The speed of data collection has gone up, because *in vitro* screening goes faster and one can do this before going to lengthier animal studies. The cost of data decreases, because animal research is expensive, although it has to be said that not all *in vitro* techniques are inexpensive either. Finally, it also is good for the reputation and image of the companies.

5.5.3.2.3.4. Tissue-Specific, Non-Invasive Toxicity Biomarkers, by Patrick Müller, Novartis, CH

Non-invasive biomarkers are the only ones that can be translated into a clinical setting.
A biomarker is a characteristic that is objectively measured as an indicator of normal biological processes, pathogenic processes, or a pharmacological / toxicological response to a therapeutic intervention. The term biomarker can be considered as an umbrella term that encompasses a variety of markers quantified by a multitude of technologies. Biomarkers are a valuable decision-making tool for the pharmaceutical industry. Ideally, a biomarker is an early surrogate marker for a (regulatory) endpoint otherwise hard to be quantified at an early stage.
For example, nephrotoxicity has classically been diagnosed by increased serum creatinine levels. Although it is most frequently used to monitor kidney toxicity, this parameter has its limitations, it is rather a parameter of kidney function. It focuses on the glomerular filtration rate and thus it is of limited value for early detection of nephrotoxicity, in particular of drug induced tubular injury as most of the drugs affect the tubules. Several biomarkers are emerging, which enable to cover the entire nephron going down from the glomerulus to the proximal and distal tubule and collecting duct. Several other histological compartments of the nephron can be covered now.
For nephrotoxicity, NGAL **(**Neutrophil Gelatinase-Associated Lipocalin) can be used as marker for tubular toxicity. It is released into the urine upon toxic inserts in the tubules. When male mice were injected intraperitoneally single doses of 5 and 20 ml/kg b.w.of a solution of cisplatin, a strong tubular toxicant, cell creatinine levels were measured in the past, but no changes were measured, even three days after dosing. Thus, cell ceratinine is not a good measure for tubular toxicity. However, when NGAL is determined in the urine, a quite pronounced increase can be seen already after three hours post-dosing and it stays elevated up to three days. The increase is constant, not just a peak which can be missed easily. This is a good example of the characteristics of a biomarker, such as a low baseline level, a high dynamic change, relatively short after a toxic insult occurs, a short latency, and the increase should last for a reasonable amount of time. There should also be a clear dose response between 5 and 20 ml/kg.
Ideally, toxicity biomarkers can be translated into clinical trials for target organ monitoring. During early preclinical toxicity studies, requirements for serum and urine sampling need to be made and based on the histopathology findings, a preselection of biomarkers should be made before the first clinical entry. If there is a correlation between the level of these biomarkers and the histopathology data, a confirmed existing biomarker for this setting is found and can be used for sensitive specific target organ monitoring in clinical trials. If histopathology changes are absent, there is a high negative predictive value.
In the regulatory field, there are two terms which are important, being validation and quantification. Biomarker validation is the formal procedure to prove the suitability of an analytical assay for a given biomarker, it refers to accuracy, precision, limits of quantification, stability of analyte etc. It is a pure in vitro procedure to validate an analytical method. Biomarker qualification refers to the correlation with and predictability for the endpoint of interest, so this refers to physiological, toxicological,

pharmacological or clinical significance. According to the FDA, biomarker qualification draft guidance, biomarkers can be distinguished as exploratory, probable valid or known valid biomarkers, depending on the level of qualification they have undergone.
Recently EMEA and FDA together endorsed qualification of renal toxicity biomarkers to monitor tubular toxicities in a preclinical setting. This was based on data from the Predictive Safety Testing Consortium (PSTC), a joint industry / authority collaboration aiming at identification of new biomarkers for early detection of organ toxicity. The outcome of the qualification process was published in an EMEA/FDA guideline, which states that clinical biomarker monitoring for nephrotoxicity can be done on a case-by-case study in first-in-human trials when renal safety concerns are raised based on animal toxicology studies.
More biomarkers will be qualified in the future, therefore guidelines are issued to outline the process that needs to be followed to qualify biomarkers, which refer to non-clinical and clinical biomarkers equally and it involves a public consultation step for new biomarkers before they are endorsed or approved. For nephrotoxicity biomarkers, this qualification was provided by two-week GLP toxicity studies in rats with various compounds. The performance characteristics of these biomarkers were assessed by receiver operation characteristics (ROC) curves, and the diagnostic / prognostic performance in terms of the AUC (Area Under the Curve) of the ROC curve was assessed on the basis of histopathological scoring of tubular kidney lesions and corresponding biomarker levels by time course investigations.
To measure biomarkers, valid assays, which can be reasonably implemented, must be available. Toxicity biomarkers allow translation from animal to man, but not all are qualified or accepted by health authorities. ELISA assays for human and monkey are available for most of these biomarkers and at least also for one rodent species. Dogs are not well covered, but dogs are often required as non-rodent species in safety testing of pharmaceuticals.
Biomarkers enable to do non-invasive or minimally invasive investigations. If biomarkers correlate very well with histopathology, they can be used for time course investigations, and thus save animals which would be required for post mortem read-outs. This applies equally to pharmacology and toxicology studies.
Biomarkers can also be used for clinical monitoring of certain toxicities which are otherwise not monitorable by classical clinical pathology parameters. Normally in this case no clinical trials can be done except with huge safety margins. Thus if such toxicities can be made amenable by biomarkers, less preclinical mechanistic studies for showing human irrelevance in such cases are needed.
The implementation and qualification of biomarkers is resource intensive and relies on animal experiments, the potential to reduce animal numbers works on the long term. At the moment, only ten qualified renal toxicity biomarkers are available on a regulatory basis. For toxicity biomarkers, there are hardly any assays available for the dog, which also is a bottleneck.

Discussion

Q: Why is there a lack of specifical biomarkers for dogs? If biomarker identification is possible for primates, then why not for dogs?

A: Müller: In academia where most of these biomarkers have been discovered, dogs are commonly not used as models, therefore most of the discovery work is based on rodents. There are also no antibodies available for nephrotoxicity in dogs.

Q: Maier: What was the driving force to change the kind of studies, quality of the data or a combination of management issues?

A: Beckmann: The quality of data is driving this development. Better data can be acquired by using 3R methods, which is a good point, next to reducing animal use. Time courses can be accessed. Furthermore, other compounds can be tested, which is not the case with terminal tests.

Q: Maier: The number of animals used in pharmaceutical companies and drug development is more or less stable. Do the new methods overall reduce the number of animals or is there a compensation for more efficient and more testing of new candidates?

A: Wright: The overall number of animals being tested for compounds is decreasing, but the number of compounds being required to move through is increasing, thus, it indeed compensates.

Q: unidentified: Animals are often re-used nowadays, but does this not raise concerns about the quality of the animals, after repeated treatments and tests?

A: Wright: The idea of re-using animals is under discussion, but the number of animals would increase dramatically if they were only to be used once. For example, for local PK, the dose of the compound administered to the animals is very small, therefore, after a certain time there is no negative impact, the animals do not suffer. There also might be some impact of the animal being handled and dosed repeatedly. But when weighing the benefits versus the negative points, the benefit of re-using animals certainly outweighs the negatives.

A: Beckmann: A differentiation between terminal and non terminal experiments has to be made. Most experiments done in safety, need histopathology, this requires terminal experiments. In non-terminal experiments, there are two categories. In classical pharmacokinetics, non-rodent species are re-used. If PK profile must be determined in a dog, it gets a vacation, recovery period and then gets in another dose. These are pharmacological doses, which do not induce any toxicity. In pharmacology models, the second category, good quality data are needed and thus the animals cannot be re-used, for example with a rheumatoid arthritis rat model there is significant swelling of the joints.

Q: Maier: Carefulness is needed when relying on the correlation between pathohistology and some of the non-invasive markers release because it can change from species to species quite remarkably, thus if different species are used. The pathohistology findings may not well correlate with some of the markers used in human.

A: Müller: The qualification process for such biomarkers is only done in one species and sometimes they must be re-qualified in other species, the qualification of a biomarker cannot be translated from one species to another.

Q: Garthoff: Is there a reluctance at the corporate level, when there is expectation of hindrances at the regulatory level, in particular when the corporate level anticipates that it will not work anyway? How can this vicious circle be broken?
If add-on alternatives in reduction do not make sense, then to what extent do alternatives themselves make sense?

A: Wright: If new techniques like this do have good scientific background, there usually is not that much resistance from the corporate level. It just has to be proven that the new techniques are equally valid.
With add-ons, people first have to be convinced about the new method, for which it has to be compared with the old method, during a period of using both methods in parallel.

Q: Rogiers: Are the non-invasive techniques of lung inhalation also used in other industrial fields? Inhalation technology is not only interesting for the investigations of drugs, but also for cosmetics, cleaning products. Are these techniques known in other fields? How can they be made known to the public?

A: Beckmann: The most important way to show new techniques to the community is by publishing data. Interest is coming from different sides now, both from academia and from pharmaceutical companies.

5.5.3.2.4. Session IV: Transgenic Animals in Pharmaceutical Research and Drug Development

5.5.3.2.4.1. Reduction in Generating and Maintaining Transgenic Animals. by Thomas Rülicke, University of Veterinary Medicine, Vienna, AT

In basic research often more than half of the animals used are transgenic, which means that breeding and maintaining of transgenic strains contribute, to a large extent, to the total number of experimental animals used for research. This is a target for reduction. In general there are two kinds of transgenesis. The first is using non-homologous DNA. This includes the most used pronuclear injection techniques and viral vectors of transgenesis with sperm cells use. A relatively recently developed technique is involved, namely the use of mobile genetic elements, the transposons. The second kind is

transgenesis by homologous DNA recombination, gene targeting in embryonic stem cells (ESC) to generate knock-out mice.
Homologous recombination starts with the injection of an ESC into a blastocyst. The genetic background of ES cells is mostly the genotype of the host blastocyst C57BL/6.
When generating a knock-out line, the chimera is usually bred with a C57BL/6 mouse. The chimera can produce three different kinds of sperm cells, sperm cells with BL/6 of host blastocyst origin, and sperm cells with ESC origin with or without imitation. The next step is to identify the animals with the targeted mutation and start to breed the heterozygous F1 generation to produce homozygous knock-outs. In the F2 generation one has about 25% homozygous offspring.
The problem is that the animals of the F2 generation and the following generations have a mixed genetic background. Genetic modifiers of the genetic background can be a source of artefacts, the genetic background can be corrected by reading a congenic strain, which means backcrossing for ten generations over three years to get a clean recipient background for the mutation. Nowadays marker assisted selection can be used to reduce the time and to reduce the number of backcross generations. But this does not reduce the number of animals.
This backcrossing can be avoided by inducing the mutation directly into the appropriate background. This is done by three consortia, KOMP (Knock-out Mouse Project) in the USA, EUCOMM (European Conditional Mouse Mutagenesis) in the EU and NorCOMM (North American Conditional Mouse Mutagenesis project) in Canada, they all use ESC with C57BL/6 origin. By now there are thousands of knock-outs in form of embryonic stem cells.
Non-homologous DNA end joining is more complicated. Pronuclear injection is most often used, with an efficiency of about 20% founders. A new method, transposon transgenesis, has increased the efficiency, with success rates of about 70% founders, which may lead to a reduction of animals. The founders are now going to be genetically re-characterised, the first results are very promising and this technique will most probably be in routine soon. Nevertheless, there are still some problems with all techniques based on this non-homologous DNA end joining.
The most striking characteristic of this approach is the random integration of the DNA, due to which there is a problem with position and effect variation. This means that conclusions regarding a transgenic phenotype can be drawn only from at least two to three independent lines, because each line produced by this technique is unique.
With random integration, about one integration event in a founder is expected and, according to Mendelian law, about 50% transgenic offspring in the F1 generation is expected, but this does not happen in reality. The great number of founders are genetic mosaic founders, which either have a low transmission or no transmission of the transgene. Sometimes, seven to eight litters have to be bred before the first transgenic mouse strain is found, which means that a lot of animals are bred, which are not needed. Another problem, due to this random integration, is the occurrence of an inappropriate expression of the transgene.
With non-homologous DNA end joining a lot of redundant reading of several independent lines with unpredictable transmission and expression of the transgene is done. This technique however, is the most used one to generate transgenic animals

nowadays. What is needed, is the possibility to direct the integration of the transgene in a neutral genomic site. This could maybe be done by enzyme mediated cassette exchange or with an acceptor strain which is prepared nuclear genomic docking site. The transgene is flanked by hazard-specific sequence motives which react by the help of site-specific recombinators of site-specific integrators, resulting in a stable integrated transgenic cassette. This is for the future, but some promising projects are running at the moment.
An interesting new possibility is the archiving of mutants. If animals are not used for a while, it is not necessary anymore to continue breeding. Animals can be archived by cryo-preservation of early embryos or germ cells. If this is not possible within the own institute, animal models can be sent to European Mouse Mutant Archive (EMMA), there are several centres around Europe, where this service is provided. This will also facilitate the incorporation of information about models into a databases, which also helps to avoid repetition of already existing animal models. This will also contribute to reduction of transgenic animals.

5.5.3.2.4.2. Transgenic Animals in Pharmaceutical Industry: Contribution to Reduction, by Rainer Nobiling, Heidelberg, DE

Transgenic animals do not only include animals in which transportation of genes has been carried out, but should include all animals in which genetic manipulation has been done, so that also all (conditional) knock-out mice are called transgenic animals. In the past, genetic modification was done with forced mutation by irradiation. Until 1990, in Germany this was not seen as an animal experiment, so no numbers are available before this time. Animals often suffered severely during these procedures and there were a lot of lethal mutations. One of the reasons that targeted manipulation of genes developed is that forced mutagenesis quite often was painful for the animals.
In the German statistics of experimental animals, there is a minimum around 1997, and after that the number increases again. The reason for this is not that at suddenly more experiments were done. The German animal welfare act was changed in 1998 and the statistical procedure was modified in 2000. In the new statistical procedure all animals have to be included, also animals killed for producing cell cultures. Before the animal welfare act did not include the animals killed for scientific purposes but also not those used for immunisation. Thus, all these changes make it really difficult to compare statistics. Alternatives, like *in vitro* work, are supported since 15 years, this number perhaps also increases, but this trend can only be seen since 2000. This all makes it difficult to compare the statistics to those of other European countries.
When looking at the German statistics, it becomes clear that the number of animal experiments increases in the last couple of years, but to a much lesser extent than the number of experimental animals. In the years ‘06 – ’07 the number of animal experiments decreased, but the number of experimental animals increased.
The British statistics show the same trend, the number of transgenic animals increases, and the number of total procedures also, but not as much. In Austria, there is also a small increase in animal numbers, but here transgenic animals cannot be separated from the rest. Switzerland also has the same main trends. There is a steep reduction until 2000 and

a slight increase since then, although Switzerland has a completely different contribution of industry than the other countries. In total Europe the number of animals used for research purposes is around 25 million at 500 million residents.
Thus, there is an increasing trend for the number of animals; for the number of experiments there is a reversed trend. This figure may be explained by the fact that the use of transgenic animals still increases. The hope is that this may refine research by gene targeting.
Animal experiments are performed in disease research, but mainly in basic research, which normally has nothing to do with diseases. Medical research, really applied research, going into industry, has rather small numbers of animal experiments, but this is increasing in the last years. Without relation to diseases, this number is much smaller than the number with animal experiments that are related to diseases. This trend did not reach the pharmaceutical industry. For the future, hopefully translational research can be transferred from academia to industry, so that more drug design can be done because there is better insight into signal systems, molecular mechanisms of signal systems and receptor systems. Improved transgenic models (cell specific / organ specific expression) will result in even better defined, smaller experimental groups of animals. Incorporation of reporter genes (dyes and more) will improve the quality of results and contribute to reduction of animal experimentations.
This development is not yet fully introduced as applied research into industry, although some innovative projects already have been launched. At present, it must be admitted that the development of these models takes more animals than the benefit that can be extracted from these animals.

5.5.3.2.4.3. SPF-Housing in Relation to Reduction, by Werner Nicklas, DKFZ, DE

SPF (Specific Pathogen Free) housing is used together with microbiological quality. A question could be: why should microbiologically standardised animals be used for experiments? Microbiological quality means dealing with infectious agents, which may cause clinical signs in laboratory animals. Other reasons to be careful are that some agents can be zoonotic, can reduce the lifetime of laboratory animals and can have an impact on physiological parameters. Many infectious agents may increase inter-individual variations, which is important because in this case more animals are necessary to get significant results or test results are statistically not significant although the animal numbers are based on the biostatistical calculation.
As an example, the variation coefficient of kidney rates of 58 groups of rats, all having the same genetic background, did not differ markedly between animals housed under highly standardised laboratory conditions and animals housed in a game reserve with maximal environmental diversity. However, in some groups the variation coefficient was much higher. All these animals were artificially infected by *Micoplasma pulmonis*. This is biological noise, which can be increased by natural infections or by animals that are not sufficiently microbiologically standardised. For these reasons, it is fair to state that infectious agents may have an impact on the results of animal experiments.

