Drug Development and Critical Analysis of the Reliability of Preclinical Studies

Wissenschaftliche Prüfungsarbeit
zur Erlangung des Titels

„Master of Drug Regulatory Affairs“

der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Dr. Jennifer Anthöfer
aus Dortmund

Bonn 2015

Betreuer und 1. Referent: Dr. Eckhard von Keutz

Zweiter Referent: Dr. Gerd Bode

Table of Contents

Table of Contents **3**

List of Abbreviations **5**

List of Tables **7**

List of Figures **8**

Introduction **9**

Drug Development **10**

I. Drug discovery 11

II. Preclinical examination 15

Toxicity studies 16

i) General toxicity studies 16

ii) Genotoxicity and carcinogenicity studies 18

iii) Reproduction toxicity studies 21

iv) Other toxicological studies 21

Pharmacology studies 22

i) Primary pharmacodynamic (PD) studies 22

ii) Secondary pharmacodynamic studies 22

iii) Safety pharmacology studies 23

Toxicokinetic & pharmacokinetic studies 24

III. Clinical trials 27

IV. Marketing authorisation and post-approval processes 29

Reliability of preclinical studies – Results **31**

I. Tools used for preclinical research 31

II. Internal variability - Design, conduct and analysis of preclinical studies 33

III. External variability and validation of animal models 35

Comparison of the study design 38

IV. Public bias 39

Reliability of preclinical studies – Discussion and conclusion 41

Shortcomings and potential solutions during implementation 42

Need for collective consciousness 44

Resolving a systematic bias 45

Further improvement suggestions 46

Summary 49

References 51

Declaration 57

List of Abbreviations

ADME	Absorption, distribution, metabolism, excretion
AG	Aktiengesellschaft
AUC	Area under the curve
BLI	Bioluminescent imaging
CAAT	Centre for Alternatives to Animal Testing
CP	Centralized procedure
CMR	Centre of Medicines Research
CMS	Concerned Member State
CTD	Common Technical Document
DCP	Decentralised procedure
DKMZ	German Collection of Microorganisms and Cell Cultures (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH)
EEA	European Economic Area
EMA	European Medicines Agency
EU	European Union
EudraCT	European Union Drug Regulating Authorities Clinical Trials
FDA	U.S. Federal Drug Administration
FIM	First in man
GCP	Good Clinical Practice
GCCP	Good Cell Culture Practice
GEMM	Genetically engineered mouse model
GLP	Good Laboratory Practice
HeLa cells	Human cervical carcinoma cell line from a patient named "Henrietta Lack"
HSC	Hematopoietic stem cells
HT1080	Human fibrosarcoma cell line
HTS	High-throughput screening
IL	Interleukin
IMPD	Investigational Medicinal Product Dossier
i. v.	Intra-venous
LC-MS/MS	Liquid chromatography-tandem mass spectrometry

MABEL	Minimal anticipated biological effect level
MLA	Mouse lymphoma cell Tk (thymidine kinase) gene mutation assay
MRP	Mutual recognition procedure
MS	Mass spectrometry
MFD	Maximum feasible dose
MTD	Maximum tolerated dose
NCE	New chemical entity
NMR	Nuclear magnetic resonance
NOAEL	Non-observed adverse effect level
NOD	Non-obese diabetic
OECD	Organisation of Economical Co-Operation and Development
PASS	Post-authorisation safety study
PD	Pharmacodynamics
PDTX	Patient-derived tumour xenografts
PET	Positron emission tomography
PPARalpha	Peroxisome proliferator-activated receptor alpha
RMS	Reference Member State
SCID	Severe combined immunodeficiency
U. S.	United States of America

List of Tables

Table 1: Recommended duration of diverse toxicology studies for subsequent marketing ([22], amended) 17
Table 2: High dose selection in carcinogenicity studies 20
Table 3: Validity scoring system used to find the optimal combination of animal models (adapted from [56]) 35
Table 4: Differences in characteristics found in both preclinical studies and clinical trials [44,58,59] 38

List of Figures

Figure 1: Main pillars of the drug developmental process 11
Figure 2: Course of the drug development process from discovering a target of issue for a special disease to the use of non-clinical trials.....12
Figure 3: Pharmacokinetic parameters showing a typical example of a plasma concentration time profile after oral administration; C_{max}: maximum concentration, t_{max}: time to c_{max}, AUC: area under the curve.....25

Introduction

The present master thesis will offer a scientific overview of the topic regarding preclinical studies in drug development, with a particular focus on the translational process from animals to humans and the potential obstacles which are necessary to be overcome. Translational research is the interface between preclinical research and clinical development and tries to turn potential innovative drugs into new medicinal products. The translational process from in vitro analyses and in vivo animal models to a first in man application in clinical trials helps to bridge the gap between basic research and the market access for medicinal products.

For a better understanding of this process, the first part of this thesis contains a detailed description of the drug development process, beginning with the drug discovery including the target validation and the finding and optimisation of special lead compounds. The most promising lead compounds are tested in preclinical studies which are specified in greater detail in the following chapter. A series of toxicological and pharmacological studies have to be performed prior to and during clinical trials in order to determine potential hazards, as well as to find the optimal dose for the use in humans. After a short characterisation of clinical trials in regard to the requirements during the application and the differentiation in four phases, a short overview of the marketing authorisation and necessary post-approval processes is provided.

In the second part of this master thesis the critical evaluation of preclinical studies and the reliability concerning the translation to the clinic are more closely elucidated. Potential disparities between the findings in preclinical studies and clinical trials and factors contributing to the lack of reproducibility of preclinical research are described in more detail. Issues such as the internal and external validity of preclinical studies, as well as the importance of using suitable and validated tools are addressed and causes of failed translation such as investigator and public bias are emphasized. Furthermore, possible improvement concerning the conduct of preclinical studies and the handling of the achieved findings are an important subject of this thesis.

Drug Development

From bench to bedside – drug development seems to be a simple process. Closer inspections show, however, that drug development is an exceedingly complex exercise in which many issues have to be taken into account. From the idea to develop a new drug through to the receipt of a marketing authorisation for a new medicinal product it can take up to 12 – 15 years, or in some cases even longer. For several medicinal products the costs of the complete development program can reach up to $ 1 billion [1–3]. Other experts even expect an average of between $ 4 and $ 11 billion for a successful launch of a substance onto the market [4]. Therefore it is of enormous importance for the pharmaceutical industry to have a sophisticated time and cost-effective strategy of the outset of the upcoming development. It should be mentioned that the success rate in drug development is very low. Only one out of 5,000 – 10,000 compounds will reach the long-awaited approval and can be launched onto the market. The more advanced the development of a new lead compound, the more expensive it will be. Well-defined hypotheses as well as valid analytical methods are prerequisites for quite a high level of success. Nowadays, "failing fast and cheap" is an indispensable basic rule of thumb. Once entering the clinical trial phase, the discontinuation of the project is difficult due to attained public notice. A detailed analysis conducted by the Centre for Medicines Research (CMR) demonstrates that failure rates of up to 95 % in Phase II and Phase III clinical trials - meaning an elimination of 95 % of the drug candidates – can be observed nowadays [5]. Therefore suitable and validated (preclinical) methods are key requirements in drug development to identify potential failures and to sort out the respective substances before entering the clinical phase.

In order to develop a new medicinal product in the pharmaceutical industry, the definition of a drug should be clarified first. A drug can be defined as '*a substance or product that is used or intended to be used to modify or explore physiological or pathological states for the benefit of the recipient*' [6]. Furthermore, due to associated risks and side effects of the drug, the use of medicinal products is strictly regulated. Drug regulation particularly evolved in the context of therapeutic disasters. In the United States President Franklin D. Roosevelt e.g. signed "The Food, Drug, and Cosmetic Act" in 1938 after an incident with the substance diethylene glycol. More than 100 people – many of them children – died after taking the new medicinal product called "Elixir Sulfanilamid" developed by a Tennessee pharmaceutical company. At that time, toxicological examination before releasing a medicinal product for sale was not required and the product was only tested organoleptic (appearance, fragrance and flavour) [7]. After the introduction of the new law, the manufacturers of new drugs were obliged to prove

to the U.S. Federal Drug Administration (FDA) that the product was safe before the launch onto the market [8]. In the European Union (EU), the Directive 65/65/EEC was enacted in 1965 after the thalidomide crisis ("Contergan") in 1962. Even in this case inadequate safety-testing was the common cause of death and also resulted in deformation of newborns [9]. Nowadays the approval of medicinal products for human use within the EU is regulated by the Directive 2001/83/EC and its amendments [10]. Since then steady developments in drug regulations emerged and still persist today. Therefore the drug development process is a consistently regulated procedure and has to pass through several stages in order to come up with a safe and effective drug.

Developing a new drug can be broadly divided into four main pillars: Drug discovery, preclinical development, clinical studies and marketing authorisation of the new chemical entity (NCE) with subsequent post-marketing activities (Figure 1).