Overt infectious diseases may result in substantial research complications, for which reason clinically ill animals should not be used for scientific experiments.
However, the effect of clinically silent infections may even be worse because they often remain undetected and modified results may be obtained and published. This happens very frequently. This might be mostly a problem in universities and research industries, but less in the pharmaceutical industry. Unfortunately, many researchers are not aware of this problem. because in most infections, the rodents have no clinical symptoms and modifications of research results due to natural infections very often occur in the absence of clinical disease. Thus, the absence of clinical manifestations has only limited diagnostic value.
Complications in the absence of clinical signs include changed behaviour, suppressed body weight, reduced food and water intake, which is a problem when drugs are given with the food or drinking water, and reduced life expectancy, tumour rates can also be influenced. Samples and tissue specimens can also be contaminated, which can be a problem when transplantations from one animal to another are done. Sperm cells can also be contaminated, and there is some proof, from swine and mice, that agents have been transmitted by sperm which has been taken from infected animals, although it is not a big problem.
Even if animals are clinically healthy, it can happen that clinical disease or even death can be induced by experimental stress. In addition, environmental factors can lead to clinical disease in previously healthy animals, for example when rats are housed in metabolic cages. They are housed singly and rats are social animals. An increase in temperature may activate infections. Interactions with various microorganisms can happen, as a synergistic effect.
The parvovirus is one of the most frequently found viruses in mice. It, however, is not pathogenic. The autonomous parvovirus can infect and destroy cells only during the S-phase of mitosis, so biological effects can occur only in rapidly growing tissues, e.g. during embryonic development and pregnancy. It may result in embryonic death, malformations, reduced litter size, which severely affects teratogenicity studies. This group of viruses can also affect carcinogenicity studies. Different allotropic strains exist of this virus. The first found strain, the prototype strain, replicates only in fibroblasts and only affects connective tissues, which is a problem in experiments for wound or bone healing. Another strain, an immuno suppressive variant, MMVi, replicates only in lymphocytes and has various effects on the immune system.
SPF is a concept, but does not say anything about the animal quality. SPF animals might differ in quality. There is no need for keeping all animals under SPF conditions, but it is necessary for those that have defined or at least known health states. To describe the health state of animals must include details about individually listed microorganisms and their absence has to be demonstrated by regular monitoring of a sufficient sample size at appropriate ages of the animals and by appropriate methods.
The use of animals that are clean of infectious agents, that may influence health or physiological and other parameters, is necessary to get better and more reliable results of animal experiments. The use of microbiologically standardised animals is also an important basis for the reduction of the number of animals. These aims can be achieved only when animals are housed under conditions that prevent the introduction of unwanted

agents. An awareness of risk factors is important. Risk factors are animals with insufficiently defined microbiological status and biological materials taken from such animals.

Discussion

Q: Unidentified: What are the effects of housing? Is standard housing better than a free environment? Is environment enrichment possible to full extent?

A: Niklas: Housing has some impact on the physiology of animals, but the effects of housing are most likely less important than the effects of infectious agents. The effects of environment enrichment are not yet fully clear, it is still too early to make valid conclusions.
Environmental enrichment is also not standardised, even bedding can be understood as environmental enrichment. Breeding animals are always cleaned, but it probably happens in every institution that once in a while some infected animals breed.

Q: Garthoff: In the regulatory part there apparently are no transgenic model sets that are sufficiently used in, for example, toxicology. How much of transgenic material is re-used in pharmaceutical industry, is it done in basic research and then supplied to industry?

A: Nobiling: Transgenic animal models are used in pharmaceutical industry, but not in the complete strength onto the regulatory. Very well defined animal models are still too young. The efficiency of marking specific receptors, the specific metabolic path, is not yet sufficient for these models to be used for reduction. But progress is being made.

A: Richmond: The trend of using transgenic animals is indeed increasing in academia and in the commercial sector.
When bringing biological products, such as cells that have been passaged, it is very important to screen the incoming materials, otherwise it is useless to define the health status of animals.

A: Nobiling: The problem with the statistics is that transgenic animals are not counted in the same way in the different countries in the EU. So, to really know in how far the use of transgenic animals is on the rise, statistics should be harmonised. The harmonisation of statistics and the counting in Europe, would be an interesting initiative. One possibility is to bring this into the revision of Directive 86/609.

Q: Garthoff: What are the application areas, the pure pharmaceutical applications?

Q: Rogiers: And what is the role of academia? Is academia doing research only for the research or is there also a real use in creating knock-outs?

A: Rülicke: The genome of the mouse is decoded, but it is unknown what is written on a specific gene. The only chance to get information about the function and the inter-relation of the genes is in mutated genes. This is one of the reasons why most probably the number of transgenic animals will increase during the next years.

Q: Rogiers: Is this helpful for the development or not? Maybe it is done not professionally enough and in too small projects, so that there is very little relevance?

A: Nobiling: It would indeed be probably better if a huge consortium of professionals is doing this, than when each single level and department in Europe tries to knock-out a specific gene. This way, the numbers of animals might still increase, but then it is certain that the mice are of high quality with a specific gene knock-out. This is a project that is running at the moment.

A: Rülicke: It is indeed being done, but it is not very well known in the public, what's going on in this field. So, it must be made more popular.

A: Garthoff: It probably will become more popular in the scientific community, when there will be publications where it is written that people got their knock-out mice from a certain company.

Q: Zucco: Patenting can be an obstacle when updating some transgenic mice. Does this European consortium has an agreement about patenting?

A: No, it is open for both academia and industry to use, an inexpensive fee has to be paid though.

Q: Garthoff: The function of genes can also be knocked out by gene silencing RNA. Is this a method that is used at the moment?

A: Rülicke: This is not done yet.

Q: Gstraunthaler: Is it possible that this in the future might be more efficient?

A: Rülicke: It is another method and it sill needs improvement, although it is a promising one. But it will not completely be an alternative to the knock-down. There is also another phenotype as in the knock-out.

Q: Kolar: What is happening in genetic engineering in other species? Are there options for reduction there?

A: Rülicke: Two publications in Nature stated that people were able to establish embryonic stem cells for rats. Maybe this will open the possibility to produce knock-outs, a genetic model with target modifications comparable to that what was done in mice. But there are no other species where it is possible to establish embryonic stem cells. So RNA

interference of knock-down is the only alternative at the moment or inducing targeted mutations in somatic cells with reproductive cloning afterwards.

5.5.3.2.5. Plenary discussion

Rogiers: Reduction is an important field for the pharmaceutical industry where a lot can be done. It seems that there is an internal willingness in the pharmaceutical industry, because a lot of cooperation exists to address the bottlenecks and possibilities. From this, and the previous meetings, a lot of ideas were gathered, which will be put in a report, addressing the bottlenecks, possible solutions and a number of proposals and prioritisations of what can be done.
It is the responsibility of the START-UP project to bring forward correct messages. These messages should also be brought to the general public to help wrong ideas out of the world, and to scientists, who often do not seem to be aware enough of new developments.
The chairs of the different sessions were then asked to tell what they found interesting in their sessions.

Pfaller: It is interesting that there is a substantial possibility to reduce animal experiments in the development of biologicals. The models with humanised immune systems are very interesting. It also is quite new. It is also interesting to learn that although using these transgenic modified animals the number can be kept fairly low.

Hendriksen: An interesting point is to make the difference between inter-, super- and extra-experimental.
Apart from that it is clear that not only new methods are important, but also a change of philosophy, a paradigm shift in the way testing is done. In pharmacological testing, for example, this paradigm shift is already happening, by introducing a lot of high throughput methods in order to screen products.

Garthoff: New data systems and systematic approaches are being developed, but what still needs to be worked on is the procedure of how to retrieve the data, how to get confidential data, and how to deal with intellectual property. This could for example be dealt with by a data trade organisation on a global scale.

Maier: In the process of drug discovery there is no real regulation, but the progress is science driven.

Van De Water: In drug discovery a lot of progress has already been made, a lot of new tools have been developed to refine and cause reduction of animals. But training and education of people who work in drug discovery really helps to cause reduction.

Rülicke: As long as there are no good statistics on transgenic animals, it is not possible to give answers on how to reduce these numbers and to which extent they can be reduced. There is hope for refinement and to do better drug design by using transgenic models, but the numbers show that the relation between basic and medical research is not good enough yet.

Richmond: Coming back to the paradox raised by Mr. Kolar, that if progress is made with reduction and replacement, then why do the animal numbers still go up?
May Petter, a contemporary of Russell and Burch, argued that better experimental tools were needed to overcome the limitations in expanding the scope of the animals that were available. New developments do not simply replace existing animal models, but also open new areas in scientific inquiry. Therefore, progress will not necessarily end animal research, it may actually create the demand for additional and different types of animal research. To get a clear view on reduction and replacement, one has to look at case studies.

Rogiers: A point that is often overlooked, is the need for better science and understanding and what this means in terms of animal use and alternative methods.

Pfaller: By having biological and biotech drugs, the spectrum is shifting to some extent and people have to be made aware of this.

Rogiers: When mentioning pharmaceuticals, most people still have the “old” pharmaceuticals in mind, namely chemical compounds. They often not think of the compounds in the pipeline, namely the biologicals or biopharmaceuticals.

Unidentified: One problem is that it is not always possible to know the benefit of certain developments, but it also is not possible to know when certain developments will not be beneficial. So at the end of the chain many experiments could be identified as inefficient or not very useful for medicine. This also needs consideration.

Q: Unidentified: Are human tissues useful in industry and academia? How can the collection of human samples be improved, to reduce and maybe even replace some of the animal studies?

A: Unidentified: In comparing the number of marketed drugs and with the number of animals used, this is a reverse relationship. Therefore the use of human samples is a must and will be supported by the industry.

A: Unidentified: Coming back to the reduction of the number of animals, it must be kept in mind that, compared to 20 years ago, society has become more complex and a lot of new questions and research areas are rising.

A: Zucco: With computer modelling very complex equations can be done in seconds, in order to simulate different functions, binding conditions from the molecular to the

physiological level. For biologicals, this would be an important and powerful tool. The topic of bio-simulation should be taken up as possible alternative method.

A: Höllriegl: When working with human tissues, for example blood, it is often the case that while answering one question, a couple of others come up, which can not be answered with this specific tissue. Therefore, these new elements are often taken to animal testing, so that the numbers still rise. It would be good if it was possible to divide the statistics in categories. Because now all the public and animal welfare organisations see a rise in total numbers and they have to trust the pharmaceutical industry that an increased amount of questions has been answered and that in some areas reduction is effectively done. The problem is that this trust is not there, so with these categories in the statistics, it would become more clear what is exactly happening.

A: Unidentified: Projects are now in the pipeline where human tissue material from surgery can be used for testing biologicals, among others. But this is a very recent development, so there are no results yet.

A: Unidentified: The availability of human tissues is a problem, not only purely because of availability, but also because of ethical problems. For example, must a surgeon ask the allowance of the patient to take a part of the heart muscles for research?

A: Garthoff: The quality and standardisation of human material also is a problem that needs discussion.

Pfaller: In cardiovascular research, material out of heart transplantations was gained, but the quality was very different between samples, all stages of cardiac decomposition were present and so on. So, analysis was done on how the muscles are reacting and this was then transferred into transgenic animals.

A: Rogiers: Also human cells are important. It is often believed that stem cells can be used for the production of all kinds of human cells and that these can be combined into organs and that therefore no more animals are needed. But this is still quite far from the stage that we actually are.

A: Dal Negro: Human tissues are already extensively used in the early phases of development of mechanistic screening. But there still is another problem with human tissues. Humans are one of the species with the highest genetic polymorphism and single cells coming from one individual are not representative for the population.

Workshop Replacement

2 - 3 October 2009, Airport Hotel, Budapest, HU

5.6. Workshop Replacement

5.6.1. Executive Summary

Replacement was defined in this workshop as avoidance of any animal experimentation in the respective part of development of a new pharmaceutical product, which is performed to demonstrate proof of efficacy, to enable assessment of safety or to ensure quality control of the production. Pre-selection of drug candidates performed by *in vitro* methods or mechanistic investigation were regarded as reduction of animal experimentation and, therefore, not discussed in detail.
In the Replacement workshop, the individual parts of drug development e.g. pharmacology and various safety aspects were discussed by experts of the pharmaceutical industry and regulatory bodies. In the same way, biologics including vaccines and sera, monoclonal antibodies and certain aspects of quality control of production (batch control) were discussed too.
A number of key points are summarised here.

- Efficacy / pharmacology
• Before a new entity can be considered for clinical trials, a successful outcome in an experimental model remains crucial. It is, however, important that this model is relevant and predictive for the proposed indication. It should deliver additional information with respect to pharmacokinetics and safety. It became clear, however, that more relevant models are needed to improve the predictability for efficacy. If no specific animal model exists, efficacy studies in animals are highly questionable.
• One of the upcoming strategies is based on PBPK modelling in animals. This strategy needs further optimisation: smaller blood volumes, sparse sampling and a population approach were among proposed strategies. Regulatory trust often is limited and is not very helpful for further improvement.

- Toxicology / Safety pharmacology of small molecule pharmaceuticals
• Safety of small molecule pharmaceuticals is primarily driven by risk assessment based on (sub) chronic repeated dose toxicity studies. This represents the key bottleneck due to its complexity. General agreement existed that even the most sophisticated combination of cell cultures or *in vitro* systems, available today, cannot provide a reliable surrogate for the complexity of a complete biological organism.
• Agreement, however, also existed on the necessity of a new safety testing strategy based upon knowledge gained, new technological developments, and existing and emerging alternatives, based upon a number of crucial points identified by the experts present:

- excluding high dose animal exposure

- more extended use of high-throughput screenings based on well-identified pathways of toxicity in human cells, invertebrate species, computational methods, *in silico* expert systems
- inclusion of *in vitro* tests focusing on mechanistic approaches, target activity, metabolism, specific organ toxicity…
- development and inclusion of batteries of sensitive, specific safety biomarkers for most key target organs with *in vitro* and *in vivo* relevance
- intensified use of imaging technology as well for *in vivo* experiments as for *in vitro* systems
- increased research efforts for better knowledge of human and animal genomes, stem cell technology and systems biology
- more efforts in data sharing between different companies, well-organised and thus without loosing competitiveness
- better integration of pharmacokinetics, safety pharmacology and toxicity studies into one package
- development of argumentation for a gradual shift from animal to clinical studies (micro dosing)
- a more focused involvement of ICH and increased efforts for more global harmonisation
- it was also thought that the possibilities for replacement was higher in pharmacodynamics than in toxicology because of the non-existence of specific performance guidelines in that field
- the need for a two generation study within reproductive toxicity was discussed and it was thought that in general this study could be omitted

- Biopharmaceuticals, vaccines and quality control

• With respect to biopharmaceuticals, classical toxicology is not of real help and alternative strategies need to be used including knock-out animals, transgenic, humanised mice.

• When no scientific relevant animal model exists, usually because of high target specificity, animal studies become highly questionable. The relevance of studying biologicals in non-human primates was heavily discussed including the necessity to justify their use for each case based on cross-reactivity and functional data.

• For efficacy assessment, vaccines and sera development, toxicity testing, some replacement methods exist and are in use, but we are still far away from full implementation in all companies due to technical reasons, hurdles regarding regulatory procedures, high costs and lack of incentives.

Some countries require animal-derived data although relevance for efficacy and safety assessment is highly criticised.

• Discussion at a global level is highly requested bringing together all different players such as ICH, WHO, ICCVAM, ECVAM. European funding for projects on biologicals was also seen as a high priority.

• A particular problem exists for veterinary vaccines for which in particular incentives are needed in order to obtain some 3R improvements.

5.6.2. Recommendations

1. In general, the efficacy of potential drug candidates needs to be studied in animal experiments before entering in clinical trials. This is, however, only true when a relevant model exists. If not, animal experiments are highly questionable and should even be deleted. Authorities should accept such justification.

2. New safety testing strategies are highly needed, based on knowledge gained, new technological developments and existing and emerging alternatives. Science should be the driving force.

3. Attention should be given to a better integration of safety pharmacology, pharmacokinetic and toxicity studies, whenever possible.

4. The development of a battery of sensitive and specific safety biomarkers with clinical relevance is a high priority and should be included in new safety testing strategies.

5. Improvements in safety testing seem feasible through different means e.g. intensified use of non-invasive imaging; application of more specific screenings of well-identified pathways; use of more human-based cells / tissues; better integration at all stages of *in vitro, in silico, ex vivo* with *in vivo* results; data sharing; availability of negative results; excluding of high dosage when animals are involved, etc.

6. Consideration of micro dose clinical trials under well-controlled conditions before finalising safety studies in animals, at least for early PK studies.

7. From a regulatory point of view, make the use of alternatives easier for defining the potency of vaccines and sera; also in quality and batch control, provide incentives and give more attention to veterinary products.

8. Global harmonisation of requirements by regulatory bodies is essential to realise progress. In particular, adequate testing of biopharmaceuticals should be clarified. Europe could take the lead in these discussions.

5.6.3. Report of the Workshop Replacement

START-UP Workshop Replacement
2 - 3 October 2009, Airport Hotel, Budapest, HU

Program:

2 October

13:00 – 13:35	**Welcome and Introduction**
13:00 – 13:10	Welcome. Lajos Balogh, hucopa, HU
13:10 – 13:25	Introduction to *ecopa* and START-UP. Vera Rogiers, *ecopa*, BE
13:25 – 13:35	Introduction to the workshop. Gisbert Sponer, set, DE
13:35 – 15:20	**Session I: Replacement in the development of new chemical entities**
13:35 – 14:00	Keynote I Friedlieb Pfannkuch, Roche, CH
14:00 – 14:40	Preclinical Pharmacology. Klaus-Dieter Bremm, Bayer Healthcare, DE and Jan-Willem van der Laan, RIVM, NL
14:40 – 15:20	Toxicology including Safety Pharmacology. Tim Hammond, Astra – Zeneca, UK and David Jones, EMEA, UK
15:20 – 15:40	Coffee Break
15:40 – 17:40	**Session II: Replacement in the development of new chemical entities (cont)**
15:40 – 16:20	Mutagenicity and Cancerogenicity. Philippe Vanparys, Altoxicon, BE and Jan-Willem van der Laan, RIVM, NL (replacing Doris Höschele, EMEA, DE)
16:20 – 17:00	Reprotoxicology. Michael Schwarz, University of Tübingen, DE and Klaus Olejniczak, EMEA, DE
17:00 – 17:40	Pharmacokinetics. Mario Monshouwer, JNJ, BE and Sonja Beken, FAGG-AFMPS, BE
17:40 – 18:10	**Free communications**
17:40 – 17:55	Replacement, by Jon Richmond, Home Office, UK
17:55 – 18:10	Conclusions of the day Gisbert Sponer, set, DE

3 October

09:00 – 10:45	**Session III: Replacement in the development of new biological entities**
09:00 – 09:25	Keynote lecture “Vital benefits for humans by replacing animal experiments?” Levente Pencz, Fauna Society, HU
09:25 – 10:05	Biologics I – sera, vaccines. Coenraad Hendriksen, Netherlands Vaccine Institute, NL and Steffen Gross, Paul-Ehrlich-Institute, DE
10:05 – 10:45	Biologics II – peptides, humanised antibodies etc. Bernd Mueller-Beckmann, Roche, DE and Beatriz Lima, EMEA, PT

10.45 - 11.05	Coffee Break
11:05 – 11:45	**Session IV: Replacement in quality control**
11:05 – 11:45	Aspects of quality control, requirements of the pharmacopoeia. Jean-Michel Chapsal, Sanofipasteur, FR and Karl-Heinz Buchheit, EDQM, FR
11:45 – 12:15	**Free communications**
11:45 – 12:00	*In vitro – in vivo*: alternative or complementary approaches? István Gyertyán, Gedeon Richter, HU
12:00 – 12:15	Replacement: Laying the groundwork for change. Katy Taylor, European Coalition to End Animal Experiments, UK
12:15 – 13:00	**Round table discussion**
13:00 – 13:10	Concluding remarks. Gisbert Sponer, set, DE

Scientific committee:
Gisbert Sponer, member of SET, Bioassay GmbH, DE
Lajos Balogh, member of Hucopa, OSSKI, HU
Bernward Garthoff, treasurer *ecopa*, Bayer, DE
Vera Rogiers, chair *ecopa*, VUB, BE

5.6.3.1. Introduction to the workshop

This workshop is the third one in a series of three, each of these is concerned with one of the R's in the 3Rs concept of Russell and Burch, namely "Refinement", "Reduction" and "Replacement". This workshop is dealing with the third R, namely the issue of Replacement. A welcome address was given by Lajos Balogh (member of NCP Hucopa, HU). Gisbert Sponer (member of the NCP SET, DE) gave an introduction to the general aim of and ideas behind this workshop. Each day of this workshop was opened with a keynote lecture. Then every theme was dealt with by two experts, of which one of the field of industry or academy, the users, and one involved in the regulatory side of animal testing.
Vera Rogiers, chair of *ecopa*, presented the structure and aims of the START-UP project and explained the role of the Expert Meetings and their input in this third Workshop.

5.6.3.2. Summary of the presentations within the different sessions

5.6.3.2.1. Session I: Replacement in the development of new chemical entities

5.6.3.2.1.1. Keynote lecture by Friedlieb Pfannkuch, Hoffmann-La Roche, CH

This presentation focused on toxicology, because this topic is most important for safety aspects of new pharmaceuticals. There is a public interest on the one hand to allow only products in the market with a well-characterised safety profile and on the other hand to reduce, refine, replace animal testing for ethical reasons. Due to legal aspects there is also an interest of the pharmaceutical industry to perform the necessary animal studies in order to avoid harm to volunteers and patients and liabilities, which can harm the reputation of the company. Industry's activities are mainly driven by regulatory requirements.
Regarding replacement, there are some limitations based on scientific reasons. In particular, cells or *in vitro* systems cannot really provide a reliable surrogate for a complex and complete biological organism, i.e. no alternative methods are available allowing reasonable safety assessment on endpoints after repeated dosing. From the regulatory side, animal studies for risk assessment at marketing authorisation stage are requested.
In safety testing, there is always a battery of studies and it will be difficult to replace them all, because there are many different tests with a different degree of complexity.
Nonclinical safety testing was in particular intensified after the thalidomide disaster in 1956. In recent years a variety of test systems and a steadily growing number of testing parameters have been introduced, consisting of biostatistical methods, GLP (good laboratory practice), PK/PD modelling, toxicokinetics, *in vitro* testing, -omics technology, transgenic animals and stem cells and regulatory needs and global

harmonisation. Many tests have to be done and the total procedure gets more and more complex. One key problem, however, is that the predictability for side effects in humans based on results generated in mice or other animals is by far not ideal. About twenty percent of the withdrawals of drug candidates due to toxicity occurs during clinical trials. This is the case when one considers the “classical” drugs with low molecular weight. However, about half of the classical toxicology tests are not of real help for biopharmaceuticals such as human proteins and therapeutic antibodies which are important new drugs in the pipeline. Another aspect is that animals may be exposed to too high doses, which causes many false positive results and has a negative impact on the potential extrapolation to man. Testing of multiple endpoints may also contribute to false positive correlations. The activities of ICH contribute to the general wish to use fewer animals in safety investigations on a global scale, but it works slowly and the actual complicated international consensus process hinders critical scientific modernisation.

Despite this evolution, radical changes in mind are needed. When developing new candidate drugs, all existing information about the candidate itself and structurally similar substances must be used. Many companies have a lot of *in silico* data and it would be good if these data could be shared, without lo sing competitiveness. When performing animal tests, they should be well targeted. Toxicology should include knowledge from both the identified human and animal genomes.

Nevertheless, a new safety testing strategy should be considered. High-throughput methods based on identified pathways of toxicity with human cells, invertebrate species, etc and computational methods, *in silico* expert systems, must be used more. Instead of exposing animals to high doses and observing a multitude of possible effects, precise questions must be raised about whether sensitive physiological processes are disturbed.

In a mid-term to far future perspective, development of batteries of sensitive and specific safety biomarkers in a collaborative approach is a promising approach. A shift from non-clinical to clinical studies, can take place if good biomarkers are available. ‘Idiosyncratic’ toxicity could better be tackled by pharmacogenomics than by animal testing.

A number of biomarkers are now undergoing the process of validation and acceptance by FDA and EMEA, including biomarkers for liver, kidney and muscles. Kidney biomarkers accepted by FDA and EMEA, include KIM-1, Albumin, Clusterin and Trefoil Factor-3 and Total Protein, β2 Microglobulin and Cystatin C. Liver biomarkers under validation by C-Path, include PON-1, MDH, PNP and GLDH (which are identified for initial cross-qualification) and alpha-GST (multiplex assay expected) which is under discussion. There are different isotypes of muscle biomarkers, involved in regulation of muscle contraction: troponin C binds Ca++ (identical in heart and skeletal muscle), troponin I is an inhibitor of actin-myosin interaction (cardiac and skeletal muscle isoforms), for this a sensitive test is available and troponin T links troponin complex (C, I & T) to tropomyosin (cardiac and skeletal muscle isoforms).

To conclude, the scientific base of safety testing must be improved, the reliability of extrapolation from animals to humans has to be relevantly improved, which may then lead to replacement of some animal tests. On the mid-term to far future perspective, the basis for a potential shift from non-clinical (animal) to clinical studies must be further developed. Batteries of most sensitive and specific safety biomarkers must be developed. The ICH process is the most appropriate platform for the replacement of animal testing.

Contribution to public risk awareness and acceptance can be achieved by an improved information policy.

Discussion

van der Laan: The ICH is said to be working slow and being complicated, but sometimes ICH works too fast and certain documents end up being not well formulated, because of time pressure. It also is a fact that the interpretation of the guidelines done by the industry, is often too strict. In general, the guidelines should rather be used as flexible advice, in particular when there is good scientific argumentation for the modifications that one wants to introduce during certain safety experiments.

5.6.3.2.2. Session II: Replacement in the development of new chemical entities

5.6.3.2.2.1. Preclinical pharmacology, view from industry, by Klaus-Dieter Bremm, Bayer Healthcare, DE

As long as there are unmet medical needs, the role of the pharmaceutical industry is to develop and provide patients with innovative, safe and efficient medicines. For ethical, scientific and regulatory reasons and despite significant technological progress, currently this mission is depending upon studies in animals. Whenever performing such studies in animals, the pharmaceutical industry fully applies the 3Rs. When alternatives, such as computer modelling or cell cultures are available, they are used. Often answers can only be given by testing in a complex living organism.
The evolution in technologies led to a paradigm shift in R&D. 20 years ago, a compound was tested nearly exclusively in animals with the consequence that only few compounds could be tested and many animals were needed. Nowadays, much more compounds are tested with fewer animals. Within these 20 years the number of animals used in pharmacological investigations has been reduced by 75%. This has been possible by using new approaches such as target identification and validation and lead compound identification and optimisation, often based on -omics.
Non-animal alternatives are used wherever possible. There are several reasons for this change, amongst which the fact that animal testing is expensive due to breeding and housing costs. Alternatives are more cost effective. Computer models and cell cultures are good tools for screening tests and are therefore used frequently. However, such models cannot provide reliable information about complicated interactions in the whole system. The combined use of animal and non-animal tests may be the best approach.
More than 95 % of all mouse genes have homologues or orthologue forms in man, which means that mice have a similar, but not identical genome compared to humans. Therefore, it always has to be kept in mind that animal models only give a hypothesis for the human situation.

The use of animal models in R&D is subject to a number of challenges. A first problem is the high attrition rate, there is a 3% chance to reach preclinical development for a NCE (New Chemical Entity) against a new target versus 17% for a validated target. New targets often lack the essential expertise about their biological functions. Due to diverting technologies, drug development for different targets becomes increasingly complex and costly. There is also an innovation gap, the pharmacological industry must create space and incentives for entrepreneurship and innovation. Polypharmacology is a major challenge for the development for the gene-to-lead drug discovery concept and clinically differentiated therapies.

Nowadays an integrated technology platform serves as a prerequisite for a successful translational medicine strategy. Screening of targets, lead finding and optimisation, are done with high throughput methods and computational chemistry. Animal experimentation is only performed in the developmental stages of a new pharmaceutical product, but this is done in conjunction with *in vitro* and *ex vivo* tests, such as toxicology profiling. This helps reducing the number of animals. For example, the use of imaging technology also in small animals provides the possibility to observe the development of a disease continuously and the response to a new drug can be observed more precisely. Using the conventional method, 20 animals were needed for a distinct investigation, whereas using the small animal imaging technique only requires 5 animals to generally achieve equivalent results.

Integrative pharmacology allows complete bio system analysis: target validation, safety, efficacy, DMPK (drug metabolism and pharmacokinetics), surrogate markers. This has become possible thanks to the availability of the human and mouse genome and transgenics. Mice can easily be genetically modified to mimic human disease. Mouse miniaturisation is another new process, which has the advantage that lower amounts of the compound are needed for the investigation. On the other hand, such technology may imply also a challenge, since physiological recordings and bio analytical chemistry measurements become more difficult.

Replacement has many advantages, because known principles can be applied to new systems, mechanistic studies can be extended in technically less challenging cellular models and are less expensive. In particular, cellular models can be used to screen large numbers of agents for activity or toxicity. The disadvantages, however, are that these alternative models are dependent on pre-existing information and it is difficult to incorporate the complexity of living organisms, due to unknown variables in physiology and pathology in living organisms. Therefore, new biological processes will require validation in living organism. Thus, regarding biomedicine aspects, animal tests will still be performed in the coming years.

To conclude, the evolution of technologies in drug discovery and development provided not only the possibility to incorporate the 3R concept in the development of new pharmaceutical products but, in fact, it has been exercised by the companies. Improved study design and focussed testing strategies help to reduce the animal numbers and replace some animal studies. However, despite all efforts, studies in animals still play a significant role in the industry’s mission to provide innovative and safe medicines for complex diseases. The pharmaceutical industry recognizes the importance of the highest possible ethical standards in animal welfare where animal testing is still necessary.

Moreover, it strives together with all stakeholders to continuously reduce the number of animal tests. Together with all stakeholders the goal of the pharmaceutical industry is to broaden the number of validated and accepted alternative methods to replace animal tests. However, for preclinical studies animal studies will be needed, at least for the next 20 more years.

5.6.3.2.2.2. Replacement of animal testing in pharmacodynamic development - a regulatory view, presented by Jan Willem van der Laan, on behalf of Doris Höschele, BfArM, DE

Primary pharmacodynamic studies are concerned with the mode of action and effects of a substance in relation to its desired therapeutic target, whereas secondary pharmacodynamic studies are not related to its desired therapeutic target.
There are no specific guidelines. In Directive (Dir) 2001/83/EC, Annex I, part I, point 4.2.1, there is a part about pharmacology studies. There are two distinct approaches possible: the first one is the investigation and description of actions relating to the proposed therapeutic use by means of recognised and validated assays, both *in vivo* and *in vitro*. New experimental techniques must be described in such detail as to allow them to be reproduced. The second approach is the investigation of potential undesirable pharmacodynamic effects on physiological functions.
The goals of pharmacodynamic studies are diverse and include elucidation of the modes of action (primary and secondary) (can be done both *in vitro* and *in vivo*), selectivity and sensitivity to the target, which is tested *in vitro* and prediction of possible efficacy in humans (tested *in vivo*). With all these studies, proof of concept is very important. Other goals are the dose-response-relationship of the effect (*in vitro/in vivo*), dose selection for non-clinical and clinical studies (*in vivo*) and selection of the dosing schedule (*in vivo*).
Pharmacodynamic studies are relevant during early drug development, when there are no or only a limited amount of toxicological and clinical data available. They can help with decision-making on further non-clinical/ clinical development of the substance, species selection for toxicological studies, dose selection for toxicological and clinical studies and give information on the MABEL (minimum anticipated biological effect level). Pharmacodynamic studies are also supportive during drug approval, when information from clinical studies Phase I, II and III are available and efficacy has been investigated in humans.Clinical data are more important for the proof of concept, especially when no animal disease model is existing.
For pharmacodynamic studies, the same rules as for other studies apply. Under article 7 of Dir. 86/609 it is stated that if good alternatives are available, then no animal tests should be performed. There is much more possibility for development of non-animal studies in the field of pharmacodynamics compared to toxicology because there is no specific guideline for performance. In Dir. 2001/83, the use of new techniques is allowed and there are few standardised models. However, in the majority of cases pharmacodynamic development of a drug relies on animal studies as the pivotal studies before starting clinical trials.

One of the reasons that animal models are still often used is that *in vitro* studies have limitations. It is for example difficult to estimate the pharmacologically active dose *in vitro*. Traditional models which one has with long-term experience are often preferred, because there is not only a lack of validated alternative models, but also experience is lacking.
As a first example, anticancer drugs are presented. Antitumour activity against human glioblastoma cells *in vitro* and in nude mice subcutaneously implanted with glioblastoma cells are both possible therapeutic options for treatment of glioblastoma in humans. However, *in vivo* distribution studies in rats showed low distribution into the brain. *In vitro* effects, on the contrary, are not under influence of the pharmacokinetics of the substance.
With anticoagulant drugs, *in vitro* studies showed selectivity and sensitivity for inhibiting coagulation factor and blood clotting. *In vivo* studies, however, are still necessary to investigate the effects on bleeding time at antithrombotic doses. These effects can only be investigated in a whole organism.
In vaccines, *in vitro* studies showed response of dendritic cells to the antigen and adjuvant. *In vivo* studies are necessary to investigate effects on protection against the invader (virus or bacteria). These effects can only be investigated in a whole organism, preferably an animal sensitive to the disease, e.g. influenza in ferrets.
The National Cancer Institute (NCI) Development program for anticancer drugs has, in the early 90's replaced *in vivo* murine p388 leukaemia model used as a pre-screen model, by an *in vitro* human tumour cell line assay comprised of different cell types. For HIV (Human Immunodeficiency Virus) infection, there are no relevant animal models, therefore the non-clinical pharmacodynamic data relies on *in vitro* studies. *In vivo* pharmacodynamic studies are complemented by *in vivo* pharmacokinetic safety studies. A question to be asked is whether animal studies are necessary, when no relevant animal model exists. A final answer does not exist so far. *In vitro* screening tests are used on a range of receptors, enzymes and ion channels in secondary pharmacodynamic studies. However, also for secondary pharmocodynamic studies in vivo studies are still requested.
To conclude: In most cases, development of a drug relies on whole pharmacodynamic animal studies. No standardised models are recommended in guidelines. The development of more animal-free methods in this field is possible, and this is and should be applied wherever possible. Final confirmation by simulation of a whole organism actually is impossible by *in vitro* tests only. If no relevant animal model of disease exists, the performance of animal studies is questionable and should be carefully considered.

Discussion

Garthoff: Has there ever been a final acceptance without animal model?

van der Laan: More clarity is still needed, by now the regulator has to decide whether or not the testing is sufficient. In HIV related research and with antibiotics, there are not always infection models available.

Bremm: Bayer would probably not step into a project where no animal model is available. Even for antibiotics, there were always models for all relevant pathogens.

Hammond: Small molecules do not exert their toxicology only through their pharmacology, but also through their chemistry, which is much more difficult to predict.
Ten years ago, AstraZeneca looked at the reasons why drugs failed during their development. Around 50% of the candidates failed, because of an exaggerated pharmacological response, but the other 50% were failing for reasons, not related to pharmacology. When the chemical structure of a compound is changed slightly, this can result in a totally different toxicological outcome. The challenge is to find the right screening methods to predict what will happen in the *in vivo* situation.

Bremm: The problem also is that often different tests must be used for different chemical core structures. Tests might work with one class of structures, but not with another one.