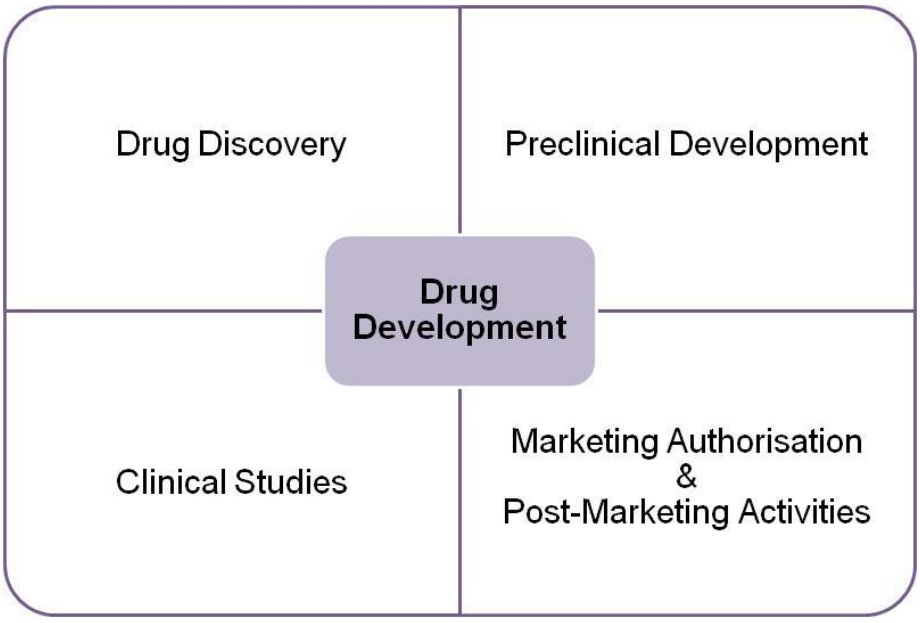

Figure 1: Main pillars of the drug developmental process

I. Drug discovery

One of the first steps in drug discovery is the target identification and validation for medical intervention (Figure 2). In early days, diseases and pathology were taken into consideration for finding NCEs. Nowadays evaluating target engagement in the target tissue is the recent strategy. A deeper understanding of failed biochemical processes responsible for the disease at issue and the associated targets on molecular level like receptors, ion channels, enzymes or elements within gene expression systems are of interest.

For finding new targets, several sources exist, i.e. genomic and proteomic data, compound profiling data or associations between genetic or phenotypic expression and the disease of issue [3]. An important factor for the efficacy of therapeutics is an adequate amount of the appropriate target structures at the site of action, a sufficient pharmacological activity, as well as the proof of target engagement. For the measurement of target engagement, biomarkers available in various forms can be used, e.g. markers to define the presence or the severity of a disease or drug effect biomarkers, which assess the degree of target engagement. For target validation, target expression in the disease-relevant cells or tissue is determined as well as the possibility to modulate this target with drug-like molecules [11–13]. The evaluation of drug binding to target proteins may be carried out using a number of different techniques. Imaging techniques like positron emission tomography (PET), liquid chromatography coupled to mass spectrometry (LC-MS/MS) or in vivo bioluminescent imaging (BLI) may be the method of choice. Biodistribution studies may be carried out to confirm binding of the drug to the relevant target structure [14]. Another approach for validation of the potential target structure is the antisense technology. Applying this method, enables modified oligonucleotides to complementarily bind to potential target mRNA structures and hence, inhibit the synthesis of the encoded protein [3]. By applying suitable target engagement assays, the repeat use of animal testing can be diminished or even avoided.

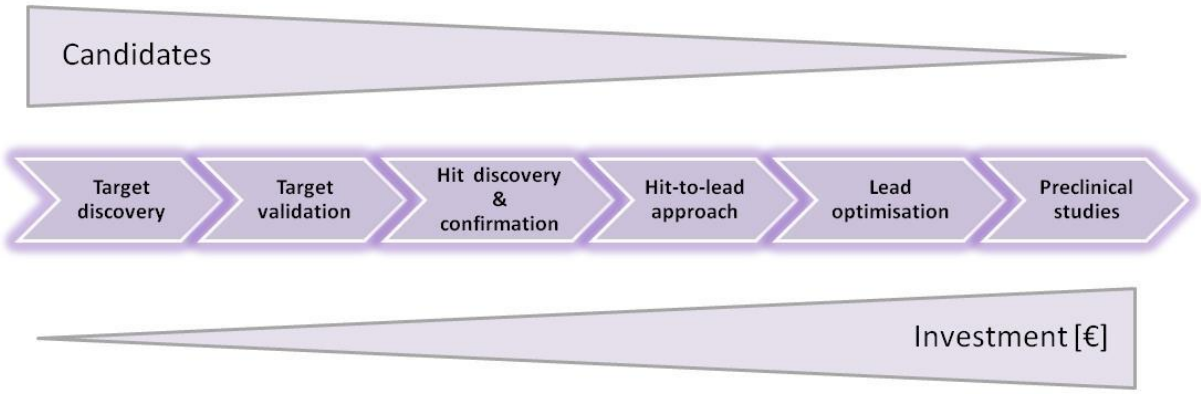

Figure 2: Course of the drug development process from discovering a target of issue for a special disease to the use of preclinical studies

One of the next stages is the hit discovery and confirmation process (Figure 2). Within this phase different methods such as combinatorial chemistry or high-throughput screening (HTS) technology are used to determine many drug-like or fragment-like molecule hits. The initial screening reveals hits which exhibit binding affinities for the identified target in the range of

micromolar concentration. Further potential assays are biochemical tests or cell-based assays in which, for instance, stable mammalian cell lines are developed over-expressing the target of issue. Affinity screening techniques like nuclear magnetic resonance (NMR) and mass spectrometry (MS) or X-ray crystallography are more suitable for the primary screening of small fragment-like molecules. If a development of an HTS technique is impossible, in silico focus-like or fragment-like screening can be the method of choice.

During the development of customized assays, different factors should be taken into consideration such as the reproducibility of the tests, the quality of the results as well as the pharmacological relevance meaning identifying compounds with the desired properties. The aim of a large number of analytical tests is to eliminate all compounds with unfavourable properties like toxicity or inappropriate pharmacokinetic profiles as early as possible.

Traditional HTS often includes a high rate of false-positive and false-negative results. Therefore for enhancing the performance, it is appropriate to confirm potential candidates with secondary assays [3,15]. One available opportunity is testing the identified hits against another member of the target family, whereby other assay conditions such as reagents or parameters remain constant. If the observed activity in the primary as well as in the secondary test is nearly identical, this can often be an indication for a false-positive result. All-or-nothing responses in dose-response curves indicate the infeasibility of possible candidates for reversible effects. The parallel use of the above mentioned technologies is advisable in order to ensure significance of the results. Representative hits will finally be re-synthesized and re-examined to verify their activity against the target of interest.

After finding initial hit compounds, these hits can be optimised towards a higher affinity interaction in the nanomolar range in order to obtain so called ‘lead compounds’ [16,17]. This can be done by methods as hit evolution where analogues of the initial hits are synthesized or hit fragmentation where promising fragments of large hits can be identified [15]. These lead compounds represent first structures for pharmacological or biological active drug candidates. Further modifications are normally necessary to optimise the lead compound for subsequent preclinical studies [18,19]. Key features beside the initially identified potency are selectivity, physicochemical properties like solubility and permeability of the respective hit as well as the opportunity of industrial production. Compounds that lack these properties can hardly be developed to a medicinal product authorised for human use. High solubility and permeability are prerequisites for a successful pharmacokinetic profile which include the absorption and distribution of the drug within the body in addition to the metabolism and excretion (ADME, please refer to preclinical examination, chapter II).

Modifications of the compounds of issue serve to maintain the favourable properties and at the same time improve the deficiencies. Subsequent analysis should ensure the selectivity of the lead compound towards the discovered target and initial studies concerning the physicochemical and pharmacokinetic as well as the toxicological profile shall guarantee a smooth transition from drug discovery to the preclinical examination (for further information see chapter II) [3,15].

Reasons for an early exclusion of a candidate could cover, amongst other topics, discovered toxicity, undesirable side effects or insufficient pharmacodynamic and pharmacokinetic profiles.

II. Preclinical examination

The most successful lead compounds found during the drug discovery program are tested in preclinical studies. Preclinical examination is the second step in the drug development process. The transition from drug discovery to preclinical testing is a smooth-running procedure due to the fact that different toxicological profiling concerning the respective compounds is already conducted during the drug discovery phase. Potential toxicities can already be predicted due to structural alerts or the mode of action of the compound.

An obvious objective of the implementation of non-clinical safety studies is the detailed characterisation of toxic effects of the respective compound in regard to potential target organs, to reveal the possible reversibility of the mode of action as well as to identify parameters for monitoring potential adverse effects in clinical trials. Another aim of preclinical examination is the identification of pharmacological properties like the mode of action (pharmacodynamic) and the metabolism (pharmacokinetic) of a substance. Furthermore, pharmacological testing should lead to an extrapolation of animal data to humans. A deeper understanding of the toxicological profile and the pharmacological properties should establish a safe initial dose for first in man (FIM) exposure in clinical Phase I trials (see Chapter III) and characterise potential adverse effects. Preclinical studies are conducted prior to and even during the clinical studies to ensure continuous monitoring.

The non-clinical safety assessment of a new medicinal product in each case implies the following studies:

1. General toxicity studies
2. Pharmacology studies
3. Toxicokinetic and pharmacokinetic studies
4. Reproduction toxicity studies
5. Genotoxicity studies
6. Carcinogenicity studies (drugs with a special concern or intended for a long time of ingestion)

Other kinds of studies such as phototoxicity, immunotoxicity, juvenile animal toxicity or abuse liability are executed on a case-by-case basis. For products using special innovative therapeutic modalities like siRNA or vaccine adjuvants, particular studies may also be waived, abbreviated, deferred or even added. The ICH guideline M3(R2) provides important insights into recommended international standards and promotes the harmonisation of

preclinical studies among the regions of the European Union, Japan and the United States. Furthermore with this guideline a reduction of in vivo experiments in accordance with the 3R principles “reduce, refine and replace” can be achieved. Nowadays the replacement of animal testing against new in vitro alternative methods represents an important milestone [20].
Requirements and characteristics of the different kinds of studies are listed in the following section.

Toxicity studies

Toxicology defines the preclinical part of the safety assessment during drug development. By conducting toxicity studies, possible hazards and risks are identified. During the risk assessment an extrapolation of the received non-clinical animal data to humans takes place where possible. In order to gather more information about the respective substance toxicology studies are conducted during the whole drug development – from drug discovery to eliminate inappropriate compounds to determining safe doses for the FIM trials and even during the clinical trials. [21].

i) General toxicity studies

Acute toxicity studies

Separate single dose or acute toxicity studies are not normally considered necessary anymore [22]. Acute toxicity can be assessed from short-term dose-ranging or dose escalation studies. Furthermore, these studies can be conducted as non-GLP (Good Laboratory Practice) studies. The main goal of acute toxicity studies is the identification of potential target organs which are toxicologically influenced by the administered substance. A calculation of the median lethal dose (LD_{50}) is no longer recommended.
Extended single dose toxicity studies can be conducted to support exploratory clinical trials (mainly single dose human trials). With this type of study, parameters like haematology, clinical chemistry or histopathological data can be evaluated [22].

Repeated dose toxicity studies

Repeated dose toxicity studies are conducted to characterise toxicological profiles, for example, to identify target organs and tissue toxicologically influenced after repeated administration of high doses. Another aim of this study type is to establish no-effect levels, like the non-observed adverse effect level (NOAEL) serving as safety margins and to determine the highest dose for subsequent toxicological studies. Repeated dose toxicity

studies have to be conducted before first-in-man (FIM) trials and therefore support the conduct of clinical trials. The studies have to be conducted in accordance with GLP.

It is generally required to perform these studies in two animal species whereby one should be a rodent (e.g. rat or mouse) and the other should be a non-rodent species like primates, mini-pigs or dogs. As to be able to draw the right conclusion and to translate preclinical findings to patients in clinical trials, the animals are always required to be the "most human-like animal species". The pharmacokinetic profile, pharmacodynamic effects as well as metabolic data should be as similar as possible to data received in human trials [21].

Table 1 indicates the duration of different toxicological study types. Generally, however, the principle shall thereby apply that the toxicological studies take as long as or even longer as the intended clinical use. Even the route of administration should be similar to the intended human usage unless a suitable justification for a deviation can be provided, for instance, the small size of juvenile animals making an i.v. injection impossible.

Table 1: Recommended duration of diverse toxicology studies for subsequent marketing ([22], amended)

Study type	Duration	Rodent	Non-rodent
Single dose (acute)	One single dose	One single dose	One single dose
Repeated dose			
Subacute	Up to 2 weeks	1 months	1 months
	2 – 4 weeks	3 months	3 months
Subchronic	1 -3 months	6 months	6 months
Chronic	> 3 months	6 months	9 months
Carcinogenicity	24 months	24 months	---

The basic principle concerning the dose selection for toxicity studies has to be followed so that the applied dose is higher than the intended dose for humans in clinical trials. Dose selection should be based on data received from other studies like pharmacology or pharmacokinetic studies. One control group and 3 treatment groups should normally be used in toxicology studies:

- Control group receiving placebo/vehicle
- Low dose group receiving the NOAEL, dose should be in the range of the therapeutical human dose, initiation of pharmacological effects
- Mid dose first reports of toxicological effects
- High dose appearance of complete toxicological profile (for high dose selection please refer to the description of cancerogenicity studies below)

During and after these studies several parameters are examined. Among other things, this includes physiological parameters like blood pressure, electrocardiography (ECG), electroencephalography (EEG) or electroretinogram (ERG) as well as haematological issues. In addition, further information is collected by application of clinical chemistry which addresses the determination of the presence of different ions and substances in the body like sodium, potassium, glucose or albumin. Today diverse biomarkers are available to further characterise the toxicological effects of the compound of issue.
After the inevitable necropsy of the animals at the end of the studies, tissue and organs are gathered. Histological examinations like H&E staining as well as immunohistochemistry further helps characterising the effects after administration of the substance [23].

ii) Genotoxicity and carcinogenicity studies

Genotoxicity describes the property of a substance to induce genetic damage on DNA or chromosomal level by a variety of mechanisms. Damage such as mutation in germ or somatic cells can lead to permanent heritable changes. Direct or indirect DNA damage which cannot be prevented by DNA repair or cell apoptosis is frequently considered to be essential for a multi-step cascade which finally can lead to the establishment of cancer.
Many in vitro and in vivo assays exist to examine the genotoxic potential of possible drug candidates [24]. A description of the standard battery to be conducted for a comprehensive genotoxic examination can be found in the ICH Guidance S2(R1). The internationally approved description of the potential assays as well as standard test protocols can be found in the respective OECD (Organisation for Economic Co-Operation and Development) guidelines. By conducting a battery of in vitro and in vivo tests the capability of detecting most of the genotoxic mechanisms regarding potential new pharmaceuticals increases. Genotoxicity assays should always be carried out according to GLP.

The first assay which should be conducted is a gene mutation test in bacteria. In most cases, the bacterial reverse mutation test (Ames[1]) using S. typhimurium is used. This test makes it possible to reliably indicate relevant genetic aberrations as well as most of the genotoxic carcinogens in rodents and humans [25]. Further in vitro tests evaluated in mammalian cells can follow and may detect gene or chromosomal damages. These include, for instance, the in vitro micronucleus assay or the mouse lymphoma cell Tk (thymidine kinase) gene mutation assay (MLA). Subsequent in vivo tests can prove or disprove the first results received by the in vitro assays. Furthermore, they should ensure that substances not mutagenic in vitro, e.g. benzol or hydroquinone, can nevertheless be identified as potentially genotoxic. Suitable tests can be micronucleus tests using rodent hematopoietic cells or tests detecting DNA strand breaks as the single cell gel electrophoresis ("Comet" assay). If two in vivo assays are conducted showing negative outcomes, this is considered as sufficient to demonstrate the lack of genotoxicity [24].

Substances which show positive results in genotoxic assays have the potential to be carcinogenic. The ICH safety guidance S1A – S1C provide information including the requirement for carcinogenicity studies and the suitable kinds of assays for evaluating the carcinogenic potential of pharmaceuticals as well as the dose selection regarding this special type of study. It is advisable to conduct carcinogenicity studies as soon as a tumourigenic potential in animals or an unequivocal risk to humans is deemed to be given. There is a need for carcinogenicity studies when there is any cause of concern like equivocal genotoxicity assays, evidence of long-term tissue retention or even a structural affinity with known carcinogens. Furthermore, carcinogenicity studies have to be conducted if the expected clinical use is continuous for at least 6 months or a repeated use in an intermittent manner is given as is the case for allergic rhinitis or depression [26]. Studies are normally not required for pharmaceuticals of topical use, short-term treatment or unambiguous genotoxic result as well as for cancer patients with low life expectancy. With this kind of study, the formation of cancer (tumourgenesis) and its mechanisms are then examined in greater detail. For this reason and due to the long period of examination, the studies have to be precise and particularly sensitive. The dosage as well as the species should be carefully selected and even parameters such as the diet or the microbial status should be closely monitored.

Historically, carcinogenicity studies are conducted in two rodent species (mostly rats and mice) over a long-term period of 2 years. If the carcinogenic potential occurs in both species, an interspecies extrapolation is possible. Rats are normally the species of choice due to their

[1] first described by Bruce Ames in the early 1970s

improved greater sensitivity in comparison to mouse models. Literature based research has shown that all known human carcinogens are positive in rat models [27].

Instead of conducting two long-term studies in two species, another practicable approach is listed in the ICH safety guidance S1B. One long-term treatment – generally conducted with rat models – and one additional short or medium-term study are adequate to determine a possible carcinogenicity of the respective substance [28]. Possible mouse models for the short and medium-term studies are the p53 knock out mouse with an inactivated tumour suppressor gene or the Tg-rasH2 or Tg.AC skin model with an activated oncogene [27].

Table 2: High dose selection in carcinogenicity studies

Kind of dose	Explanatory notes
Maximum tolerated dose	Can be predicted from a 3-month dose-range finding study in which a minimal toxicity can be observed (e.g. target organ toxicity, decrease in body weight).
25-fold AUC* ratio	25:1 exposure ratio of rodent to human plasma AUC of the parent compound or the appending metabolites.
Saturation of absorption	The mid and low doses used for the carcinogenicity study should take into account the saturation of metabolic and elimination pathways.
Pharmacodynamic endpoints	Dose-limiting pharmacodynamic effects like hypotension or the inhibition of blood clotting can be considered.
Maximum feasible dose	Maximum feasible dose by dietary administration is considered to be 5 % of the diet.

* **area under the curve**: suitable pharmacokinetic endpoint which takes into account the plasma concentration of the respective compound and the residence time in vivo

In each of the stated studies, 3 treatment groups and one control group are normally used. The route of administration should be the same as that used in humans. For finding the appropriate high dose a 3-month dose-range finding (range of five dose levels) has to be conducted first. For determination of the high dose all available relevant animal and human data should be taken into consideration. In the past, dose selection was almost exclusively based on the

maximum tolerated dose (MTD). Nowadays, alternative kinds of doses can be used as high doses in cancerogenicity studies (Table 2).

iii) Reproduction toxicity studies

Reproduction toxicity studies include reproductive toxicology which analyses the risk for male and/or female fertility as well as developmental toxicology which examines the toxic effects for newborn and unborn in more detail. Reproductive failures in adult infertility, miscarriage or birth defects could appear. Adverse effects induced during pregnancy like teratogenicity is another issue regarding the offspring.