Gross: In research related to biopharmaceuticals such as monoclonal antibodies often no animal models are available. Knock out models can be produced, but this remains difficult. There are examples where no appropriate animal models are available.

van der Laan: New compounds in the pharmaceutical field are new by intention. This is not the case in the chemical industry. This makes that there is by definition some unpredictability in the *in vitro* systems.

Vanparys: In the past, at J&J there was a compound, which induced cataract. After testing in lenses, it became clear that this effect was caused by an inactive enantiomer, which was not active in the proposed indication. Now, J&J also performs chemical screening.

5.6.3.2.2.3. Replacement in Toxicology and Safety Pharmacology, the reality, by Tim Hammond, AstraZeneca, UK

For ethical, scientific and regulatory reasons, and despite significant technological progress, the development of innovative, safe and effective medicines is dependent upon some studies in animals. These are a small but vital part of the development of a new medicine. The principles of the 3Rs are applied as an integral part of drug discovery and development. Replacement is inherent in screening cascades that are used to select the best potential new medicines for animal studies. Law requires regulatory toxicology, because it actually is the best way to protect patients and volunteers.
Most compounds are killed before the clinical testing phase. Scientists apply a “reductionist” approach nowadays, thanks to target identification, due to which more compounds can be tested with less animals. One problem, however, is that the predictability is decreasing.
Replacement is good business, because using animals is expensive, resource intensive and time consuming. Alternatives offer many advantages, such as lower costs, a lower

amount of testing material is needed and high throughput testing is possible. However, *in vitro* reductionist approaches cannot mimic the complexity of an intact organism. *In vitro* systems are also poor predictors for systemic and chronic effects in man.
Data mining, *in silico* prediction, are fast working tools for better utilization of internal and external unstructured data. Reduction should be seen in context of better knowledge management. However, this system is not yet perfect, since the understanding of pathways must get better.
Animal tests for toxicology and safety pharmacology are required by law, and scientifically supported by ILSI (International Life Sciences Institute) concordance data. Animal testing is not perfect, but currently it is the best way to protect volunteers and patients. Animal use in regulatory safety assessment represents less than 10% of the total animal use in research and development of new medicines, but the focus for replacement has been very much on toxicology. The question could be posed whether this focus is not disproportionate.
Replacement is science and technology driven, not legislation driven. Despite being a small percentage of the total animal use, there are a lot of efforts and funding possibilities (ECVAM, EPAA, FP's…) in safety assessment for replacement. Replacement alternatives of other sectors, however, will not always be relevant for the pharmaceutical sector, due to the differences in properties of the chemical entities. Global regulatory acceptance is also very important, before alternative methods can replace animal testing.
In vitro methods in safety pharmacology and toxicology are mainly used for compound selection. They are also good for understanding mechanisms, and are as such used for in depth profiling of activity at primary target and secondary effects, kinetics and metabolism, general cytotoxicity, hepatoxicity… They are also used in regulatory safety testing, for example in cardiac toxicity (hERG (Human Ether-a-go-go Related Gene)) and phototoxicity (3T3 NRU (Neutral Red Uptake)).
In the first example, with respect to cardiac toxicity, a refined cascade to prevent sudden death associated with heart arrhythmia has been developed. The screening cascade now used, is that first a virtual screening with a hERG computer model is done, then a screening for hERG channel binding in cell lines, next testing for the effects on the heart rhythm in animals and finally in man. This works well, but the problem is that even the best *in vitro* tests only reach a prediction accuracy of about 80%.
Potential phototoxicity, is tested *in vitro* with 3T3 NRU PT cells, which is an ECVAM validated replacement for *in vivo* phototoxicity testing. This test is adopted into EU regulatory guidance. Also an *in vitro* photo-clastogenicity test (ChromAb test with CHO (Chinese Hamster Ovary) cells) and an *in vivo* photoallergy test are performed. 50-80% of the drug candidates demonstrate some absorption with 290-700 nm, >50% of the drug candidates would be detectable in either the skin or eye. Therefore, most drug candidates need photosafety testing. The problem is that *in vitro* photosafety assays are oversensitive. Using the 3T3 NRU PT assay, 45% of the compounds were tested positive, whereas using photoxicity animal studies, only 15% of these 3T3-positives were tested as positive. If 3T3 NRU PT is used in decision-making the danger exists that, potentially valuable drugs might get lost.
The use of non-human primates is addressed in the Wetherall Report (report by Sir David Wetherall, commissioned by the Medical Research Council, The Royal Society, the

Academy of Medical Sciences and the Wellcome Trust (UK), Times, Dec. 16, 2006) as follows: “There is a strong scientific case for the carefully regulated use of non-human primates where there are no other means to address clearly defined questions of particular biological or medical importance.” The SCHER Report (Scientific Committee on Health and Environmental Risks, Scher Report on Non-human Primate Research, 86/609/EEC) (January 2009) states that the use of non-human primates may be essential for specific pharmaceutical development, for getting more understanding of infectious diseases, such as AIDS, as well as for getting better understanding of complex neurobiology and for studies regarding xenotranplantation.
Some future topics are stem cells, computational chemistry and systems biology. Stem cells are an interesting field with a lot of potential, but it is not reality yet. Computational chemistry is already valuable today. Systems biology is promising for the future.
Despite enormous advances in technology very few replacements currently exist besides those for some local or acute effects, but safety assessment for pharmaceuticals is primarily driven by risk assessment from chronic systemic toxicity studies. The primary bottle neck is the complexity of the scientific challenge.

Discussion

Garthoff: What are the limitations of sharing data?

Hammond: Duplication of data is not the same as sharing. Replication is sometimes done, duplication not. A balance between competitiveness and sharing should be reached. Data sharing is not limited by competitiveness, but by time.

5.6.3.2.2.4. Chances and Limitations With Respect to Replacement of Animal Experimentation in the Development of New Pharmaceuticals, by David Jones, MHRA, UK

MHRA (Medicines and Healthcare products Regulatory Agency) is the executive agency of the Department of Health safeguarding public health by ensuring that all medicines and medicinal devices on the UK market meet appropriate standards of safety, quality and efficacy. MHRA does not perform animal testing itself, but believes that some animal use will remain necessary for safety evaluation purposes at least for the near future. MHRA is opposed to the indiscriminate use of laboratory animals. If appropriate parameters are monitored, duplication of studies may be avoided. Optimum design in gathering data will reduce the number of animals required.
When performing safety pharmacology studies, it is important to adopt a rational approach to select the studies. Regulatory guidance is only one way of achieving an objective, there might be a better way. The specific studies that should be conducted and their design will vary based on the individual properties and intended uses of the pharmaceutical. Safety pharmacology is not a box ticking exercise.

Safety pharmacology studies could also be used in the design of toxicological and other preclinical studies. Safety pharmacology is some of the most important non-clinical studies, to be conducted prior to first use in man. Since pharmacological effects will vary depending on the specific properties of each tested substance, the studies should be selected and designed accordingly.

There are changes in the use of animals, for example, although the dossiers are getting larger, more and more *in vitro* studies are used, which proves that some applicants are not just "ticking boxes". More *in vitro* work is done in the early research stages, but little in the development stages. There is not much real replacement, but things are changing in refinement and reduction.

In ways of refinement, toxicokinetics could be very useful in safety pharmacology studies. For example, a caudal venepuncture is often done. When repeated withdrawal of blood is needed to properly define PD/PK relations. Extra satellite animals may be used for this purpose. Dried blood spots is a "new" development in this area, even though the technique has actually been around for over 40 years. It is an easy way of collecting, shipping & storing blood samples. Animal or human blood is bled directly onto a collection card, which lyses cells and denatures proteins. The cards are air dried & stored or shipped desiccated at room temperature. Discs are punched out of the dried blood spots for analysis. The advantage of this is that reduced blood volumes are used, which is useful in juveniles. It also reduces the animal usage, as serial sampling can be done in one animal, rather than collecting composite bleeds from several animals. There is also the possibility to remove satellite rodents. A number of pharmaceutical companies are currently using this technique and inter-laboratory validation exercises are underway.

When planning experiments, it is important to carry out a proper statistical analysis to determine how many animals are needed. Also all aspects of the experiment must be properly designed. If experiments fail and have to be repeated, additional animals have to be used, which defies the goal of reduction. A case-by-case vision is often the best way to design experimental studies.

A lot of scientific effort has been devoted to developing new, non-animal techniques, which can be used in experiments instead of animals. There have been some notable successes, but overall, progress has been disappointingly slow. One reason why *in vitro* methods are used more often now is that they are cheaper.

The ICH M3R2 guideline on Nonclinical Safety Studies, which was approved in June 2009, should facilitate the timely conduct of clinical trials, reduce the use of animals in accordance with the 3R principles and reduce the use of other drug development resources. Although not discussed in this guidance, consideration should be given to use new *in vitro* alternative methods for safety evaluation. These methods, if validated and accepted by all ICH regulatory authorities, can be used to replace current standard methods.

According to this guideline, in safety pharmacology, consideration should be given to inclusion of any *in vivo* evaluations, e.g. as additions to general toxicity studies, to the extent feasible, in order to reduce animal use. Stand alone acute toxicity studies are not recommended.

Other efforts include the acceptance, in some cases, of 6 month rather than 9 month non-rodent toxicity studies in the USA and Japan. Exploratory clinical trials in humans

section are added. These can be initiated with less, or different, nonclinical support than is generally warranted for clinical development trials. It is also accepted that reproductive toxicity studies are not needed until later in clinical development.
It is also accepted that the conduct of any juvenile animal toxicity studies should be considered only when previous animal data and human safety data, including effects from other drugs of the pharmacological class, are judged insufficient to support paediatric studies. The guideline also recognizes that existing phototoxicity guidelines are inappropriate.
The new ICH S9 Guideline and Addenda to S6 Guideline should also have impact on animal use.

Discussion

Vanparys: In the M3R2 guideline, it is stated that acute toxicity testing is not recommended anymore. Is this only for general toxicity, or also for genotoxicity testing?

Jones: Single dose toxicity testing for genotoxicity is still needed, but not acute toxicity testing anymore. Death is no longer the endpoint and just giving animals high doses to see what happens, is not done anymore.

Unidentified: *In vitro* models for high throughput screening are indeed cheaper than *in vivo* models, but when complex batteries for replacement are used, they might be as expensive, or even more, as *in vivo* models.

Jones: When comparing a single *in vitro* study against a single animal model, then *in vitro* models are cheaper. Validating *in vitro* models is indeed expensive.

Hammond: The challenge of companies is to select compounds, *in vitro* systems can screen thousands of compounds. If a system is found on which decisions can be based, it is worth for the investment, because then the company can get better compounds.

Pfannkuch: Less than 10% of the animals are used for toxicology. Shouldn't there be a focus on those fields where the highest amounts of the animals are used?

Bremm: A large part of the animals is used in basic research and kinetics studies. Workgroup 2 of EPAA is doing an analysis to see where there are other opportunities to reduce the number of animals. Toxicology is not the major driver in this analysis.

Unidentified: Do EPAA or any other organisations give guidelines on the use of juvenile animals?

Jones: Yes, there are juvenile guidelines and workshops focussing on these guidelines. One problem is that the FDA validates models by letting companies run studies to

investigate whether the approach works or not. Companies do studies that might have no relevance at all. This makes no sense to EU regulators.

Maier: When comparing *in vitro* to *in vivo* studies, they are always compared to each other, while in fact they should be compared to the relevance to the human. This is often not possible, because harmfully positive compounds are not tested in humans.

Jones: The problem is that statistics ignore investigational drugs that have never been tested in humans due to potentially harmful findings in preclinical studies. Animal studies do not show real "human" side effects, that's what the clinical trials are for. Animal studies are targeted to support the start of clinical trials. Clinical trials are of key importance, but test groups always remain small and are limited in covering the variety of the "whole" population. Also, in reality people don't always follow the prescription, which can cause side effects (overdose for example). Nothing can be done about this and regulators or industry are not to be blamed here.

Hammond: The normal concept of screening cascades is going at the beginning from *in vitro* models to man at the end. Nowadays, there is a second approach, since often compounds that are already on the market, might be identified later on to cause harmfully positive signals in test systems. This information is then also used again in computational models.

5.6.3.2.2.5. Mutagenicity and Carcinogenicity, by Philippe Vanparys, Cardam, BE

Genetic toxicology testing has gone through an evolution, starting with purely *in vivo* testing, to nowadays, mostly *in vitro* testing.
At Johnson & Johnson (J&J), in 1978, the dominant lethal test was used, since genetic toxicology was still new. This test used 1.000 mice. Nowadays only one *in vivo* assay is performed.
In drug evaluation, regulatory authorities require three tests, the Ames and mouse lymphoma tests, which are *in vitro* tests, and the micronucleus test, which is an *in vivo* test. These tests are used to detect gene mutations and chromosome aberrations. The question is whether this *in vivo* test can be replaced.
A number of initiatives to reduce and replace animal tests in genotoxicity testing have been taken and are underway. The ones that will be discussed here are the ECVAM workshop, the ICH draft genotoxicity guideline S2 (R1), the REACH Integrated Testing Strategy and the 5th International Workshop on Genotoxicity Testing.
On the ECVAM workshop, two conclusions were reached. The integration of the micronucleus assay into Repeated Dose Toxicity (RDT) studies should be standard when RDT studies are foreseen for the test compound. A recommendation was made to combine acute micronucleus and comet assay studies into one study.
The ICH has drafted a genotoxicity guideline S2 (R1). ICH drafted a testing cascade, starting with a test for gene mutation in bacteria (Ames test), then there are two options. Using i*n vitro* mammalian cell tests, mouse lymphoma or *in vitro* chromosome aberration

test, can be performed. Depending on the results, there are two possibilities. If this test is negative, a micronucleus test integrated in RDT can be carried out. Under these conditions, only one endpoint is needed. If the test is positive, two endpoints are needed, one done in a separate *in vivo* study. If no *in vitro* mammalian cell tests are performed, also two endpoints are needed. ICH also discussed the integration of micronucleus test as endpoint into RDT studies. This counts for 14 to 90 day RDT, dose range finding and teratogenicity studies, however there are still some issues, about the highest dose to be tested and the satellite group of positive controls. In acute toxicity studies, no positive control animals are needed. Micronucleus scoring can be done in blood by using flow cytometry, acridine orange staining and manual scoring.

The recommendation of the REACH Integrated Testing Strategy is that genotoxicity tests can be integrated into RDT, if it is scientifically justified.

The International Workshop on Genotoxicity Testing, focussed on the reduction of false positive results *in vitro* and the reduction of animal usage. A new subgroup was set up to develop consensus recommendations for integrating genotoxicity endpoints into RDT studies and to study multiple endpoints in short-term studies, in light of recommendations of ICH, ECVAM and animal welfare organisations which all are in favour of reducing animal usage.

The first topic was the combination of micronucleus and comet assay into acute toxicity. A possible design of a bone marrow & blood micronucleus test plus comet assay in tissue of choice was presented. To take blood, it is best to use rats. They are treated at time 0 and blood is taken in order to get base line values, thus, the animal acts as its own control. Then it is treated at 24h, blood is taken again. It is treated a third time at 45h, comet is sampled after 3 hours, animals are sacrificed at 48h, a blood sample is taken again. Micronucleus tests are done at each time. Bone marrow is collected at 48h for micronucleus testing. Comet assay can be done on stomach, liver or other material. This is technically feasible and scientifically acceptable as an alternative to the separate assays.

The second topic was the integration of micronucleus testing into RDT studies. This is scientifically acceptable, but there may be situations where an acute study is preferable (for example with severe bone marrow toxicity).

The third topic is the integration of comet assays into RDT studies. This is considered scientifically acceptable. The liver comet assay complements micronucleus tests in blood or bone marrow in detecting *in vivo* genotoxins. Practical issues still need to be considered.

Conclusions on the replacement in genotoxicity testing are that the only routine in *in vivo* testing is the micronucleus test. Replacement of the acute micronucleus test by an *in vitro* micronucleus test for pharmaceuticals is not expected for the near future as the *in vitro* system does not mimic the possible metabolism of the test compound which may occur *in vivo*. The *in vitro* test is not mature enough to replace the pivotal regulatory *in vivo* test for human safety testing. Integration of the micronucleus and combination with comet endpoint in RDT studies can be seen as a partial replacement of acute micronucleus test. Implementation of *in vitro* micronucleus and comet tests in the discovery phase of drug research to deselect the genotoxic compounds should also be seen as a partial replacement of *in vivo* (Mammalian Erythrocyte Micronucleus Test) micronucleus and

comet testing. Integration is in compliance with OECD Guideline 474, strongly encouraged at the ECVAM workshop and recommended by ICH and REACH ITS.
For pharmaceuticals, there is no alternative *in vitro* test, which can replace the 2-year carcinogenicity studies, because the complex mechanisms for carcinogenicity induction and *in vitro* models do not mimic metabolisation pathways in humans. The Cell Transformation Assays (CTA) are the only available *in vitro* alternatives for carcinogenicity testing which may play their role in chemical testing at low human exposure conditions but cannot serve as a final test in the pharmaceutical industry.
In the long future, no accepted and validated *in vitro* test will be available for full replacement of the long-term carcinogenicity tests in drug research.

Discussion

Maier: Why are positive controls needed?

Vanparys: Positive controls are not necessary if one has experience with the test, but since comet assays are still new, positive controls are necessary.