For examination of reproductive toxicity mammalian species and strains already used for other pharmacological and toxicological studies should be used. This allows the comparability with the results already received. With data gathered in repeated dose toxicity studies, important information concerning fertility, in special male fertility can frequently be provided. For reproduction toxicology studies, the use of one rodent and one non-rodent animal species is recommended. Rats are the rodent species of choice as for reasons of practicability and a large amount of knowledge about these animals. Rabbits are frequently used as non-rodent species for embryotoxicity studies [29]. Other in vitro methods like tissue, organs or cell systems may be used to increase the knowledge. As in the studies already mentioned above, the route of administration should be the same as the intended route in humans. The dose applied in reproduction toxicity studies should be carefully selected based on the data from already conducted studies. If no data is available, preliminary studies are recommended [29].

iv) Other toxicological studies

Other toxicological studies are conducted, if required. Immunotoxicity which may be investigated during repeated dose toxicity studies, identifies adverse effects of drugs on the immune system as immunosuppression which can lead to infectious diseases or malignancies, hypersensitivity or autoimmune reactions to self antigens. To determine potential immune reactions, different parameters like antibodies (IgM, IgE, IgG, etc.) are quantified, lymph nodes are weighed or lymphoid cell morphology is analysed [23].

If local tolerance studies are evaluated, this should be done with identical formulation intended for later marketing. Even these studies can often be conducted during repeated dose toxicity studies whereby there is no necessity for additional studies. With local tolerance studies, the tolerability of sites in the body which come into contact with the respective drug is investigated. These days, there is regulatory acceptance for using in vitro methods

scientifically validated as part of the whole testing strategy. Potential sites for investigation can be the skin, the eyes and ears as well as unintended paravenous or intra-arterial administration [30].

Photosafety studies are another type of toxicological studies which are only conducted on request. The best time for investigation is preferably before the exposure of large numbers of subjects in clinical phase III trials. Two different effects are addressed in the ICH safety guideline S10 regarding the photosafety evaluation of pharmaceuticals: phototoxicity and photoallergy. Phototoxicity defines an acute light-induced tissue response to photoreactive chemicals whereas photoallergy describes an immunologically mediated reaction like the formation of protein adducts which subsequently may induce a photochemical reaction [30]. Photosafety studies are only conducted if at least one of the following criteria is met:

1. The absorption of the substance is within the range of 290 – 700 nm
2. Generation of a reactive species following absorption of UV/visible light
3. Sufficient distribution of the respective substance to light exposed tissue

Pharmacology studies

Pharmacology studies can be divided into primary and secondary pharmacodynamic as well as safety pharmacology studies.

i) Primary pharmacodynamic (PD) studies

Primary PD studies are intended to provide a deeper insight into the mode of action and the effect of a compound in regard to its desired therapeutic target [22]. Primary PD studies are in general not conducted in accordance with Good Laboratory Practice (GLP) due to the time of execution. Primary PD studies are frequently already carried out during the discovery phase whereby in vitro and in vivo testing is applicable [20]. This type of study is mainly focused on the efficacy of a substance.

ii) Secondary pharmacodynamic studies

Secondary PD studies are mainly focused on the mode of action and the effects of the relevant compound which are not related to its desired therapeutic target. This type of study is important in terms of safety issues.

The difference between primary and secondary PD studies can thoroughly be illustrated by means of statins which are used as lipid lowering drugs. Statins are HMG-CoA (3-hydroxy-3-methylglutaryl-coenzyme A) reductase inhibitors which lower the cholesterol level by inhibiting the enzyme HMG-CoA reductase. Whereas the primary PD studies focus on the desired pharmacological effect – lowering of cholesterol – the secondary PD studies investigate the undesired therapeutic effects like lowering the ubiquinone and dolichol level which can cause adverse effects on the muscles and the testicles [31].

iii) Safety pharmacology studies

Safety pharmacology studies are conducted to identify possible undesirable pharmacodynamic effects of a compound on selected physiological functions which may have an impact on human safety. The exposure can fall within the therapeutic range and may even exceed it. Another objective of safety pharmacology studies is the evaluation of pharmacodynamic and pathophysiological effects observed in toxicology studies as well as the investigation of the exact mechanisms of these effects [32].

A distinction can be made between three types of safety pharmacology studies which are described in the following:

a) Core battery

The core battery of safety pharmacology studies which should be conducted in accordance with GLP is mandatory in order to investigate before first administration in humans. The core battery implies organ systems which are important with respect to life-supporting functions and are therefore most critical for life. This includes the cardiovascular, respiratory and central nervous system. Thereby, in vitro studies on isolated tissue, cells, receptors, ion channels or enzymes are an initial method to investigate potential pharmacological effects in concentration ranges of the respective substance on which an effect seems probable. For subsequent in vivo studies, the expected clinical route of administration should be used and the animals should ideally not be under anaesthesia [31,32].

Safety pharmacology studies are normally performed by a single dose administration, whereby the exposure should at least be similar or even higher than the potential therapeutic concentration in humans.

b) Follow-up studies

The follow-up studies for the core battery may provide a deeper insight into kinetic conditions, potential repeat dose administrations or suitable animal species. Follow-up studies need not necessarily be conducted before first administration in human beings but are carried out before the potential approval of a drug [31].

c) Supplemental studies

In supplemental safety pharmacology studies organ systems not addressed in the core battery are investigated. This is notably done with other major organ systems such as the gastrointestinal, renal or the immune system.

Toxicokinetic & pharmacokinetic studies

The pharmacology and toxicology studies described above mainly focus on desirable and undesirable effects of a compound and its influence on the body. Toxico and pharmacokinetic studies are conducted prior to and/or during the clinical development phase to determine the detailed effects caused by the body after application of the substance of issue as well as to clarify observed suspected, undesirable effects.

The main task of pharmacokinetic studies is to find an optimal dose level and to provide information about the dose-effect relationship. Therefore, different processes in the body are investigated and intensive information about the absorption, distribution, metabolism and excretion (ADME) of the substance is generated. The bioavailability of a substance is linked to its absorption. A poor solubility or the inability of a substance to permeate different tissues like the intestinal wall reduces the absorption of a substance. The distribution mostly takes place after transfer of a substance into the blood stream where it is transported into the target tissue or organs. By entering the body, the substance is metabolised, mostly by redox enzymes of the liver like cytochrome P450. So called “pro-drugs” for example do not carry out their function unless they are metabolised to the active substance. Excretion of substances usually takes place through the kidneys (urine) and to a lesser extent through the faeces. Further parameters which may be carried out during pharmacokinetic studies are the plasma half-life, the clearance and the mean residence time of a substance as well as protein binding or steady state conditions [33].

Where the compounds show a long half-life, an incomplete elimination or unexpected organ toxicity, the conduction of repeated dose tissue distribution studies are recommended. These

studies may provide information about the distribution and accumulation of a substance and its metabolites in the body. This can be useful for planning and the interpretation of further pharmacological or toxicological studies [34].

The term "toxicokinetic" is defined as '*the generation of pharmacokinetic data, either as an integral component in the conduct of non-clinical toxicity studies or in specially designed supportive studies, in order to assess systemic exposure*.' [35]. This means not all pharmacokinetic data but rather selected data is of special importance for the interpretation of toxicological concerns and the calculation of safety margins. Toxicokinetic studies are generally carried out in a much higher dose than those used in pharmacokinetic studies and are mainly focused on systemic exposure. Toxicokinetic studies can already be integrated in repeated dose studies and/or carcinogenicity studies whereas pharmacokinetic studies are independent animal studies [33].

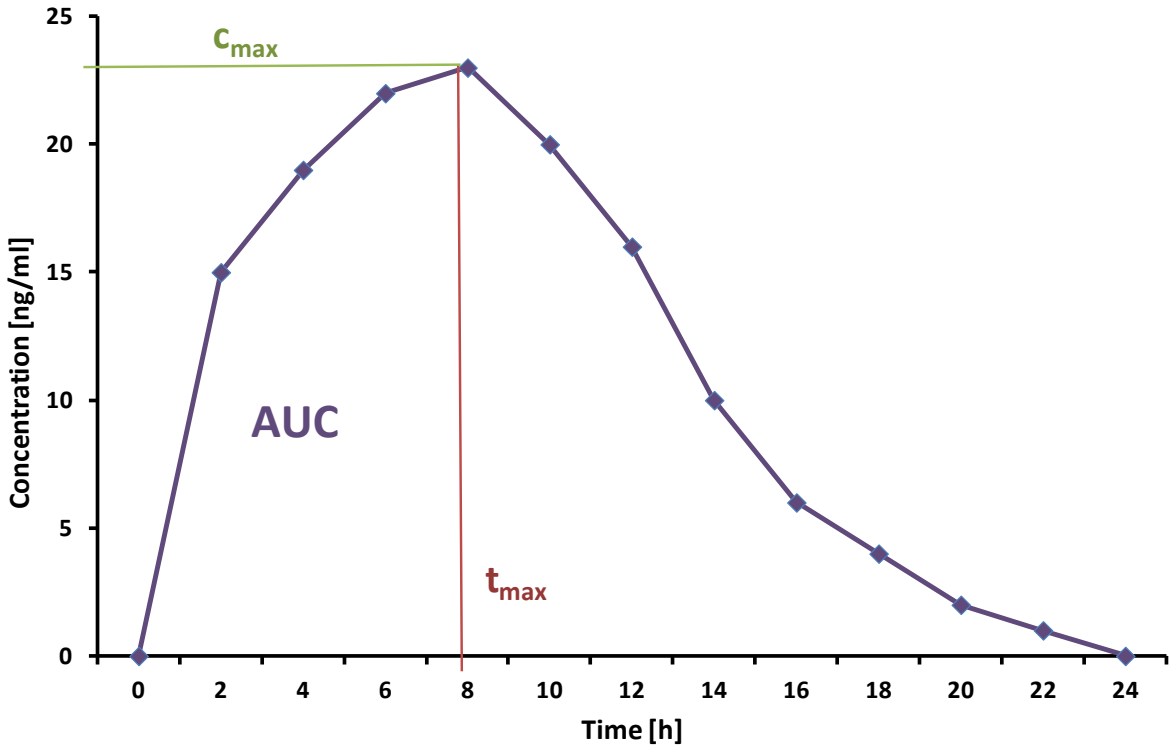

Figure 3: Pharmacokinetic parameters showing a typical example of a plasma concentration time profile after oral administration; C_{max}: maximum concentration, t_{max}: time to c_{max}, AUC: area under the curve

During toxicokinetic studies GLP must be adhered to and only selected kinetic data is gathered and determined during preclinical safety studies, i.e. the maximum serum concentration (C_{max}), the time at which the maximum is observed (t_{max}) and the total drug exposure over the time (area under the curve, (AUC), figure 3) which can be equated with the

bioavailability of a substance. The aim of all preclinical studies is to find the “most human like” animal species in order to be able to extrapolate from animals to humans. An extrapolation is best carried out if the received pharmacodynamic and pharmacokinetic data is interspecifically consistent.

Following the description of the drug development process, this master thesis deals with the issue of translational research from preclinical testing to the clinical application as well as with the reliability of preclinical findings and potential improvements concerning this translational process.

III. Clinical trials

As a next milestone in drug development, the most promising candidates of the preclinical studies are tested for their human compatibility in clinical trials. These interventional trials are intended

"to discover or verify the clinical, pharmacological and/or other pharmacodynamic effects [...], and/or to identify any adverse reactions [...], and/or to study absorption, distribution, metabolism and excretion [...] with the object of ascertaining its (their) safety and/or efficacy." [36]

Before clinical trials can be conducted, they first have to be approved by the corresponding competent authorities and require a favourable opinion of the respective ethic committee.
For this reason, the sponsor of the clinical trial has to follow an internationally accepted and harmonised scheme which was introduced by the Clinical Trials Directive 2001/20/EC. A EudraCT (European Union Drug Regulating Authorities Clinical Trials) number has to be requested which is a unique reference for the respective trial. Furthermore, filing an IMPD (Investigational Medicinal Product Dossier) to the competent authority becomes necessary. This dossier includes information related to the quality, manufacture and the control of the investigational medicinal product (IMP) including placebo. An additional NIMP (non-Investigational Medicinal Product Dossier) has to be submitted to the authority in case of planning to use medicinal products as background or rescue medication. Other documents to be submitted cover, among other things, the Clinical Trial Application (CTA), the Investigator's Brochure, the Clinical Development Plan or the PIP (Paediatric Investigation Plan) decision. All clinical trials have to be carried out in compliance with Good Clinical Practice (GCP).
The clinical trials can be divided into four main phases:

Phase I — In Phase I trials, often called first-in-man trials (FIM), first administration of the potential new medicinal product is conducted on a small number of subjects. The IMP is normally tested in healthy volunteers unless performing trials with special drugs for diseases such as cancer.
The aim of this type of trial is to find a suitable route of administration and an optimal dose range with the lowest dose at which the therapy is still effective and the highest dose without causing any harm. Further information on pharmacokinetic and pharmacodynamic relationship, the bioavailability of the

drug, sometimes tested in fasted and fed state, drug interaction or the effect of gender and age, can be investigated in Phase I trials.
The applied dose has to be justified by the general and safety pharmacology as well as by the toxicological data determined in the preclinical studies (see chapter III).

Phase II — First evaluation of the efficacy and short-term safety of the drug in several hundred patients with the disease of interest for a limited treatment duration of days to a few weeks. After proof-of-concept studies, the optimal dose and the length of the treatment is determined in dose finding studies. Another topic is to identify common side effects.
This type of trial is frequently placebo-controlled and double blinded meaning that neither the health care professionals/investigators, nor the subjects of the trials are aware of the treatment.

Phase III — Phase III trials are large trials to confirm the results concerning safety and efficacy under clinical conditions. Therefore, several hundred to thousands of patients with the respective disease undergo short and long-term treatment, for a period of up to 12 months. The trials are mostly randomised and double blinded. Only the optimal dose determined in Phase II trials is used to further determine the proof of efficacy and the long-term safety of the drug. Other parameters to be investigated are the benefit-risk ratio as well as the demonstration of superiority or non-inferiority of the drug compared to products already on the market.

Phase IV — Phase IV trials which can be interventional studies after the receipt of a marketing authorisation investigate the benefit-risk as well as the cost effectiveness and treatment optimisation of an unlimited patient number. Efficacy and safety issues like rare serious adverse events not detectable in Phase III trials due to limited subjects and interactions of medicinal products are also addressed. For these trials no special patient supervision is necessary but a short protocol must be drawn up.

After conduction of each of the four trials the “end of trial“ must immediately be notified to the competent authorities.

IV. Marketing authorisation and post-approval processes

As mentioned in Chapter I of this thesis, obtaining a marketing authorisation for a new medicinal product is necessary to place it on the market and make it accessible to the general public. In the European Economic Area (EEA) the Directive 2001/83/EC and its amendments, as well as the Regulation (EC) No. 726/2004, provide the legal and administrative basis for medicinal products for human use, their authorisation and supervision.

Depending on the medicinal product to be approved and the pharmaceutical company's strategy, different marketing authorisation procedures exist. Medicinal products can fall within the mandatory or the optional scope of the centralised procedure (CP) and following approval by the European Medicines Agency (EMA) the applicant receives a Community marketing authorisation. If the medicinal product does not fall within the mandatory or optional scope of the CP, the applicant has the opportunity to request for an authorisation in more than one Member State of the EEA. Suitable procedures are the mutual recognition as well as the decentralized procedure (MRP/DCP). In both cases, the applicant is able to choose a so-called Reference Member State (RMS) - the competent authority which composes a draft assessment report. Afterwards, the Concerned Member States (CMS) further selected by the applicant have to approve or reject this assessment report and consequently, the respective marketing authorisation. In case of the MRP, the medicinal product is already approved as a national marketing authorisation in one Member State which simultaneously fulfils the role of the Reference Member State. This procedure is based on the mutual recognition by the chosen CMS [36,37].

Different types of legal basis regarding the application type exist, on which the applicant can request for a marketing authorisation:

- Full application according to Article 8(3) of Directive 2001/83/EC
- Application of generic and similar biological medicinal products according to Article 10 of Directive 2001/83/EC
- Well established use application based on bibliographic literature according to Article 10a of Directive 2001/83/EC
- Fixed combination application according to Article 10b of Directive 2001/83/EC
- Informed consent application according to Article 10c of Directive 2001/83/EC

The current requirements for a full application according to Article 8(3) of Directive 2001/83/EC are defined in Annex I of this Directive. The application to be submitted to the competent authorities has to follow the Common Technical Document (CTD) which is an internationally agreed format for preparation of the application of a marketing authorisation submitted in the three ICH regions EU, USA and Japan. It consists of a total of five modules, in which different information has to be provided [38]:

- Module 1: Administrative, regional and national information
- Module 2: Summaries and Overviews of Module 3 – 5 prepared by suitable qualified experts
- Module 3: Chemical, pharmaceutical and biological documentation concerning the medicinal product of issue
- Module 4: Non-clinical study reports
- Module 5: Clinical study reports

Even after launching the drug onto the market, a continuous update of the marketing authorisation is necessary to ensure the protection of public health at any time. The classification of different types of variations may be derived from the Classification Regulation (EU) No. 1234/2008 as well as the related Classification guideline on the operation of the procedures laid down in Chapters II, III, IV of Commission Regulation (EC) No. 1234/2008 [39]. The list of variations includes variations which can be classified as minor variations of Type IA or as major variations of Type II. If one or more conditions for the Type IA variations are not fulfilled, this change can be submitted as a minor variation of Type IB. Whereas variations of Type IA can be implemented before approval of the change by the competent authority, variations of Type IB and Type II must be authorized in advance.
For the maintenance of the marketing authorisation, in addition, the applicant has to perform a series of other tasks like e.g. implementation of pharmacovigilance obligations and its continuous monitoring, as well as conducting potential interventional or non-interventional post-authorisation studies (PASS) to investigate the safety and efficacy of the medicinal product in an unlimited number of subjects (please also refer to Chapter III).

Reliability of preclinical studies – Results

The Institute of Medicine's Clinical Research Roundtable differentiates between two definitions of translational research. The first definition describes translational research as *"[...] the transfer of new understandings of disease mechanisms gained in the laboratory into the development of new methods for diagnosis, therapy, and prevention and their first testing in humans*." [40]. In other words, this definition refers to the transfer of knowledge, mechanisms and special techniques obtained from basic research to the proposed practical application in humans. Scientific discoveries generated in the laboratory at a molecular or cellular level may be transformed into knowledge and are brought to market for diagnosis, prevention and treatment of diseases and therefore shall improve the health and healing of diseases in human beings [40,41]. Another definition for translational research is worded as follows: *"[...] the translation of results from clinical studies into everyday clinical practice and health decision making."* [40]. The results received from the research described above are only the starting point for this second definition. Afterwards, these findings have to be translated into practice meaning the correct implementation of new treatments and research knowledge to reach the appropriate group of patients [40].

More and more questions arise from recent literature as to whether translational research from preclinical testing to clinical application can really be executed reliably, whilst special concerns for improving this translational process are discussed in detail. Causes of failed translation and differences of outcome in animal models and clinical trials of obviously promising interventions are discussed in the following. Contemporary, a closer look is taken at potential improvement in this field.

Promising new molecules designed by computer-based methods aspires to prevent and control a lot of diseases. Many of these compounds reach the clinic but during Phase II or Phase III trials, at the latest, many of them fail mostly due to the lack of efficacy. In many cases animal models are then blamed for being the limiting factor as predictors of responses in humans. However, other factors should also be taken into consideration which may undermine the reliability of preclinical studies [42].