5.6.3.2.2.6. Carcinogenic Risk Assessment of Human Pharmaceuticals: A European Regulatory Perspective, by Jan Willem van der Laan, RIVM, NL

In the old paradigm of carcinogenicity studies of human pharmaceuticals, the duration and exposure was enhanced as much as possible to have a maximum risk on the development of tumours in the animals. Life time studies were performed, being 24 months in rats and 18 or 24 months in mice. Two species were used to correct for species differences.
The assumption was that genotoxic, DNA-damaging compounds were covered by a genotoxicity battery (ICH S2), including Ames test, mammalian cell assay and *in vivo* assay on chromosome aberrations. The presumption was that any genotoxicant will be carcinogenic, unless proven otherwise (ICH S1A). However, non-genotoxic compounds, not acting via DNA damaging mechanisms must be studied in two year studies in rodents at high exposure (MTD (Maximum Tolerant Dose) or multiple of AUC (area under the curve), as defined by ICH S1A, B, C.
50% of all chronically used human pharmaceuticals induce tumours in rodents. But, only 20 human pharmaceutical carcinogens have been identified by epidemiology, although a lot of epidemiological studies have been carried out.
There are several issues in carcinogenicity research, one being that rodents are more sensitive than humans. Often the pharmacodynamic activity is the basis of the carcinogenesis. Another problem is that of age-related tumours in rodents. Indeed the tumours arise in the 18 to 24 months of the treatment period. The duration of the treatment is also under discussion: IARC (International Agency for Research on Cancer) compounds induce tumours within 12, and exceptionally within 18, months.

Non-genotoxic human mechanisms of action of carcinogenicity include proliferation of cells, suppression of apoptosis and induction of cell damage with subsequent restoration. Examples of this are chronic cell injury, immunosuppression, increased secretion of trophic hormones and receptor-mediated mechanisms.
The ILSI project Alternatives in Carcinogenicity Testing used transgenic and knock-out animals, as well as the *in vitro* SHE (Syrian Hamster Embryo Cell Transformation) assay. Ten to twenty compounds were evaluated using a transgenic mice strain. The SHE assay was too sensitive to be used for regulatory purposes, but might be useful in early screening of new compounds. It has no predictive value for human carcinogens, but for rodent carcinogens.
There is hardly any new approach for this cost-intensive project, despite the involvement of regulators from various authorities. The number of studies for regulatory purposes is small. It seems that the regulatory affairs people from companies are conservative It is also possible that there is a lack of evidence that the knock-out and transgenic mice are a good alternative.
There are a few other approaches such as QSAR (Quantitative Structure Activity Relationship). Biomarkers are not yet predictive, unless a battery is available.
FDA has an office of pharmaceutical sciences, with an informatic and computational safety analysis staff. MDL (Molecular Design Limited) QSAR contains an integrated set of tools for similarity searching, compound clustering, and modeling molecular structure related parameters, which is based on the atom-type electrotopological state (E-state) connectivity between molecules and other descriptors, which can be statistically correlated with toxicological or biological endpoints. This is based on the FDA/CDER (Food and Drug Administration/Center for Drug Evaluation and Research) database with 1285 compounds. Validation was done with 108 compounds, of which 86 were pharmaceuticals. A cluster with data of male and female rats and mice was made. This gave an estimation of rodent carcinogenic potential of the test compound from a single cell to a two plus cell analysis (FDA is considering a mouse study having two cells, males and females, the same for the rats, 2x2 cells). This database provides some hope that the outcome of a single species study for carcinogenicity could be predicted. The coverage was 92%, the specificity, sensitivity and predictability were in the range of 70%. The authors stated that “electrotopological E-state descriptors and MDL QSAR software are promising new *in silico* approaches for modeling and predicting rodent carcinogenicity and may have application for other toxicological endpoints”. Limitations for carcinogenicity include the observation that tumour findings may arise by chance. It is an application of multi-cell analysis and most reliable predictive models are developed for compounds that produce multi-species tumours.
The major issue in this approach is that rodent carcinogenicity is the gold standard and proliferation is the main mechanism. There is not enough focus on human carcinogenicity.
The accuracy of the life-time rodent bioassay is unknown. Emphasising rodents as gold standard, neglects the high number of false positives, known not to be human carcinogens due to their mode of action. Emphasising the importance of multi-species carcinogens, neglects that 3 out of 6 human carcinogens tested are only positive in one species.

Epidemiological data are needed, based on a broad ranged of human carcinogens and human non-carcinogens.
The genomic approach is a technical approach. It can be used as genomic profiling, pathway analysis, as a translational approach and to learn more about human cancer mechanisms. In the translational approach, animal pathways associated with tumours are compared to human pathways. It can also be used in human cancer mechanisms to look at disturbances in human cell metabolism associated with the pathogenesis of cancer, thus, mechanisms involved in human cancer are analysed by genomic techniques. May be these data should be used as the starting point for toxicological analysis. In screening, there are some issues to detect these pathways. Questions are: "should *in vivo* or *in vitro* studies be used?", "must transgenic animals be used?", "what is the duration of treatment and is it possible to do risk assessment?"
Replacement is not possible in the short term as another gold standard is lacking and human pathways need to be characterized.

Discussion

Maier: If an *in vitro* test would be performed as a lifetime test, it would not be accepted. If *in vivo* studies are performed in lifetime, they are accepted. In this matter, there should be a new way of thinking.

van der Laan: An *in vitro* study as replacement of an *in vivo* lifetime study is insufficient if it is just predicting the outcome regarding the carcinogenicity during the lifetime of the *in vitro* system . The SHE assay predicts the outcome of the rat assay, but it is not acceptable, because the rat assay gives more information (for example which organs are bearing tumours). This information is helpful in finding the mechanism. This is not possible with the SHE assay.

Garthoff: It is regularly heard that industry doesn't want rodent carcinogenicity as a gold standard and that new info is needed. But a lot of data is available, for example from translational medicine. Then the question is who will make work of sharing this data and compiling a new standard?

van der Laan: The database is not yet at the level that carcinogenicity can already be predicted. This is the way forward, this should be worked out now and when this is ready and working, then it is time to share data and hopefully it will be done.

Hammond: Industry did already some data sharing (with PPAR-γ agonists (peroxisome proliferator activated receptor), i.e. glitazones). All data related to these compounds illustrated the complexity of the problem. Doing pathway analysis is extremely complex.

5.6.3.2.2.7. The ReProTect Framework Program: New Innovative Approaches for Evaluating Fertilization, Implantation and Prenatal Development, by Michael Schwarz, University of Tübingen, DE

ReProTect is an EU FP6 project targeting the development of a new approach in hazard and risk assessment of reproductive toxicity by a combination and application of in vitro, tissue and sensor technologies. It is divided in four research areas: fertility, implantation, prenatal development and crosscutting technologies, with the goal of reduction and/or replacement of animal use in reproductive toxicity testing.
Not a single test is able to present effects on the reproductive cycle, but it might be possible to divide the cycle in parts and develop tests that pick up critical endpoints in the reproductive cycle.
In spermatogenesis, there are two (Leydig and Sertoli cell) CASA (Computer Assisted Semen Analysis) tests and a ReProComet assay. In oocytogenesis, granulosa cell and *in vitro* oocyte maturation tests exist. In fertilisation there are *in vitro* fertilisation assays. In implantation, a variety of tests can be used, relying on human explant tissues. In embryonic development, the compound has to reach the embryo to be active, so there are groups working on placenta perfusion assays, QSAR tests, EST (Embryonic Stem Cell test) and a ReProGlo assay (stem cell based reporter). Tests also exist, that are designed to measure effects on the endocrine system.
As a first example the tests for female fertility were presented. For the *in vitro* maturation, the bIVM (bovine *in vitro* maturation) test can be applied, then bovine oocytes are incubated for 24 hours, then maturation is measured and during maturation the oocytes are exposed to the test chemical. *In vitro* fertilisation (bIVF, bovine *in vitro* fertilization) can also be done. The molecular targets for these compounds are quite different, such as microtubuli, signal transduction, gap junctions, DNA, endocrine system or ribosome. In most cases IVM is more sensitive than IVF, because the process of maturation is more complex than that of fertilisation. This test gives mechanistic information, which can help to interpret what a chemical might do when going into *in vivo* studies.
The mouse follicle bioassay can predict effects on ovulation, fertilisation, but also on the endocrine system, because one of the read-outs consist of secreted hormones.
The ReProGlo assay is a cell-based toxicity pathway test for the prediction of chemically induced embryotoxic effects. There is a limited number of signaling pathways for early development, but disturbances in these pathways lead to disturbances in development. Mouse embryonic stem cells were transfected with the reporter for the β Catenin pathway. The test is performed in a high throughput system, where the cells are incubated with the test chemical for 24h. The endpoint is measured after a known invasive cytotoxicity assay within two hours. LiCl and retinoic acid, both not cytotoxic, were used as positive controls. When the cells are incubated with a series of valproic acid derivatives, which differ in teratogenic activity, there is a good correlation between teratogenic activity *in vivo* and the effects in the assay. The assay can be combined with metabolic activation.
The project was started with two already validated tests for embryotoxicity, the mouse Embryonic Stemcell Test (mEST) and the Whole Embryo Culture (WEC). The problem

is that there is no metabolic capacity and, especially for the mEST, the applicability domain is unclear. These problems could only be solved partially in ReProTect. Several assays were developed, receptor binding or cell-based reporter systems, for the detection of endocrine disrupters. These tests are reproducible, transferable between labs and are predictive. Several assays were also developed to predict adverse effects on female and male fertility.
The potential use of these assays is in early drug development, such as in-house use for prioritization during lead compound optimisation. Some companies use the EST in this context. They also give information on mode of action analysis for compounds that have demonstrated reproductive toxicity *in vivo*. This may also be helpful for regulatory decision making. Alternative tests may lead to a reduction in the number experimental animals, but presently not to a replacement of the animal assays.

Discussion

Maier: There are 26 tests, is there one test for each segment of developmental reproduction?

Schwarz: Quite a number of the tests failed to do what they were designed for. One of the tests, a receptor binding assay, is in validation. A variety of other tests is close to pre-validation. A feasibility study, with 10 core tests, is being performed. Ten blinded chemicals are running through these core tests. At the end there is the hope that the effect of these chemicals which they might have *in vivo*, can be predicted by the *in vitro* methods. At the end of the test series the identity of the chemicals will be unblinded and the false positives and negatives will be explained. False positives and negatives are often found in alternative tests, but it is nearly never tested where they come from.

5.6.3.2.2.8. Reproduction toxicology, by Klaus Olejniczak, BfArM, DE

The Commission Directive 2003/63/EC of 25 June 2003 amending Directive 2001/83/EC for reproductive and developmental toxicology, stipulates that the investigation of possible impairment of male or female reproductive function as well as harmful effects on progeny shall be performed by appropriate tests, being it *in vitro* or *in vivo*. These tests comprise studies of effect on adult male or female reproductive function, studies of the toxic and teratogenic effects at all stages of development from conception to sexual maturity as well as latent effects, when the medicinal product under investigation has been administered to the female during pregnancy. Omission of these tests must be adequately justified. Depending on the indicated use of the medicinal product, additional studies addressing development when administering the medicinal product on the offspring may be needed. Embryo/foetal toxicity studies shall normally be conducted on two mammalian species, one of which shall be other than a rodent. Peri- and postnatal studies shall be conducted in at least one species. If the metabolism of a medicinal product in particular species is known to be similar to that in man, it is desirable to

include this species. It is also desirable that one of the species is the same as in the repeated dose toxicity studies. The state of scientific knowledge at the time when the application is lodged shall be taken into account when determining the study design.
The ICH has generated guidelines for guidance on reproductive toxicology, (Detection of Toxicity to Reproduction for Medicinal Products and Male Fertility) and for risk assessment of medicinal products on human reproductive and development toxicities.
This means that fertility studies in males and females and early embryonic development studies in rats have to be done, as well as embryo-foetal development, normally done in rats and rabbits. Prenatal and postnatal development, including maternal function, is also a rat study.
The evaluation of results includes the fertility of males and females, which is the parental toxicity, embryo-foetal development, in maternal toxicity (F0) and F1 embryo-foetal toxicity, including teratogenicity, pre- and post-natal developmental maternal toxicity and F1 developmental toxicity.
In the fertility and early embryonic development to implantation study in rats, male rats have a premating dosage period of 2 or 4 weeks, then a cohabitation period for three weeks, while still receiving dosage. Thereafter, they are sacrificed. Female rats have a premating dosage period of 2 weeks, still get dosage until day 7, by which gestation is presumed. At day 13, foetal evaluation, external only, is done using a caesarean sectioning.
In embryo-foetal development studies, called teratogenicity studies in rats, the females have a dosage period from day 6 until day 17 of presumed gestation. At day 20, a caesarean sectioning is performed, allowing foetal evaluations on the external shape, soft tissues and the skeleton. In rabbits, the only difference is that the dosage period is until day 18 and caesarean sectioning is done on day 29.
In pre- and post-natal development studies, including offspring preweaning period, F0 female rats have a dosage period from day 6 of the gestation, until day 21 of lactation. F1 generation is then looked at both pre- and postweaning period.
Based on the current regulation for pharmaceuticals, replacement of reproduction toxicology animal studies by *in vitro* studies or other methods is not recommended.
However, fertility studies could be replaced by repeated dose toxicity studies to facilitate the treatment of patients with advanced cancer. In some cases, a juvenile animal toxicity study could be replaced by a modified pre- and postnatal study. For insulin analogues, the use of *in vitro* embryo culture studies may be considered. For anticancer pharmaceuticals in the case that an embryo-foetal developmental toxicity study is positive**,** a confirmatory study in a second species is usually not needed.
Reproductive toxicity testing makes a huge contribution to the estimated costs (70%) and the number of animals (90%) used for compliance with the REACH legislation. For pharmaceuticals no second generation study is needed, but this might be needed for chemicals. Reproductive toxicity testing in chemicals therefore uses a lot of animals and it would be interesting to look for alternatives in this field.

Discussion

Vanparys: There is a lot of discussion on these numbers (i.e. 70% and 90%).

Olejniczak: ECHA (European Chemicals Agency) indeed says that the number of animals is lower. But the reproductive part of toxicity testing for chemicals is really high.

Hammond: Most of the risk-assessment done by companies is a combination of pre-clinical studies and clinical trials. Carcinogenic and reproductive risk assessment are based almost entirely on nonclinical data.

5.6.3.2.2.9. Performing animal experiments in support of human pharmacokinetic research: Sense or non-sense?, by Mario Monshouwer, Johnson & Johnson, BE

In the process of going from a molecule to a drug, there are several fields in which pharmacokinetic research is important. In discovery and preclinical development, there is PK/PD (pharmacokinetics/pharmacodynamics) in pharmacology models, optimising of drug-properties and the prediction of human PK and dose needed to start in Phase I of clinical trials. A lot of animals are also used in formulation assessment. Toxicokinetics is started early and goes on until carcinogenicity studies. A lot of satellite groups are used here, so this might give opportunities to reduce the number of animals. PK/PD is also investigated in humans.
In these fields, reduction and refinement are performed, but when it comes to replacement, it is doubtful whether it is possible at all. In human PK/dose predictions and toxicokinetics there might be possibilities for replacement**.** Indeed, the challenge regarding the prediction of the human dose/PK consists in the fact that the prediction of clinical effectiveness relies on validated disease models, which is difficult. Preclinical safety/toxicology profiling is built on the assumption that similar drug exposures will provide similar preclinical and clinical outcomes. Significant species differences in ADME (absorption, distribution, metabolism, excretion) profiles, however, make it difficult to compare results. To predict the human dose and PK, allometry, PBPK (physiologically based pharmacokinetics) and exploratory IND (Investigational New Drug) can be used (meaning early exploratory clinical trial on few volunteers at low doses in order to get some pharmacokinetic information).
Allometry relates physiological functions for various species with their body weight. The problem is the poor correlation of bioavailability between animal and human data, but good results can be obtained with this method. It is a simple, robust and quantitative method. It works well for drugs with renal elimination and for drugs with similar elimination processes across species, but it fails when elimination pathways differ between species (drug metabolizing enzymes, efflux/uptake transporters). It only predicts average parameters. No mechanistic information is obtained and it does not tell anything about possible metabolites. It is also data and animal intensive.
PBPK really describes a profile, instead of giving only some parameters. One of the current strategies for PK predictions is based on a PBPK model in animals, which is

optimised with different parameters to get an optimal profile to go to humans. Modelling can also be used to simulate drug-drug interactions. PBPK modeling provides a full pharmacokinetic profile and mechanistic information on the pharmacokinetic behaviour of a potential drug candidate. A PBPK model allows to vary input parameters (sensitivity analysis).
Exploratory IND can be very appropriate for resolving pharmacokinetics issues. Human microdosing can be performed and radiotracers can be included, so that excretion pathways can be followed. This way it might be possible to get rid of a human mass balance study later on. There are limitations, the dissolution rates can be very different at oral therapeutic doses, the saturable first pass metabolism at therapeutic doses. But it is a potential approach to address uncertainty on human pharmacokinetics.
Regarding toxicokinetics the challenge is how to incorporate this program into the main group devoted to toxicologic evaluation without jeopardizing the toxicologic evaluation. The relatively extensive blood withdrawal in satellite rodents may bias the kinetic results. Some requirements exist to allow sampling from main groups. Blood volumes for bioanalysis need to be reduced, sparse sampling and a population approach should be considered.
To this end, dried blood spot analysis is used, the advantages from this technique are that they require reduced blood volumes and reduced costs. The biggest hurdle here is that the group exercising this practice has no benefit from it, but has to invest more time in it, so it is difficult to organise it within a company.
With sparse sampling/population approach, when using the population – toxicokinetics approach, less samples are needed during repeated dose toxicity studies. Reducing time points at which samples are taken to five per animal, instead of eight, still allows to reconstruct the model.
To conclude, empirical approaches will provide only pharmacokinetic parameters such as clearance, volume of distribution and half-life, whereas PBPK modeling simulates and predicts the time-course of drug concentrations. Therefore, to provide valuable information to the customer, mechanistic models should be the preferred choice. Poor simulation does not mean that the model is useless, one has to look to the non-fitting parts and use the model to identify what data/process is missing. The advances in dried blood spot analysis are promising to allow sampling from main group devoted to toxicological evaluation. This is even more powerful if used in combination with a population approach.

5.6.3.2.2.10. Regulatory requirements for non-clinical pharmaco- & toxicokinetic testing of human medicinal products, by Sonja Beken, FAGG-AFMPS, BE

Before regulators get involved, a lot of work is already been done, *in silico*, *in vitro* (such as permeability, metabolite identification, protein binding…) and some confirmatory *in vivo* experiments, during lead generation and optimisation. This already reduces the number of animals used and early drug attrition.
There are several regulatory guidelines, such as the “Pharmacokinetics and metabolic studies in the safety evaluation of new medicinal products in animals EudraLex vol. 3B

(NfG 3BS11A)". The objectives of this guideline are that PK and metabolic studies should assess the levels of the substance and of its metabolites and their kinetics in blood, body fluids and organs to obtain information on the relationship between target organ toxicity and the blood, body fluids and organ concentrations of the substance. They must also assess the possibility of enzyme induction and of accumulation of the substance with repeated administration. This should be done to choose where possible the animal species to be used in toxicological studies on the basis of their similarity to man in handling the medicinal product, and to determine the relevance of these toxicity studies to man. The animal species used should be those that are appropriate to predict human ADME. A preliminary study of kinetics and metabolism of the medicinal product in a few human subjects could provide useful information in choosing the animal species to be used in repeated dose toxicity studies. All this is done at therapeutic doses.

The outcome includes absorption, distribution, half-life, plasma protein binding… Few *in vitro* methods are used, but they are always add-ons for the *in vivo* experiments. *In vitro* methods are used for absorption, plasma protein binding, pattern of metabolites and enzyme induction.

According to the "Toxicokinetics: the assessment of systemic exposure in toxicity studies (ICH S3A)" guideline, the primary objective of toxicokinetics testing is to describe the systemic exposure in animals and its relationship to the dose levels and the time course of the toxicity study. This helps to set safety margins for human safety. Secondary objectives are to relate the exposure achieved in toxicity studies to toxicological findings (relevance of findings with respect to clinical safety) and the selection of animal species and findings at doses exceeding those in the clinical use (to optimise the design of subsequent non-clinical toxicity studies).

Data may be obtained from all animals on a toxicity study, in representative subgroups, in satellite groups or in separate studies. Normally, samples for the generation of toxicokinetic data may be collected from main study animals, where large animals are involved, but satellite groups may be required for the smaller (rodent) species. Toxicokinetic data are not necessarily required from studies of different duration if the dosing regimen is essentially unchanged, but this has to be assessed on a case-by-case basis.

The revised "Non-Clinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorisation for Pharmaceuticals (ICH M3)" guideline includes strategies to reduce the number of animals in accordance with the 3R principles. It includes new recommendations regarding PK, such as exploratory clinical trials (microdose trials) this may reduce the animal PK data package. Only TK (toxicokinetic) data are needed, for phase I and II studies. The starting dose is based on the NOAEL (no observable adverse effect level) in the most appropriate animal species, which is based on allometry. These measures allow for reduction rather than replacement.

PBPK modelling is useful in lead identification and optimisation, clinical candidate selection and early drug development before and after entry into man. However, there is a need for *in vivo* verification in non-clinical species before making human predictions. *In vivo* data in non-clinical species drive model refinement and validation. There is a lack of experience of PB/PK (physiologically based pharmacokinetic) modelling in regulatory

decision making, which has as a consequence that regulators do not really trust this approach.
In vitro models are used for early identification of metabolites, metabolic stability, enzyme induction/inhibition, transporter mechanisms… They are also used to identify species differences and select species. They can also predict human ADME. The "Note for guidance on the Investigation of Drug Interactions (CPMP/EWP/560/95)" is being revised and linked with the activities of the Validation Management Group of ECVAM on Validation of Toxicokinetics and Metabolism. There will be a more detailed part on the *in vitro* work that has to be done in this kind of studies. *In vitro* studies are used to support First In Human studies up to phase II in combination with toxicokinetics, but *in vivo* PK data are still required before including a large number of human subjects or treating them for long duration (phase III). So, this is also rather reduction than replacement.
As of today, no pure replacement approaches are implemented or foreseen for non-clinical PK & TK, from a regulatory point of view. *In silico, in vitro* and *in vivo* methodologies are used as complementary in a tiered approach, which leads to a reduction of animal use. PK modeling approaches are used for calculating First In Human dose levels using data-based kinetic models. PB/PK modeling approaches lead to reduction of animal use rather than replacement, but there is a need for increased knowledge and/confidence regarding PB/PK modeling for increased application in regulatory decision making.

Discussion

Maier: Initial human data can be used to choose the representative animal model, can it also be used to choose the correct *in vitro* assay?

Beken: This is possible, if human data are available it drives the further non-clinical studies, be it *in vitro* or *in vivo*.

5.6.3.2.3. Free communications

5.6.3.2.3.1. Replacement, by Jon Richmond, Home Office, UK

Better science and efficient regulation are needed to get better and safer medicinal products.
In 1959, Russell & Burch defined replacement as any scientific method employing non-sentient material which may in the history of animal experimentation replace methods which use conscious vertebrate animals. They distinguished between two types of models. High fidelity models very closely resemble, in biological and structural terms, the system of interest, e.g. when interesting in man, a chimpanzee is a high fidelity model. But they stressed that what is required to produce relevant findings are in fact

high discrimination models which correlate in terms of stimulus and response to very precise elements of what is observed in the target species.

Russell & Burch also distinguished two types of replacement: absolute and relative. In absolute replacements, no sentient animals are required, such as is the case in *in silico* systems. Relative replacement still uses animals, but in non-painful procedures, such as in *ex vivo* methods and in non-recovery procedures. However, the distinction is often blurred.

Replacement methods are not just substitutes for animal models, typically they are more advanced and relevant models that supersede animal test methods. As they better predict what will happen in man, they can not give precisely the same results as animal models. Ideally replacement test methods are developed from a detailed understanding of the relevant biological mechanisms. For regulatory purposes these methods have to be validated, which is expensive and time consuming, even though the established animal models were never scientifically validated. Therefore, the replacement methods generally represent better science. In the future, they should be the tools of choice to better understand the biological mechanisms and relevance of the results.

Alternative methods typically are more technically advanced, cost effective, reliable, scientifically validated, and more easily scalable than animal models. They can also overcome the limitations, increase the scope of animal-based methods and are relevant to research and testing.

There is a range of replacement options, for example strategies can be developed to avoid the need to generate new animal-based data – examples of this have emerged from ICH. Systems can be developed to allow elements of evidence gathering, analysis and decision making without new animal test data. Methods and models providing the scientific insights sought without causing pain, suffering, distress or lasting harm to sentient animals, can be developed.

Replacement strategies and systems can be used to amend or harmonise regulatory requirements, to ask whether certain tests are relevant and to avoid checklist approaches. Tiered and hierarchical approaches can be used. Scientific objectives can be changed, so that no animal tests are needed. Other possibilities are reviewing published work and data sharing. It is also possible to decide not to do a certain test anymore, as was done in some cases with the abnormal toxicity testing in biologics, which actually gave very little information.

Replacement methods and models can include the use of non-sentient organisms, including immature forms, *in vitro*, *ex vivo* and *in silico* systems, human studies. Education and training in new technologies help in getting knowledge about these models spread.

A question that has to be asked is whether certain methods really are replacements. They might in fact be adjuncts, tests done in addition rather than in place of animal use, or bottle-openers, which overcome bottlenecks in animal based R&D, but thus produce more lead compounds for further development, or even drivers, setting the scene for formulating additional scientific objectives that might require animal use.

Discussion

Garthoff: ICH, in pharmaceuticals, is ahead of other fields. Is this really so? This depends on the criterion.

Richmond: ICH delivers on two levels. Firstly, the harmonisation of test requirements at global level, which is important, because if for example a certain animal test does not have to be done in the EU, but has to be done in the USA, then the animal test is only relocated, but still done. Secondly, the ICH takes a more evidence and science based approach to regulatory decision making.

5.6.3.2.4. Conclusions of the day, by Gisbert Sponer

Intensive work has been done to develop new alternatives and to replace animal tests, as well as reduce them, where possible. Today, there is, however, very little chance that all animal experiments can be replaced by alternative methods in the near future. Most of the speakers indicated that alternative methods can provide better insight into mechanistic pathways but they are so far not able to provide the broad spectrum of information which is needed for benefit/ risk assessment of new drugs being in the developmental phase.

5.6.3.2.5. Session III: Replacement in the development of new biological entities

5.6.3.2.5.1. Keynote lecture "Vital benefits for humans by replacing animal experiments?", by Levente Pencz, Fauna Society, HU

There are three approaches to animal welfare, being basic health and functioning, natural living and affective state, which include the emotions and basic happiness of the animals. These three approaches overlap each other. Three examples for mice were given. Concerning health, when a mouse receives veterinary care, it is supposed to be healthy, but when it cannot play and is fearful, its natural living and affective state are not optimal. Pets should have an optimal affective state, but they can be sick, so in that case its health is not optimal. Wild mice have optimal natural living, but they can be afraid of predation, so in that case that their affective state is not optimal. In testing animals, a balance between these three approaches must be sought.
An ethical question is how it can be morally justified to cause distress, discomfort, suffering or pain to animals, limiting their freedom and killing them after ending the experiment. Answers exist from three different points of view: contractarian, utilitarian and animal rights.
In the contractarian, human centered, view, there is no objection to animal experiments and the potential benefits of animal-based research are harvested as much as possible.

The only animal welfare concern here is that experiments causing public concern, are avoided.
In the utilitarian, egalitarian, view, a cost – benefit analysis is made of the use of animals and a balance is sought. This has as a consequence the duty to improve the research involved. In this view, the use of the 3Rs approach is important. Animal rights groups deny the trade-off, which exists between the interests of man and those of the animals.
In the animal rights view, it is unacceptable to treat sentient beings as a tool to achieve one's goals. Animal rights must be respected and the dignity of the animals must be preserved. Testing bans emerge from this point of view.
However, there is room for compromise, when two requirements are taken into account, namely research should provide vital benefits for humans and animals should be taken care of as much as possible.
No vital benefits are received with the development of slightly different drugs (due to patent matters) or with diseases that not necessarily need to be cured, such as baldness. Also cosmetics can be classified here, and also medications for diseases which are caused by abnormal lifestyle, for example smoking. When taking into account these considerations not all animal studies would be viewed in the same way, i.e. there are studies saving human lives and there are those that make it only more comfortable.
Research should provide vital benefits and a way to achieve this is to put every effort to replace experiments which have no vital benefits. The question is also whether animal experiments are always the best way of achieving one's goals.
However, there are some encouraging developments, such as volunteer studies in pain research. The identification of areas to prioritise for replacement is also an important development. Failed molecules and withdrawn drugs should be made available. A multidisciplinariy approach helps to avoid unnecessary animal testing. A better classification of sub-groups can increase focus and power. Networks of local human tissue banks can also increase the possibilities of *in vitro* testing in human tissues. In some fields, such as toxicology, there are also opportunities for replacement. Use of the microdosing approach in human beings can also avoid unnecessary animal studies. A key factor in this process is the international regulatory harmonisation and requirements.
To this end, critical discussion within the scientific community is needed. Companies should be proactive in putting more effort to find replacing methods (for example by internal validation). Companies should also engage with other stakeholders (animal welfare representatives, policy-makers, academics) to find solutions. To conclude, animal welfare should be put beyond individual company interests.

Discussion

Rogiers: When talking about making available failed and withdrawn drugs, is it meant in the sense of making available results and to what use?

Pencz: The goal is to make negative results available, so that molecules or techniques that failed, will not be tested in other companies again, as often happens now. By avoiding double studies of failed drugs, reduction in animal studies can be reached as well.

Sponer: Can a clear edge between vital and non-vital benefits be made? In some cases, it is very clear, but in other cases, as with cancer, it can take years before it becomes clear whether a drug really improves the survival of the patient or it does not.

Pencz: In cases like diabetes and cancer, the benefits are vital. In many other cases, it is questionable, such as with baldness and other non-diseases. But it is indeed debatable what is vital and what is not.

van der Laan: Is microdosing replacement? It is mainly an early way to go into humans, but after that it is decided whether all tests should be done or not. So, it is rather a reduction of animal use in screening.

Pencz: This is a recent development and it has potential to work as replacement, but I'm not an expert in this field.

5.6.3.2.5.2. Regulators view on preclinical requirements prior the initiation of clinical studies, by Steffen Gross, Paul-Ehrlich Institute, DE

The Paul-Ehrlich Institute is responsible for processing of applications for marketing authorization and subsequent applications, approval of clinical trials for medicinal products for human use within the scope of the Paul-Ehrlich-Institute and official testing and release of batches of medicinal products.
This presentation focuses on monoclonal antibodies, sera, and immunoglobulins, etc...
There are different guidelines to follow in this field, being the European Pharmacopoeia, CHMP (Committee for Medicinal Products for Human Use) and ICH (International Conference on Harmonisation) guidelines. Of the ICH guidelines, especially the M3 and S6 guidelines are important. The ICH S9: non-clinical development for anti-cancer drugs: for small molecules and biotech products is for patients with a short life expectancy. There are also footnotes in these guidelines, which state that they may not apply for monoclonal antibodies.
According to the declaration of Helsinki, for ethical, scientific, and regulatory reasons and despite significant technological progress, current development is still based on studies in animals.
It is possible to reduce, and even replace, animal experiments, this starts with early basic research. *In silico* methods and literature searches, can help to look for structural alerts and to predict toxicity based on the mode of action. Then one can go to the first safety assessment, by toxicological profiling *in vitro*, as there are *in vitro* assays for cardiotoxicity, genotoxicity and general toxicity.
In basic research, the properties of the antigen have to be investigated, factors such as the physiological role, existence of knock-out models, solubility, exclusive expression of the antigen at the cell surface, the expression pattern in normal tissues and ubiquitous/specific expression.

Also the properties of the antibody have to be investigated. Questions to be asked here are, for example, whether the antibody is an specific monoclonal inhibitor, is there Fc-binding, cross-linking etc

Then it becomes important to find out whether an animal model is available or not. Sequence comparison can be carried out, downstream effects and signalling pathways can be searched for. Cross reactivity studies should be performed in order to find relevant animal species and to interpret kinetics or certain toxicity findings. If there is no cross reactivity, a surrogate model should be considered.

Sometimes *in vitro* studies can be more predictive than animal studies, as was shown after the Tegenero case, where substantial cytokine release could be shown with human materials using *in vitro* methods.

Safety evaluation programs should normally include two relevant species. However, in certain justified cases one relevant species may suffice, for example when only one relevant species can be identified or when the biological activity of the biopharmaceutical is well understood.

Other ways for reduction include safety pharmacology in toxicological studies, the design of repeat dose toxicity studies could allow the reduction or omission of single dose toxicity studies. No additional animal studies are requested in cases where no gain in information about the product might be expected. The design of the ani,mal studies should be carefully planned. Non-GLP (good laboratory practice) studies, or studies in non-relevant animal species should be avoided.

In a classical three-study design in reproductive and developmental toxicity, male fertility can be tested as part of general toxicity studies. In embryo/foetal development, there is no transplacental transfer during organogenesis, so these studies are not useful in the field of moAbs (monoclonal antibodies). Therefore, these studies are often taken together with those of pre/postnatal development. The need for reproductive/developmental toxicity studies is dependent upon the product, clinical indication and intended patient population.

Standard carcinogenicity studies are generally inappropriate for biotechnology-derived pharmaceuticals. However, product-specific assessment of carcinogenic potential may still be needed depending upon duration of clinical dosing, patient population or biological activity of the product. When there is a concern about carcinogenic potential, a variety of approaches may be considered to evaluate the risk.

When introducing proteins into a patient, most often there will be an immunogenic reaction. In some cases this might have no effect, but in most cases there will be an effect, such as neutralising of biological effects with the consequence of compromising further therapy, altered PK/PD (pharmacodynamics/pharmacokinetics), cross-reaction with native proteins and induction of adverse symptoms. However, thanks to the humanisation of the therapeutic genes, the immunogenicity has decreased. But with new constructs such as bispecific antibodies, diabodies etc, there might be a higher potential for immunogenicity again. Factors influencing immunogenicity of proteins include concomitant treatment, dose, route and frequency of administration and duration of therapy, but also immune status of the patient and factors of the proteins themselves, such as molecular structure and product impurities. These factors are difficult to predict, because these are complex molecules with a high molecular weight and known heterogeneity and there might be as well process- as product-related impurities. An

important paradigm is that the process defines the product. Small changes in the process, might lead to differences in the proteins and even to new products.
Pre-clinical tests have to be performed, if comparability tests are not sufficient. These should include PK/PD, clinical efficacy, specific safety, pharmacovigilance and immunogenicity studies. Comparison of immunogenicity can only be made in clinical trials. Finding the relevant species is important, because many biotechnological products are species-specific. Due to this, the predictability of non-clinical animal models is low. Nevertheless, immunogenicity evaluation should be included in non-clinical studies, because it helps to interpret toxicology findings and to optimise strategies to lower immunogenicity and comparability exercises and to study the consequence of cross-reactivity of immune responses against endogenous proteins.
The Paul-Ehrlich Institute also performs batch release and control. For example, with anti-tetanus immunoglobulines, 48 animals are used per batch, last year 147 batches were tested, this amounts to more than 7.000 animals. A project to replace this with *in vitro* tests is now going on and the first results seem promising. Implementation of a commonly accepted *in vitro* alternative method for human tetanus immunoglobulin batch release purposes will ensure better implementation of the 3R policy.