I. Tools used for preclinical research

Prior to the beginning of preclinical examination one needs to ensure the use of suitable tools. For cancer models, for instance, it is of tremendous significance to point out the detailed

requirements for primary cells or cell lines used in the study and to assure compliance with these requirements.

Especially for frequently used cell lines like the fibrosarcoma cell line HT1080 or the breast cancer cell line MCF-7 there will be limitations imposed by different factors. Due to the different handling of the cell lines across various laboratories an intra-laboratory cell line heterogeneity may appear. Continuous culturing using different culture media with additives like fetal bovine serum or various growth factors, which force the growth of the cells, as well as different passaging rates and the use of non-physiological culture conditions (low cell density, hardly any cell-cell contact) play a decisive role for the development of potential sub-populations. In the worst case this can result in genotypic and phenotypic drifts or changed cell characteristics like variations in cell growth or clonogenic activity [43,44]. Furthermore, serially passaged immortalised cell lines may lose their molecular complexities and therefore are often no longer capable of reflecting the characteristics of solid tumours [45]. Continuing morphological similarities and the unawareness of the research groups in many cases prevent a thorough analysis of the cell lines before their use in preclinical studies.

Another challenge can be cell cross-contaminations. In 1981 Nelson et al. demonstrated that many of the examined cell lines showed cross-contamination, especially regarding HeLa cervical cancer cells [46]. As mentioned above, the conduct of cell line characterisation and authentication prior to application in animal models is still a rare event and therefore probably several thousands of studies contain misleading and even incorrect results due to the use of inappropriate tumour models [43]. Not only cell cross-contaminations, but also contaminations of the cells by bacteria, viruses or mycoplasma may distort the results. Mycoplasma infections appear much more frequently than commonly assumed. According to analyses by the FDA and the German Collection of Microorganisms and Cell Cultures (DKMZ), mycoplasma contamination rates of 15 % and 28 % were verified, respectively [47,48]. Studies conducted in Japan even found a contamination of up to 80 % [48]! Mycoplasma infections can trigger cell death, alteration in proliferation rates or can have an influence on the cell metabolism resulting in altered RNA and DNA synthesis. These are just some examples of negative effects which can appear during a mycoplasma contamination. Once infected, it is very difficult to eradicate this infection from cell culture and is frequently associated with disproportionally high effort [44].

Another fact which should be taken into consideration during analysis is the cells' origin. In most cases, cell lines are not derived from primary tumours but from tumour metastases and therefore these demonstrate a more aggressive behaviour. Further consideration should be

given to the question which cell line of a particular type of cancer shall be used. Between cell lines of the same type of cancer, major differences can occur, i.e. the rate of proliferation or different kinds of mutations. A mutation leading to an increased expression of enzymes like the hepatic enzyme p450 cytochrome can lead to potentially different pharmacokinetic behaviours which in consequence can alter the interaction with the drug to be examined [49,50].

II. Internal variability - Design, conduct and analysis of preclinical studies

First of all, a study can only be as good as the investigators who carry it out. Therefore, the expertise of the respective investigator as well as the correct performance and interpretation of the study results may be essential factors for a successful translation from animal testing to clinical trials in human. The investigator bias and a potential lack of scientific rigor play important roles regarding the poor predictability of many of the preclinical studies conducted. In basic research, especially in an academic environment, scientific rigor and a precise hypothesis are rarely applied due to constantly changing investigators. These investigators are mostly students trying to finish their scientific master thesis or PhD thesis. After completion of their thesis they are able to apply for the respective title and other students coming through continuing the work of the pervious students. Investigators and even hypothesis will change and a consistency can hardly be achieved. Every student wants to make best use of the received data, often at the expense of quality. Of course, this phenomenon can be transferred to preclinical studies where investigator bias and a lack of scientific rigor may also occur. Impressive examples for unsatisfactory reproducibility of key findings of studies conducted early in drug development are provided by a review article from Prinz et al. (2011) from Bayer HealthCare [44]. The authors compared published data with in-house findings and found out that in almost two-thirds of the studies inconsistencies were found that '*either considerably prolonged the duration of the target validation process, or in most cases, resulted in termination of the projects*' [51].

A minimisation of failure can already be achieved by suitable target validation (see Chapter I Drug Discovery) followed by intensive investigation of the proof-of-concept during preclinical studies [42,52]. Over-optimistic conclusions about the efficacy of a potential substance and in the following, the failure to translate these finding to the clinic may have several causes. Variability and bias are main challenges during the design, conduct and analysis of preclinical studies. Adequate internal validity of the animal experiments ensure an elimination of the possibility of bias and also assure that disparities observed in various

animal groups merely exist due to the treatment under investigation [53]. Furthermore, following the rules of internal validity the reliability and reproducibility of the studies and the received findings are significantly improved.

Four main types of bias can be distinguished which could be circumvented by just respecting clinical standards, in most cases [54,55]. The first type of bias, the so-called selection bias, can be defined as the biased distribution of animals to the respective treatment groups. The best method to prevent an investigator from choosing the "right" animals for particular treatments is the use of randomisation. In clinical trials, this method is applied in nearly every trial whereby a computer-based generation of random numbers is commonly used. In preclinical studies manual methods like tossing a coin are also acceptable for randomisation. Besides random allocation of the animals, the method of blinding is an important tool for preventing other biases like performance, detection and attrition bias [53]. In the clinical environment, double-blind trials are commonly used meaning that neither the investigator nor the patient are aware of the treatment allocation. The appliance of double-blinding does not make sense in regard to preclinical studies due to the unawareness of animals of the treatment conditions. Nevertheless, keeping the investigator who performs the experiment, collects the data and analyses it unaware of the treatment allocation, prevents him from systematically supplying additional care to special treatment groups (performance bias). Furthermore, by knowing the allocation of the individuals to different treatment groups, the investigator may – consciously or not – be tempted to assess the findings differently, according to the respective treatment group (detection bias). The last bias potentially having an important impact on the reliability of preclinical findings is the attrition bias. In clinical trials the eligibility criteria for the inclusion and exclusion of patients are usually defined prior to initiation of the respective trials. Whenever possible, this principle should also be taken into account for preclinical studies. The predefinition of eligibility criteria as well as the unawareness of the treatment allocation, prevents investigators from a selective exclusion of animals which may as a consequence lead to false positive results. In addition to the determination of the characteristics leading to the enrolment of animals in the study, an intention-to-treatment analysis is generally preferred. This type of analysis defines the handling of missing data and outliners, and is favourable due to the analysis of all animals included in the experiments regardless of whether an exclusion of single individuals took place or not. This approach can also lead to the prevention of the misuse of statistical analysis and, consequently, to the elimination of attrition bias [53].

In addition to the main types of bias described above, it is also of enormous importance to select the right sample size of animals. The group size should, on the one hand, be large enough to provide significant results but should, on the other hand, respect legal requirements keeping the number of animals as low as possible.

III. External variability and validation of animal models

The compliance with the internal validity and the resulting elimination of bias is no guarantee for a successful translation of preclinical findings to the clinic. It should be realised and always taken into consideration, that the animals used in preclinical studies are only models for special human diseases which in most cases cannot reflect the complexity of the particular disease and therefore results in differences between the model and the clinical trials. Therefore, it seems all the more essential to find a suitable mouse model or a combination of models and to define the right scientific questions. To find an optimal combination of different models, the validity scoring system of Sams-Dodd can be applied [56]. Here five different criteria are considered, each with four options scored from 1 (low validity) to 4 (high validity).

Table 3: Validity scoring system used to find the optimal combination of animal models (adapted from [56])

Criteria	Options	Score
Complexity	Molecular	1
	Cellular	2
	Tissue	3
	In vivo	4
Disease simulation	No	1
	Pharmacological	2
	Complex	3
	True	4
Face validity	No	1
	One symptom	2
	One core symptom	3
	> 1 symptom	4
Predictability	No	1
	All or nothing for certain pharmacological principles	2
	Graded response for certain pharmacol. principles	3
	Graded response for all pharmacol. principles	4
Species	Non-mammal	1
	Non-human mammal	2
	Non-human primate	3
	Human	4

This validity scoring system allows compiling a combination of animal models which should have maximal validity and therefore can potentially come closer to the clinical situation. The first criterion, the complexity of the test system, plays an important role for the examination of the relevant mechanisms of a particular disease. With in vitro models on the cellular level (score 2), for instance, it is possible to detect simple effects on special cell types, receptors or ion channels whereas in vivo models make it possible to investigate the whole effect or cascade (score 4). For disease simulation the 'true' simulation is considered as the most appropriate choice which, unfortunately, is very rarely applicable as is the case for some infectious diseases [42]. In most cases, the disease and/or the concerning symptoms are induced by the use of special drugs. Face validity means the similarity in biological mechanisms between the animal model and the human disease. Whereas some models used make it possible to model only one of the many symptoms associated with a particular disease, other models can detect one or even more of the core symptoms. The predictability or predictive validity can be defined as the gradation of pharmacological responses. An all or nothing response is of limited informative value as a graded pharmacological response is. The last criterion, the species, considers the similarity between the model and the respective human disease. Using a non-mammal model (score 1), for instance, creates the risk of considerable differences between the pathophysiological mechanisms of the model and the human disease. The closer the model comes to humans, the greater the opportunity to develop different biological mechanisms regarding the respective disease. For some biological substances such as particular antibodies or interferons/interleukins, it is important to bear in mind that they only show cross-reactivity in non-human primates and that rodents are not suitable in this case. However, attention should be drawn to the fact that the validation has to be earmarked in order to be considered successful. The determination of similarity in biological mechanisms is far more valuable while examining pathomechanisms as is the case during the simple demonstration of predictability of animal models [42].

The use of immunodeficient mouse models allows more insight into the human immune system and cancerous diseases. Most of these models only have a very small number of B- and T-lymphocytes which are essential tools for initiation of an immune response. Without a functional immune system immunodeficient mice, for instance, are able to receive human tissue or tumour xenografts without any graft rejection. The immunodeficiency of these mouse models contemporaneously constitutes a significant disadvantage. Without a functional immune system, the translation to immunocompetent patients as well as the assessment of the results received may be significantly hampered.

Further obstacles for a successful translation from animals to humans may be marked species differences in sensitivity regarding special substances. Some substances, for instance, peroxisome proliferators like phthalate ester plasticizer or some pesticides show a carcinogenic potential in rodents but non-responsiveness in humans. After chronic administration tumour induction was predominately seen in rats and mice whereas other rodents like guinea pigs as well as monkeys and humans appear to be relative insensitive to these substance class. In this special case, these differences could be explained by an increased incidence of the peroxisome proliferator-activated receptor alpha (PPARalpha) in rodent liver cells compared to humans. The strong activation of PPARalpha by peroxisome proliferators in rodents can lead to liver growth and, finally, also to carcinogenesis [57]. In addition to the peroxisome proliferators, other substances like dopamine agonists and antagonists or β-agonist bronchodilators are also well known as rodent carcinogens, whereas humans show an insensitivity or even non-responsiveness when exposed to these substances [27].

Another common cause of reduced external validity of preclinical studies can be the ignorance of other differences in age, gender or even health status. Diseases such as Alzheimer's or Parkinson's disease mainly occur in elderly people. Therefore, young and healthy animals are not entirely suitable for the induction and investigation of the respective disease. Diseases affecting both genders are frequently tested in only one gender in rodents due to ethical reasons. Furthermore, diseases and injuries like burn wounds (e.g. for the investigation of wound healing) under study are often induced with insufficient similarity to the real conditions in humans. This may cause an incorrect translation from preclinical data to the clinic. Differences in starting the treatment or the application of doses which show satisfying results in animals but are toxic in humans, are other obstacles which need to be overcome [53].

For the reasons given above, it often makes sense to confirm and substantiate positive findings found in one animal model by subsequent studies with other slightly different animal models.

But regardless of how sophisticated an animal model is - the study design is of at least the same importance to receive convincing results. In the following, recurrent differences found between preclinical studies and clinical trials are described.

Comparison of the study design

The prediction of substance characteristics by comparing findings of preclinical studies and clinical trials is frequently restricted due to differences in the conduction and analyses of both study types. Examples for potential differences and methodological problems between preclinical studies and clinical trials were compiled by Hartung, 2013 [44]. An excerpt of some important differences can be found in the following table:

Table 4: Differences in characteristics found in both preclinical studies and clinical trials [44,58,59]

	Preclinical Studies	Clinical Trials
Subjects	(often) one gender young animals small groups often homogeneous genetic background (inbred)	both gender patients of all ages large groups heterogeneous background
Disease models	Acute model artificial diseases (e.g. Parkinson's disease) (often) immunodeficient knock-out mice	Chronic disease Acquired illness Active immune system
Study design	Tools used in clinical trials often not existing	Randomisation, Blinding (Double-Blinding), placebo-controlled, reporting of loss
Handling	Optimal circumstances (housing, nutrition) Stressed animals	Variable everyday life

Whereas in clinical trials several thousand individuals are included in the study, and in most of the cases only one endpoint is determined, in preclinical studies only a minimal number of individuals are used primarily due to ethical reasons and often a large range of endpoints is examined [44]. Furthermore, the individuals used in preclinical studies are mostly inbred strains of the same age and potentially the same sex and therefore, cannot reflect the wide spectrum of human diversity. And whereas patients taking part in clinical trials are constantly exposed to different disturbing factors of the environment, the animals in preclinical studies

are given optimal circumstances in housing or nutrition [44]. If a surgical intervention becomes necessary care should be taken to designate one responsible person to have as little variability as possible.

Generally, during the conduction of preclinical studies the efficacy and the safety of the substance under investigation are assessed in different experiments. The efficacy of a substance is normally investigated in animal disease models whereas the safety is usually examined in healthy animals. But monitoring the safety in healthy animals is often not expedient due to less sensitivity for potential side effects. Disregarding side effects in the disease models on the other hand, may lead to overestimation of the efficacy and therefore, to the production and interpretation of false-positive results. In consequence, the investigation of safety issues should, whenever possible, be included in animal efficacy studies [42].

IV. Public bias

Nowadays it has unfortunately become an internationally accepted standard in all fields of research to only deliver excellent results. Conferences and publications are mainly used to advance the own position instead of really discussing the potential issues emerging during the studies conducted. Pressure keeps rising to deliver positive results in order to receive further funding or to be able to publish the results in journals of high impact. Publication in such top-tier journals in turn is the driving force for success and many careers of investigators depend on the number, as well as the impact of published articles. All these incentives can lead to public bias expressed in reporting only selected studies and potentially overestimating certain treatment effects [53].

Moreover, most editorial boards of scientific journals are not willing to accept and publish negative (preclinical) data and there are rarely opportunities to present such negative findings at scientific conferences or meetings. Frequently, journal editors and reviewers unintentionally encourage public bias by looking for the ‘perfect’ story with a complete data set and clear explanations of the findings. Imperfect stories with scientific gaps are not tolerated and often result in publication in journals with lower impact [60]. To publish negative findings would be an important step in order to learn from these outcomes and to possibly minimise unnecessary studies carried out by other research groups.

The awareness of public bias in the reporting of clinical studies has been studied intensively. However, the presence of public bias in preclinical studies is somewhat neglected. As already mentioned in chapter II ‘Internal variability - Design, conduct and analysis of preclinical studies’, scientists from different companies like Bayer HealthCare attempted to determine

the reproducibility of preclinical key findings by comparing the published data with their own established findings. The reproducibility of research findings in many instances leaves much to be desired. Many of the non-reproducible preclinical studies formed the basis for further publications of other research groups and few of them even led to the implementation of clinical studies [60].

Reliability of preclinical studies – Discussion and conclusion

Preclinical examinations cover various kinds of toxicological and pharmacological in vitro and in vivo studies. Toxicological safety studies try to figure out toxic effects of the compound to be evaluated on potential target cells and organs. Furthermore, parameters for the monitoring of potential side effects and the reversibility of the mode of action are examined in greater detail. The essential characterisation of the mode of action and the metabolism of a compound (pharmacodynamic and pharmacokinetic properties) on the other hand are the main aims of pharmacological in vitro and in vivo studies. The evaluation of all data collected should enable an extrapolation from animals to humans as well as establishing a safe initial dose for FIM studies in clinical Phase I trials. Preclinical studies are conducted before and even during the clinical trials to ensure a continuous monitoring.

Well-conceived preclinical studies are a critical element of translational research and the subsequent successful marketing authorisation of a new medicinal product, thus making new and urgently needed medicinal products available to patients. In earlier days several therapeutic disasters occurred which particularly highlight the importance of preclinical research. In 1962, inadequate safety testing, probably due to the absence of regulations in this field, lead to an increase in teratogenic deformities of newborns after thousands of women took thalidomide marketed as Contergan® by the pharmaceutical company Chemie Grünenthal (now Grünenthal GmbH) for reducing morning sickness during pregnancy. The use of new chemical entities during pregnancy was not strictly controlled and medicinal products were not attentively tested for potential negative effects for the foetus, probably because of the assumption that the placenta barrier could not be passed [61]. Too little scientific knowledge and, in consequence, a lack of preclinical studies on pregnant women resulted in a failure of translational research. The situation is different, however, in case of TeGenero. The TeGenero AG was a biotechnological company in Würzburg which developed the new CD28 superagonist antibody TGN1412. This humanised antibody was capable of activating T-lymphocytes and was treated as a promising candidate for the treatment of T-cell deficiencies. It differed from ‘normal’ antibodies only in the epitope-binding site. Preclinical studies in rodents with previous antibody variants as well as with TGN1412 itself, could demonstrate safety and efficacy of the class of superagonists. Even repeated dose toxicity studies conducted in non-human primates cynomolgus and rhesus monkeys had not disclosed any inconsistencies. After conduction of all necessary preclinical studies and collection of a large amount of data, the FIM dose was calculated on the basis of the ‘Minimal Anticipated Biological Effect Level (MABEL) approach. A low dose of TGN1412 was administered

simultaneously to six healthy human volunteers in a clinical Phase I trial. Only a short time after the application, all six men started to suffer from severe adverse effects triggered by a 'cytokine storm'. Besides reconsidering clinical regulations and guidance, this unforeseen outcome once more reflects the difficulties of translational research and the presence of hurdles which frequently have to be overcome [62,63]. Therefore, it is of great importance that basic rules are obeyed in preclinical research ensuring increased prospects of success in the translation from cells and animals to humans in clinical trials.

Shortcomings and potential solutions during implementation

At the beginning of preclinical examination, the first hurdle which needs to be overcome involves potential shortcomings in the in vitro testing of a compound. Intra-laboratory cell line heterogeneity and a lack of quality control and quality assurance are major issues that can be observed. Differences in the handling and passaging of immortalised cell lines and primary cells frequently lead to the formation of potential sub-populations with less molecular complexities and possible changes in genotypic and phenotypic characteristics [44]. Incorrect performances of in vitro tests, often due to less experience of the respective investigators, further reduce the quality of the studies. Investigator bias and a potential lack of scientific rigor play important roles concerning the poor predictability of a lot of preclinical studies. Mycoplasma infections and cell line cross-contaminations are frequent companions in cell culture and continuous morphological similarities or unawareness of the research groups in many cases, prevent a thorough analysis of the cell lines before their use in basic research and preclinical studies. An essential approach for improving the in vitro testing may be the standardisation of the cell culture practice across different research groups. This could possibly be achieved by definition of special guidance like the Good Laboratory Practice (GLP). Whilst GLP gives only limited guidance for in vitro testing, in 1996 T. Hartung (CAAT-Europe), for instance, started a promising initiative toward Good Cell Culture Practice (GCCP) resulting in a GCCP guidance drawn up in 2005 [44,64]. The standardisation of workflows, the validation of reagents and analytical methods, as well as the awareness of the researchers regarding the need for continuous cell line authentication and periodically checking cells for contaminations is an initial step for optimising the translational process.