Discussion

Olejniczak: About the S9 guideline: non-clinical development for anti-cancer drugs: for small molecules and biotech products, this guideline is for patients with a short life expectancy.

Gross: The guideline should be carefully read, as well as the footnote, as there is some room for interpretation.

van der Laan: In the presentation it was said that there is a need for carcinogenicity testing, growth factors and immunosuppressants, but might it be a possibility to accept a certain risk, to bring this in the risk management plan and no further data are needed?

Gross: The Paul-Ehrlich Institute would accept it, but there should be a justification and discussion.

Garthoff: In the TeGenero case, afterwards *in vitro* tests were done and certain indications were found, because they knew what they had to look for. This information is not available with new biologicals. How do you deal with this?

Gross: The company did not want to do *in vitro* tests before, for some reasons, otherwise they could have picked up the indications. The cynomolgus monkey also was not as predictive as expected. It is not yet possible to go directly from *in vitro* to clinical studies.

5.6.3.2.5.3. Replacement in the development of new biological entities, by Bernd Mueller-Beckmann, Hoffmann – La Roche, DE

Biologics are mainly used in serious disorders and represent about 20% of the market of pharmaceuticals. This presentation exemplifies on oncology. Since cancer is the number one cause of death worldwide, there is a need for better medicines. New cancer medicines have shown to improve prognosis for the patients, so the median survival time in cancer increases and biologics play the most important role in this medical progress.
Small molecules have a low molecular weight, they have a chemically defined molecular structure. Biologics, in particular antibodies derived from biological sources, have a high molecular weight and they are mostly a mixture of molecules. They also have a high target specificity, and thus a high species specificity, which limits the choices of relevant test species mostly to non-human primates. Biologics have long half-life times. In small molecules, toxicity of metabolites also must be considered, whereas biologics are degraded proteolytic, so there is no toxicity induced by metabolites.
Preclinical efficacy testing is not formally regulated. It addresses molecular, subcellular, cellular and *in vivo* models of disease pathways. Due to the high specificity, relevant disease models are rarely available. Due to this, a project specific scientific approach is needed. The reasons to do these tests are to support scientific rationale and confidence in a hypothesis that can be translated into the clinic and to demonstrate target interaction with expected effect that may translate into efficacy, to motivate clinicians and patients to enroll in clinical trials with a reasonable chance that patients will benefit. It also provides legitimization for ethical committees to give their consent to administer a new drug candidate to humans. Preclinical efficacy testing provides experimental proof of concept, determines effective doses and concentrations, but often suffers from weak predictability.
Replacement of efficacy studies may be disputed, but meaningful efficacy studies can open new avenues for the benefit of patients.
For example, in a xenograft study in mice bearing breast cancer, animals treated with pertuzumab or trastuzumab continued to develop a tumour, but the combination of both antibodies cured 6 out of 10 animals, without return of the tumour. This combination was also tested in a clinical setting and here about 50% of the patients reacted positively on the combination. The same combination was also tested in mice with lung cancer and here 3 out of 10 animals were cured, without return of the tumour.
Preclinical safety testing is regulated by guidelines from regulatory bodies. These are mandatory to enable clinical trials. They are done in *in vivo* models, in two species, rodent and non-rodent. This is done to identify an initial safe dose for subsequent dose escalation schemes, to identify potential target organs for toxicity, for studying reversibility, to identify safety parameters for clinical monitoring and to define pharmacological and toxicological effects prior to and throughout clinical development.
The relevance of an animal model (ICH S6) may be demonstrated by comparison of sequence homology of epitope or receptor, target binding properties or biological effects including downstream effects. Data from non-relevant species are not required. If no relevant animal model is available, alternatives might be considered, such as transgenic animals.

Selection of the relevant toxicology species for safety testing in small molecules is guided by pharmacology and metabolism. For biologics, homology of target biology, target binding and target distribution, tissue cross-reactivity and pharmacological activity (functional and biological relevance) are the selection criteria. Biologics feature high target specificity and therefore rarely cause unspecific toxicity, but the high species specificity mostly causes that rodents or non-rodents are no relevant species. Often only non-human primates are the only relevant species, but chimpanzees cannot be used for toxicology studies. If tissue cross-reactivity is done to identify the relevant species, and only the chimpanzee turns out to be relevant, then what has to be done? Skip the *in vivo* testing?
In another example, a pharmacokinetic and safety study with monoclonal antibodies was performed using cynomolgus monkeys. Here unexpected observations demonstrated the relevance of *in vivo* studies, since there was a discrepancy between the expected and observed pharmacokinetic profile and an unexpected fatality of the highest dose was observed, which would not have been detected without in vivo animal testing. *In vivo* efficacy studies may be partially replaced by *in vitro* experiments, but the legitimate interests of involved parties must be considered, namely the physicians, the patients, the ethical committees and the sponsors. *In vivo* safety studies are indispensable for proven scientific, ethical and regulatory reasons, they cannot be replaced.
Revision of ICH S6 guideline is underway, five topics are selected for update via an addendum to ICH S6: species selection, study design, reproductive/developmental toxicity, carcinogenicity and immunogenicity. An update of ICH S6 (R1) may offer opportunities for reduction of animal studies, but not for replacement.

Discussion

Rogiers: Toxicology of biopharmaceuticals is often overlooked in discussions. The difference between the toxicology of small molecules and biopharmaceuticals is often not made. This should be brought up more often.

Mueller-Beckmann: Usually, studies with biologicals are conducted in non-human primates, but these are not necessarily the default species. The use of non-human primates must be justified for each case based on cross-reactivity studies and functional data.

Garthoff: Accessibility to negative results is an issue, and the use of primates is a very sensitive issue, for these reasons it would be good to bring the reasons to the public why this research is done.

van der Laan: One of the dilemma's for compounds only specific in chimpanzees is the guideline on risk mitigation: what is the certainty of the safety of the medication? If there is a low certainty and no animal data are available, then what other options would be available beside direct testing in humans? Some companies then decided to discontinue

biologicals only specific for chimpanzees, because they feel that they cannot go into humans without animal data.

5.6.3.2.5.4. Replacement in the development of new biological entities, by Beatriz Lima, iMED.UL & INFARMED, PT

Animal experiments are needed, but they can sometimes be replaced by other methods without compromising scientifically based assessment of efficacy, safety and quality.
To be able to progress, it is important to identify the bottlenecks, concerning the implementation of replacements, such as scientific, ethical reasons and global requirements. It also is important to identify fields where there are realistic chances for replacement.
Most commonly with peptides, proteins, oligonucleotides, metabolism is usually not a concern, there is more concern about the mode of action. Since they are human specific molecules, there often is a decreased animal PD (pharmacodynamics) responsiveness, toxicity is mostly related to exaggerated PD and the molecules are often immunogenic in animals. Due to this, there is a decreased number of relevant species for safety/efficacy assessment. Some companies are now focusing on less human specific biologics, so that they can easier be tested on animals.
The main requirements for the development of biopharmaceuticals are to check for relevant species for the human system, with respect to responsiveness (pharmacodynamics) and it has to be qualitatively defined. In case no relevant species exist, homologous or transgenic models must be looked for. As always, irrelevant animal experimentation must be avoided.
In vitro and *in silico* approaches can be used in PD investigations focussing on drug-target interaction, understanding of response cascades and cross talks. *In vitro* approaches also can give information about quantitative estimation of concentrations needed at target level. In PK (pharmacokinetic studies), *in vitro* methods can give information to evaluate target distribution and tissue binding and to predict drug biodistribution patterns.
In safety evaluation, *in vitro* and *in silico* methods can be used to identify secondary targets for potential secondary or adverse effects, to anticipate safety windows based on effects on the target and efficacious concentration in the target. Predictions of the consequences of excessive PD response, such as cytokine release, can also be made with these methods.
The bottlenecks for replacement are due to insufficient scientific knowledge to justify replacement: tests need to be as predictive as *in vivo* models. From an ethical and scientific standpoint, extensive comparative studies are needed to prove that the new methods work. Science drives regulatory acceptance, but due to this, it always takes long before new methods are actually put to use. Local regulatory acceptance can trigger globalisation, but since most companies are global, new methods will only be used when they are accepted worldwide.
For biopharmaceuticals with a new mode of action, unless irrelevant for humans, animal models cannot be discarded in the present state of the art. Whole body physiology and physiopathology is different from the sum of individual interactions and responses at all

parts. Well designed studies in relevant animal models have shown predictability of human responses. New modes of action cannot be understood from local activities only. Whole system testing is needed. For these reasons, animal tests cannot be discarded unless a direct human test is performed, which can be done in some exceptional cases or could be done by micro-dosing.
With known mode of action, provided that the primary and secondary targets are identified, the drug interaction is fully characterised qualitatively and quantitatively, *in vitro* experimentation did not provide unknown concerns and further *in vivo* data will not impact on benefit/risk evaluation, animal experimentation might be avoided: comparative evaluation of existing products could help the discussion.
Some questions about *in vivo* testing can also be raised: is there a need for repetition of reproductive toxicity of drugs from the same well understood pharmacological class? Is there a need to do a two year carcinogenic study for growth factors, hormones expected to be tumourigenic due to their mode of action? Better use of, for example, -omics and *in vitro* systems, might overcome further mechanistic *in vivo* studies.
Biocomparability is a big discussion in biologics. A lot of similar biologics are produced nowadays and animal studies are used to detect differences in pharmacology and toxicology. It could be questioned whether this is useful, since the products are so similar. For this goal, *in vitro* characterization might be sufficient. Here might be good opportunities for replacement.
One way to overcome hurdles in development and usage of replacement methods is validation. Appropriate comparative evaluation of *in vivo* with alternative models for different classes is an important way to go forward. A realistic approach is also important: the areas where alternative methods can be developed. Another possibility is to also first simplify certain tests, before they can be replaced by other methods.
There are realistic chances for the replacement of certain methods, but complete abolishment of animal studies is impossible at present.

Discussion

Garthoff: Since biologics are coming up, it is thought that there will be an increase in animal experimentation and probably also in the use of primates. Will this still increase?

Lima: This might happen, indeed. For this however, the evolution of making smaller, less specific biologics is very interesting, since this can overcome the increase in the use of primates. The use of transgenic animals might also help, although the problem is that their physiology is not always understood very well.

5.6.3.2.6. Session IV: Replacement in quality control

5.6.3.2.6.1. In vitro assays as an alternative to animal testing in quality control, by Jean-Michel Chapsal, Sanofi Pasteur, FR

Batch control testing of vaccines amounts to about 10% of all animal use in the pharmaceutical industry, using 10 million animals worldwide every year.
In the vaccine industry, animal tests are used in vaccine development, a small amount in production and mostly in batch control testing, in particular, safety and potency testing. Routine batch control testing is responsible for 80% of animal use. In Sanofi Pasteur, 98% of the animals used are rodents and chicken, 95% is used for quality control and 5% in research. Primates are used only at about 0.1 to 0.2% of animal, they are used only for some controls. In official control laboratories animals are also used, which seems to be double testing.
Since 2000, the use of animals has diminished by about 40%., although the vaccine sales had a substantial growth.
There are 3 major reasons to implement alternative methods to animal assays. There is an ethical reason to replace animal experimentation. New methods can help to reduce the assay variability, which is important for the quality of the product. Replacement of animal tests can also reduce the costs of quality control. However, replacing tests will be more difficult than reduction.
Replacement alternatives include *in vitro* artificial systems, lower organisms, such as bacteria, fungi and plants and *in silico* models. Biophysical and biochemical techniques, for example antigen quantification, toxoid residual toxicity and irreversibility (use of binding and enzymatic activities) and characterization of antigens can also replace animal tests. Certain tests that are no longer needed due to increased production consistency, such as, the abnormal toxicity can be deleted.
In replacing animal models in vaccines batch control testing, safety is a short term target as most of these tests are directed to known mechanisms and manufacturing consistency, for known products. Primary monkey cells are used for effective inactivation of polio vaccine with bioengineered continuous cell lines. MAPREC (Mutant Analysis by Polymerase chain reaction (PCR) and Restriction-Enzyme Cleavage) is a method for detecting change in the sequence of oral polio vaccine. These *in vitro* methods are used for viral vaccines. For bacterial vaccines, the residual irreversibility of toxoid has to be shown. The *in vivo* test for residual/irreversibility of tetanus toxoid is replaced by *in vitro* enzymatic and/or binding tests, the inactivation *in vivo* test for residual/irreversibility of diphtheria toxoid is replaced by *in vitro* VERO (African green monkey kidney epithelial cell line) cell tests.
For general safety tests, the abnormal toxicity test can be removed, although some countries still request for it. Endotoxins-Pyrogens tests can be replaced by *in vitro* assays with human monocytoid cell lines or cryopreserved human whole blood when the validation of these methods have been demonstrated. One of the limitations is a lack of scientific knowledge on some of the assays.
Replacement of potency tests in vaccine batch control testing is a long target, because it is a biological induced effect, with complex antigens. Some *in vitro* models are available.

For immune-sera, the *in vivo* seroneutralisation test is replaced by an ELISA (Enzyme Linked ImmunoSorbent Assay) test.
With respect to antigen quantification of vaccines. Some tests are performed by ELISA-technique when a well known correlation with clinical data exists. Antigen characterization can be done by biochemical-biophysical assays. Artificial immune systems are *in vitro* biomimetic of the human immune response. The limitations here are a lack of scientific knowledge on the antigens, for example in pertussis, the toxoiding impact on conformational epitopes involved in the potency of the vaccines and in general immunological mechanisms are not well understood.
Potency testing, has however seen a shift towards reduction and replacement in the last decade. Firstly, the *in vivo* challenge potency assay was carried out with a multi-dose assay, later with a single dose assay. Furthermore, *in vivo* serology is now analysed by ELISA, ToBI (Toxin Binding Inhibition), VERO cells and multiplex methods. For the future, there are two possible *in vitro* systems, characterization tests specific for each antigen and artificial immune systems.
With an increasing knowledge of the vaccine manufacturing process, there is also a shift from *in vivo* evaluation of protective immune response to an *in vitro* evaluation of antigenicity.
In general, the limitations to replace animal tests are that it is mandatory to validate the new methods by comparing the data generated by existing assays with those obtained with new *in vitro* tests. The problem is that even *in vivo* data are quite variable and that sometimes the gold standard mechanisms are unknown. From a companies point of view, there is also the risk that considerable effort is put into a proposal that does not get accepted by regulators. It could also be favorable if regulators would be more strict on implementation delay in routine testing. Implementing assays that may be difficult to maintain with respect to reagents, standards, is another limitation. Cost/benefits balances are another point, because going to replacement needs investments from companies.
In safety tests, replacement could probably be achieved. In potency tests, *in vitro* models will not replace animal models, because the impact of main manufacturing changes have to be monitored. But *in vitro* models will replace animal models, e.g. more data regarding consistency tools offer the chance to reduce the use of animal models for quality control, such data are being brought up for new vaccines, however, clinical data must be linked to biochemical-biophysical data to avoid animal use.
The successful implementation of replacement for regulatory testing depends on high quality science and on understanding, recognition and implementation of the change by all stakeholders. Global harmonisation is also needed. There are too much organisations, so that it is not always clear who to go to with certain problems or questions, all stakeholders must be involved, which demands discussion and there is a need for European funding for projects for biologicals, since most EU projects are in toxicology.

Discussion

Sponer: Is abnormal toxicity still in pharmacopoeia or is it totally banned?

Chapsal: It is still mentioned in the monographs, but for nearly all vaccines, these tests are removed. But outside the EU, abnormal toxicity testing is often still required.

Buchheit: A lot of research on 3Rs is done in companies, but it would be better if this research is brought more to a global or European basis. With the involvement of EDQM (European Directorate for the Quality of Medicines and HealthCare), OMCLs (Official Medicines Control Laboratories) and other partners, this could be organised, so that at the end alternative methods would be available not only for one manufacturer but for the whole community.

Chapsal: There are discussions, with EPAA, to establish a European vaccines group.

Lima: The need for toxicology testing for different manufacturing sites was mentioned. Is this done to identify quality problems?

Chapsal: Raw material origin can be an issue in terms of toxicity, since it is not always according to standards. The problem is that it is not always known when the material comes from an other manufacturer.

Garthoff: Random testing might be a solution.

Chapsal: Random testing of the products in general would be good. If something has slight derivations, there might be a change in the final product, thus, for this reason random testing would indeed be good.

Gross: A discrimination between batch release and marketing authorization application must be made. If the marketing authorization application is changed, a comparability approach must be performed, in which the product must be characterized again. In batch release, if the method is validated, then it is possible to replace a method with an *in vitro* assay.

5.6.3.2.6.2. Current animal tests in the European Pharmacopoeia (Ph Eur): Perspectives for replacement, by Karl-Heinz Buchheit, EDQM, Council of Europe, DE

In individual monographs in the European Pharmacopoeia (Ph. Eur.) animal tests are currently listed during development, validation of production (during development and after major production changes) and for final bulk/final product (each batch); animal tests are used for safety and efficacy tests. The majority of tests is used for biologicals as compared to classical chemical medicines.
Tests for pyrogens are applied for parenteral preparations and blood derived products on the final product. These can be replaced by the endotoxin test, which is preferred, however, this only detects pyrogenic endotoxins, so one must be certain that there are no non-endotoxin pyrogens in the product. The pyrogen tests are also used for some

vaccines, also on the final product, and some old antibiotics, on the final lot. For most antibiotics there is however an endotoxin test in the Ph. Eur..

The test for abnormal toxicity is still in the general monograph for allergens, but is to be used only for allergens derived from moulds and which are for parenteral use. This test is done on the final product, but as of January 2010, this test will be moved to the production section of the Ph. Eur. Then this test has to be done only once during validation of production or when the production changes markedly. This is already the case for some vaccines, some antibiotics, aprotonin and for botulinum toxin. Here, the Ph. Eur. states: “The production method is validated to demonstrate that the product, if tested, would comply with the test for abnormal toxicity”.

Specific toxicity is a major problem for vaccines. Here, animal tests are used in diphtheria and tetanus vaccines, in the production section. The Ph. Eur. states: “The production method is validated to demonstrate that the product, if tested, would comply with the test.” The acellular pertussis vaccine monograph requests a test for specific toxicity in the production section, thus not for each batch. The same monograph requests a test for pertussis toxin on the final product, but it may be omitted if the results on the final bulk are satisfactory.

Animal tests used for potency assay are mentioned in the Ph. Eur. monographs for diphtheria, tetanus and hepatitis A/B vaccines, on the final product. In the latter case, *in vitro* tests are already widely used. For botulinum toxin, animal tests for potency are also done on the final product, however, the Ph. Eur. allows also an *in vitro* test after validation. For certain hormones, an animal potency test on the final product is mentioned by the Ph. Eur. However, this does not have to be done if an *in vitro* test is developed.

The Ph. Eur. is based on the “European Convention for the protection of vertebrate animals used for experimental and other scientific purposes” (Council of Europe, 1986). The Ph. Eur. is committed to the reduction of animal usage wherever possible and encourages to seek for alternative methods. In the general statements, it is written that the tests and assays described are the official methods. With the agreement of the competent authority, however, alternative methods of analysis may be used for control purposes. This also applies to animal tests.

In the general monograph on human vaccines, production section, it is stated that “Where justified and authorized, certain tests may be omitted where it can be demonstrated, that the production process consistently ensures compliance with the test.” If a persistent production process can be shown, not all tests need to be done.

Many animal tests can be replaced for individual products after proper validation against the Ph. Eur. method and with agreement of the competent authorities.

The Ph. Eur. is continuously reviewing all texts with the goal to replace animal tests. Replacement can be reached in the Ph. Eur. by introduction of an *in vitro* test, by moving a test “upstream” from the final product section to the production section or by the deletion of a test.

Examples for replacements include the use of a bacterial endotoxin test for the pyrogen test. The abnormal toxicity test, has been moved upstream in the case of aprotonin or has been deleted, as was the case with heparin. For extraneous agents testing, cell culture assays can be used instead of animal tests.

Furthermore, the Biological Standardisation Programme of the EDQM has as one goal the validation of alternative methods and has successfully introduced over the recent years a number of *in vitro* assay in the Ph. Eur. for example, the *in vivo* potency assay was replaced by a physico-chemical test for somatropin and an *in vitro* ELISA test for tetanus immunoglobulin is under validation to replace the challenge test in mice.
For veterinary vaccines, more attention is needed in the future, there are not much replacements in this field yet.
In 2007, EDQM did a survey to see whether the alternatives are applied. Six internationally active manufacturers still apply the abnormal toxicity test for vaccines for human use, although it is deleted in the Ph. Eur. This is done because this test is still required by WHO (World Health Organisation), FDA (Food and Drug Administration), India, China… In asking the manufacturers why alternatives were not often used, the reasons mentioned were too high costs, unclear validation (although clear guidelines are provided), personnel has to be trained again, equipment has to change, too time-consuming and the legislation is retarded for implementing these tests.
The major obstacles to implement alternatives are that the validation process is long and costly, there are no incentives, there is a need for variation of the marketing authorisation and there is a lack of global harmonisation.
Recommendations are to focus on veterinary vaccines in the future and authorities should encourage replacements. A production consistency approach should be fostered. International partners should be involved in the development of alternative as early as possible.

Discussion

Sponer: It seems that in this field there really is place for replacement, whereas this seemed limited in the other fields so far presented.

Buchheit: One way of replacement is to do collaborative studies, as a kind of generic validation, and subsequent inclusion of the test in the Ph. Eur. But in some cases it is impossible to show a good correlation between the *in vivo* and *in vitro* tests. In such cases, each manufacturer can still develop and validate an *in vitro* test for his own product. This is a second way to reach the goal.

Rowan: For veterinary vaccines, incentives are needed, because the values of sales are often small. Harmonisation is important, because a lot of tests are still performed, although not needed in Europe anymore, because other countries and organisations still require it. VICH (International Cooperation on Harmonisation of Technical Requirements for Registration of Veterinary Medicinal Products) is working on harmonisation for veterinary vaccines now.

Buchheit: On the other hand, in some cases VICH proposes to use more animals than the Ph. Eur.

5.6.3.2.7. Free communications

5.6.3.2.7.1. In vitro – in vivo: alternative or complementary approaches?, by István Gyertyán, Gedeon Richter, HU

If a pharmaceutical drug is under development for curing a disease, a clinical indication must be defined, the mechanism of action, i.e. the drug target must be defined and the chemical structure of the molecule needs to be known. There is a hypothesis, that with a certain mechanism of action, some positive effects will be induced in the respective disease. The final proof of concept can only be recognised in clinic trials, but, of course, every drug developer wants to have information about efficacy of the new drug as early as possible.
A candidate for clinical trials should be effective, act on the specified target (*in vitro*) and show efficacy in the (model of the) target disease (*in vivo*); and should be safe.
To reach this goal a hierarchic screening cascade is used, which starts with *in vitro* assays, has a mix of *in vitro* and *in vivo* assays in the middle and ends with *in vivo* assays. This cascade is used to test for efficacy, safety and appropriate ADME (absorption, distribution, metabolism and excretion) properties. In each of these steps, *in vitro* tests can be used.
In screening for the mode of action, mainly receptor binding studies (with cloned human proteins) and studies for effects in cellular systems (in transfected human proteins) are performed. This can be done with high throughput screening methods. The goal of these methods is to build up structure-activity relationships on the drug-target interaction (affinity, selectivity, functional activity). The advantage is that these methods are relatively simple systems, can be used for human targets and have a high capacity. But, they are not genuine alternative methods because there are no animal alternatives.
In studies for ADME parameters there is another approach. In particular, there are *in vivo* measurements, which can start with pharmacokinetics, measuring tissue levels, whole body autoradiography, metabolite identification and mass balance of the product and its metabolites. Studying what happens with the compound in the living organism, cannot be replaced by *in vitro* methods. However, there are indirect methods which make a reduction or even replacement possible to a certain degree.
The indirect approach uses *in vitro* measurements. The physico-chemical parameters largely determine the passage through biological membranes in different tissues. There are intestinal absorption models (CaCo-2 cell monolayer), blood-brain barrier penetration models, metabolism in liver microsomes or hepatocytes, effect on Cytochrom P- enzymes (drug-drug interaction liability) and plasma protein binding tests. The advantages are that human targets can be studied and their high capacity. The disadvantage is that the complex drug-organism interaction is decomposed to its elements. The goal is to give predictions for the *in vivo* properties, decreasing risk and cost of ADME parameter studies.

In vitro methods serve as important tool for lead optimisation. They are not alternatives to the *in vivo* methods, rather have a complementary, pre-screening role. The *in vitro* to *in vivo* transition success rate may range from 10% to 90%.
Investigations for efficacy in the proposed disease is carried out with animal models, although there are some problems such as species differences, differences in pathomechanism (e.g. genetic, drug-induced). These models also only partially mimic the complex human disease. However, these models are continuously improved, concerning predictive and construct validity. Translational science and importance of biomarkers may help to choose the best models. This is a fully *in vivo* domain, which cannot be replaced.
In safety testing, *in vitro* methods are used for mutagenicity and pyrogenicity. They can also be used in skin corrosivity and irritation, phototoxicity and percutaneous absorption. The advantages are animal welfare, but also the high speed with which they can be done and the reduction of costs of safety testing. In this field, there is place for real alternatives, so this is a challenge for the future.
Chances of replacement in drug development, lie in the development of technology and a deepening of knowledge of mechanisms of diseases. The biggest limitation is that a highly organised living system cannot be modelled by the aggregation of its components.

Discussion

Rogiers: Are there possibilities for integrated testing strategies?

Gyertyán: In ADME, many things can be measured which are quite predictive, but the predictability of *in vitro* tests is not always certain. Plasma protein binding can, for example, protect certain compounds from decomposing, which sometimes cannot be identified in *in vitro* tests.

5.6.3.2.7.2. Replacement: Laying the groundwork for change, by Katy Taylor, European Coalition to End Animal Experiments, UK

While we wait for replacement methods to become available, there are some problems for the implementation of those that are already available. Unsuccessful implementation of alternatives can be caused not only by scientific obstacles, but also by discordance between regulatory guidelines and practice exercised in the industry.
Replacements or removal of redundant tests can fall through the gaps, due to a delay in harmonisation, lack of clarity in guidance and confusion about the responsibility for the final steps in implementation.
Redundancy is here used as an example. While international harmonisation is going on, there are regions who are reluctant to clarify or update their own guidance in the meantime. However, harmonisation can take years. In the meantime, there is the European Directive 86/609/EEC, which prohibits experiments on animals if another scientifically satisfactory method is available. If the EU has already decided that a certain

test does not need to be done anymore, then it should not be done anymore, EU should then not wait for the harmonisation.
An example of this is single dose studies for pharmaceuticals. The use of dose escalation studies as a 'replacement' for these has been allowed at ICH level since 1997, but presumably this was not taken up due to regional disharmony and lack of clarity. The EU pursued greater clarity and harmonisation at ICH (M3(R2), 2009). However, the EU 1987 guideline demanding single dose tests remains on the EMEA website, which leads to confusion amongst both national regulators and industry about whether they are required or not.
Replacement methods may be incorporated into guidelines and monographs in the Ph. Eur. or the need to perform specific tests may be removed from each monograph, but for many reasons the animal test is still conducted. For example, many companies still prefer the pyrogen test above the bacterial endotoxin test and still perform abnormal toxicity testing. The question is whether test choice is always based on science or whether conservatism plays a role.
Concerning the responsibility for the final steps, in some cases, companies may need to demonstrate in-house validity of the replacement test or redundancy of the original animal test, as is the case with abnormal toxicity, botulinum testing and batch testing. But it appears that there is a need for encouragement for the companies to do this. If a good consistency in production can be shown, these tests can be waived, but industry does not seem to take these waivers up. The question here is whether the alternative tests must be enforced or whether the industry must be encouraged. Maybe EMEA can waive the fees that companies have to provide if they want to change their authorization files. There also often seems to be a lack of information on current practice so that regulators are not actually aware of whether alternatives are being used or not.
While there are international harmonisation delays, it is necessary to get a clear message about possible alternatives out at the regional level. There is a lack of clarity concerning guidelines, a greater clarity is needed when animal tests are really needed or not. Not using animals must also be promoted. In REACH, the message is already given that animal testing must be the last resort. It must also become clearer on who has the responsibility for the final steps. The current practice must be monitored, to make clear where there are problems. Investment in faster updates to guidance is also needed. The question is also, whether certain changes must be enforced or whether industry must be encouraged.
3Rs statements by industry or regulators can encourage the recognition of the importance of the 3Rs, both by users, and by the public. The requirements for validation can be clarified and promoted, especially concerning the importance of reference drugs, applicability domain and labelling/regulatory purposes. 3Rs statements can be included in guidelines, flexibility for alternatives can be included as well as a general encouragement not to use animals. Monitoring of actual practice is important, because it is unclear who is doing this. To conclude, the implementation phase is a very important phase that must not be overlooked when alternative methods become available.

Discussion

Buchheit: The Ph. Eur. attempts to harmonise requirements for medicines on the European level, currently for 37 member states. The disadvantage of this approach is that products and production processes from less well-developed countries also have to be taken into account. For this reason, different methods are given, for example for pyrogens some manufacturers have to use the pyrogen test due to their production process, while others can use the bacterial endotoxin test.
Another point, sometimes it is difficult for the manufacturers to know what they have to do to implement an alternative method. For this reason, the Ph. Eur. started to give guidelines for all new techniques to be implemented.

5.6.3.2.8. Round Table Discussion

Moderator: Bernward Garthoff
Panel: Jan Willem van der Laan, Lajos Balogh, Philippe Vanparys, Levente Pencz, Karl-Heinz Buccheit

Garthoff: Some topics that might deserve some more discussion are for example the implementation of new methods. Toxicity in the 21st century is a vision of the USA, so they should implement first.
Often the same items are mentioned and everybody in meetings seems to be aware of those, but are they also known outside, by the public?
Veterinary drugs are often forgotten as a result of human pharmaceutical development.
Science drives regulatory acceptance, but who, in science, can make messages clear in a non-scientific way? And where is the forum to do this?
When looking for ethical drugs, often solutions for less prominent issues will be found. One drug that was developed for a certain cardiovascular disease was a side-product of another product, which was used for baldness.

van der Laan: The statement that if there is no global harmonisation, then there should at least be a regional position is true, but the problem is that products are made for the global market, so some studies are only done for other markets and not for the EU. As long as there is no global harmonisation, regional harmonisation does not help a lot.

Vanparys: It might be good to have a taskforce at ICH, existing of people from industry and regulators from the three regions, involved in looking to possible alternatives. Then, if somebody wants to develop an alternative they could inform with ICH whether it is worth to go for it. ICH could then also follow the development and validation of the new method. When ICH covers this, the new methods might be implemented sooner. All regulators must then be involved from the beginning on.

van der Laan: This is not correct, since ICCVAM (US) (International Coordinating Committee on the Validation of Alternative Methods) had a discussion on the local lymph node assay. ECVAM (Europe) took over the ICCVAM position, so it is possible to implement methods without involvement from the ICH.

Richmond: The three –VAMs and Health Canada have signed a concordat which should both speed up and result international acceptance of validation study findings. ECVAM's internal and scientific advisory committee structures are being restructured. This should improve prioritization and the management of validation studies. Inviting early input from regulators should result in the development and validation of methods which the regulators will want to implement as fast as possible.

Vanparys: Due to the involvement of both ICCVAM and ECVAM the process is going even slower than in the past.

Buchheit: In the field of vaccines, WHO has to be taken in account as a major player, because vaccination in the third world follows the guidelines of WHO, so they will not necessarily accept decisions by ICH. Every country involved in WHO can have a veto, so if countries like India, Korea etc. want to have certain animal tests, then so it is.

Garthoff: Veterinary drugs seem to be forgotten.

Balogh: Veterinarians are doing a lot of research on spontaneously occurring diseases in man, here the animals can then also be a model for the human situation. For example in cancer there often are good similarities between the animal/canine and human situation, the same applies for endocrine diseases such as Cushing, thyroid diseases and others. In these cases it would be interesting to involve veterinarians more in the process of development.

Sponer: Heart failure is also a good example, as dogs can also suffer from it. The problem is that in the daily veterinary practice the dogs cannot be used as test animals, since there is no adequate control group. Thus, it is not easy to get valid data

Balogh: It would be difficult, veterinarians would have to be very well equipped and trained. Voluntary owners, who would prescribe for these test, would be needed. People with dogs with a final stage cardiac failure might want to be willing to let their dog be used for tests. Of course, criteria should be made to select the animals.

Garthoff: Is it possible to transfer *in vitro* assays across sectors? Take for example the EST.

Vanparys: The EST has been developed and validated with a certain set of compounds and it was so promising that a lot of companies implemented it, but when they started testing it with pharmaceuticals, the output was not as expected. Some companies still use it. When *in vivo* teratogenicity studies are done and there are teratogenicity problems, the

EST is used and if the compound is positive in the EST, then this can be used as a tool to select analogues. This way it is partial replacement, because only non-teratogenic drugs are tested in the clinic, so no animals are wasted for compounds that would not make it. Some companies also use it for routine testing.

Garthoff: Are there possibilities to make available data of failed drugs, to share negative results?

Pencz: It could provide additional information to the development process of other drugs. To reach such information from the companies it would be needed to break their confidentiality on this issue.

Garthoff: The most mentioned problem here is that it is difficult to share results and at the same time keep competitivity.

van der Laan: Data sharing is important, not only for background data of control groups, but also to have ideas on what went wrong. This has also been mentioned at IMI (Innovative Medicines Initiative). There are indeed difficulties with confidentiality, but for the commercial interest, the clinical data are the most important. All data of products that are stopped, are just sitting in the archives of the company and those should be made available.

Garthoff: Will a set of biomarkers ever be available for replacing toxicological testing?

van der Laan: Probably not, because a set of biomarkers for humans should be very large, so that aspects of local damage probably could not have a read out. Animal studies can be integrated more and more, together with *in vitro* studies for certain decisions. *In vitro* approaches have a place in the development of compounds, but not at the final stage, there *in vivo* confirmation is needed.

Vanparys: With new techniques, -omics, *in vitro*, *in vivo* and clinical studies and *in silico* tools should be used combined and if some parameters pop up in all these tests, this might be a biomarker. Industry could put up a group of scientists who can validate the biomarker issue. If this is validated, extrapolations from the *in vitro* situation to man may be made.

5.6.3.3. Conclusions of the workshop, by Gisbert Sponer

In this conclusion, both possibilities and limitations were clearly presented.
There are a lot of possibilities to reduce animal use, but only few for real and complete replacement. Only in vaccine testing, there seem to be good chances for replacements.
However, some important ideas were presented. The need for better science was mentioned. Better science means better experiments, better protocols, clear questions,

better interpretation of the data etc. Better science gains more knowledge about physiology and pathophysiology, which can provide new chances for drug development and more implementation of the 3R concept.
Making available negative results is another important issue.
International harmonisation of the requirements of the regulatory bodies is important, but difficult to achieve, because a variety of different players are involved.
Cross talk between scientists and regulators is important, this can already be considered before research or animal experimentation is started, this may help to interpret correctly guidelines and to answer other questions.