Equally important, however, is an adequate training of the investigatory staff because a study can only be as good as the investigator who carries it out. Therefore, adequate training and supervision of scientists new in the field of preclinical research are essential requirements. Correct performances and interpretation of study findings form the foundation of successful

translational research but variability and bias are main challenges during the design, conduct and analysis of preclinical studies. Systematic evaluations of the quality of animal studies were carried out and demonstrated that even highly cited animal research gave hints about weaknesses in quality [65–67]. Methods like randomisation and blinding were rarely reported and eligibility criteria (characteristics for inclusion and exclusion of animals) or intention-to-treat analysis were seldom applied. Furthermore, the repetition of experiments did not take place and statistical tests were not appropriately used [68]. Studies ignoring those tools showed greater differences between the study groups examined and had an almost five times higher probability of reporting positive findings. Nevertheless, even in case of high impact publication, only one third of the preclinical studies could be translated to the clinic [67].

Possible improvements in preclinical research may be achieved by adopting standards similar to those applied in clinical trials [45,53]. Apart from the biases also known to exist in clinical research, the quality standards used in clinical trials are very high. Clinical trials have to be authorised by competent authorities and require ethical review. Furthermore, they are conducted under GCP conditions and carried out by trained health care professionals. Randomisation, blinding and placebo controls are permanent components realised in most of the clinical trials conducted [44]. Clinical trials are often conducted in a multicenter approach whereas preclinical studies are frequently performed by only one research group in one laboratory. The multicenter approach is another factor which helps to save transparent quality control. Therefore, compliance with clinical standards may be an important step on the path towards effective translational research and robust preclinical findings. Another significant aspect is to get feedback from ongoing clinical trials. This may help matching preclinical studies with clinical findings and therefore, improve the translational process. In this context the choice of endpoints can be another significant aspect. Endpoints in preclinical studies often differ significantly from the endpoints requested for clinical trials. These differences may be due to ethical reasons. While the overall survival is a commonly used endpoint in human clinical trials, its application in most preclinical studies is not suitable. Instead, for preclinical cancer models the tumour size is often the crucial parameter and primary endpoint. This endpoint clearly permits the distinction between health and disease [42].

Furthermore, health care professionals should be encouraged to design and conduct clinical trials similar to the preclinical studies previously performed. This replication of condition under which efficacy has been observed in animal models might also lead to an improvement in the translational process [53]. In this context phase 0 trials should be mentioned as an important tool. These clinical trials defined as exploratory IND (investigational new drug)

studies are recommended by the FDA and conducted prior to phase I trials [69]. They may help to gather essential information about the further course of action and optimise the design of subsequent phase I trials. Micro-doses are administered to patients in order to perform intensive pharmacodynamic and pharmacokinetic analyses. The data collected may help to better interpret findings of potential 'imperfect' preclinical animal models and to prevent misleading of the investigators [70,71].

Need for collective consciousness

All participants (investigators, publisher, clinical staff, etc.) should be aware of the fact that the animal models used in preclinical research will never represent the real state of human diseases. First of all, in preclinical studies it is a frequent practice to use young animals of only one gender and a homogenous genetic background. This does not reflect reality as patients taking part in clinical trials are often a heterogeneous group of young and old, women and men. Diseases like Parkinson's, Alzheimer's or cancer are frequently generated by administration of pharmacological substances or injection of immortalised cell lines, respectively. The complexity of these diseases and the related immune reactions are scarcely reproducible. Inter and intra-species pharmacokinetic variability is in many cases the main reason for failure in the translational process and the transfer of research findings into the clinic [2]. Different absorption of the compound under investigation, a very short or long half-life, possible enzyme inhibitions or other differences in biological processes should be taken into consideration not to cause unnecessary costs and delays in drug development. Further differences might be possible co-morbidities from which (elderly) people might often suffer. Co-medication is necessary which, in turn, can potentially affect the therapy of the respective disease for which the treatment is intended. The therapeutic intervention is another crucial point which needs to be considered and which should not be earlier or later than in humans. As a matter of fact, this cannot always be adopted as can easily be illustrated, for instance, using the example of Parkinson's. Whereas in humans it is indicated as an acquired disease and takes years to develop, in animals it is artificially induced by pharmacological substances and instantaneously treated.

The determination of the FIM starting dose can also be a critical parameter. In some cases high doses can have an effect on the motor activity of animals, and at the same time frequently having an influence on their behaviour. Unawareness and inexperience of the investigator could therefore easily lead to false (positive) findings. Another hurdle for a successful translation from preclinical studies to the clinic may be species differences in

responding to special substances. As already mentioned in chapter III ‘External variability and validation of animal models’, peroxisome proliferators like phthalate ester plasticizer or some pesticides, for example, show a carcinogenic potential in rodents but non-responsiveness in humans. Other substances like dopamine agonists and antagonists or β-agonist bronchodilators are also declared as rodent carcinogens while when administering these substances to humans insensitivity or even non-responsiveness is shown.

Due to the mentioned differences between animal models and humans and other forms of behaviour which can occur, it is well advised to always confirm and reproduce positive findings of preclinical studies by means of subsequent examinations with other valid in vitro tests or animal models. Especially in case of critical substance classes where the opportunity of different behaviours in animals and humans exists, pharmacokinetic and pharmacodynamic parameters should be given particular attention. If, for instance, the molecular pathway for a molecule is the same in both species but there is a disparity in the pharmacokinetic and pharmacodynamic, the use in humans should be reconsidered and further information should be collected by means of an intensive literature research. Besides weaknesses in quality of study design and conduction, many published findings are only based on one single study with one animal model and performed in one laboratory. This approach hardly provides meaningful results and it is therefore even more important for scientists, physicians or other persons relying on study findings to be sceptical and always evaluate and validate these critical before building on these findings in connection with further research or clinical trials.

Resolving a systematic bias

Apart from weaknesses in the design, conduction and analysis of preclinical studies and reduced external validity due to disparities between animal models and humans, a systematic problem exists which is assumed to play a major role regarding the reliability of preclinical studies. Nowadays, the number and the impact of publications are considered to be key drivers for a successful career in science. Publications and conferences which should normally offer a platform for discussions concerning potential issues raised during the conduction of the respective preclinical studies are now mainly used to improve the image and increase the own awareness level. The lack of reliability and reproducibility concerning preclinical studies and their publications are mainly due to the fact that the publication system is arranged in a way as to publish as fast and as many publications as possible, preferably searching for the ‘perfect story’. By looking for excellent results, publication bias is encouraged and inevitably

leads to an overestimation of treatment findings. During publication processes, i.e. in many instances original data is removed whilst positive data sets are emphasized.

The system has to change in a way as journal editors and peer reviewers should only support publications of validated preclinical studies conducted and analysed with high quality standards as used in clinical trials. Tolerating and giving incentives for fast publications of excellent stories is frequently combined with poor quality and non-validated methods or analyses. Furthermore, attention should also be given to negative outcomes as learning from these findings can possibly minimise unnecessary in vitro and animal studies. Investigators ought to have the possibility for publishing negative findings, even in high-impact journals. Regardless of the outcome of preclinical research, to keep preclinical findings from the public is unethical in two different ways. Animals are used in vain and researchers are deprived of data which might give important insights for the conduction of future studies.

Many approaches of confirming and reproducing research findings, some of these even having triggered clinical trials, lead to disappointing results. Bayer HealthCare, for instance, examined findings of primary and secondary publications, only 25 % of which could be confirmed and validated. It could be observed that studies for which preclinical research results could be confirmed, controls, validating reagents, analytical methods and publishing the whole data sets played an important role for investigators. Not reproducible studies frequently indicate shortcomings like non-blinded investigators, not validated methods or using only one (mouse) model [60].

Publication bias in clinical literature has already been reduced by the establishment of clinical trial portals like ClinicalTrials.gov or Clinicaltrialsregister.eu. The central registration of clinical trials offers the advantage of having access to all important (positive and negative) findings, regardless of their publication [53]. In case of preclinical studies, the foundation of a central register summarising research findings might therefore be another significant step for improving the translational process.

Further improvement suggestions

The type of preclinical model used for gaining new insights into the safety and efficacy of a new chemical drug is crucial for the robustness of the data obtained. In vitro models are less informative than in vivo models due to the missing opportunity to investigate the complex effects and cascades taking place in the human body. With in silico preclinical studies there is another possibility of gaining more insights into the functionality of newly developed molecules or substances. This computer simulation-based testing methodology has become an

important element of preclinical proof-of-concept testing and has already been accepted as a substitute to in vivo studies, for instance, for examinations concerning type 1 diabetes mellitus. With computer-based simulations, investigators have the ability to recreate extreme scenarios which cannot be examined in animal studies mainly due to ethical reasons. Therefore, in silico testing can help to further close the translational gap [72,73].

Back to in vivo models, the immunodeficiency of most of the mouse models used in e.g. cancer studies, constitute a significant disadvantage as the translation to immunocompetent patients is hampered and immune response can only be partly investigated. An interesting alternative to immunodeficient models are genetically engineered mouse models (GEMM). Through the use of genetic engineered techniques like molecular cloning or gene targeting, transgenic mice are generated which, for instance, are carrying cloned oncogenes or lacking tumour suppressor genes and have been proven to be successful models for human cancer. In contrast to the immunocompetent models, tumours developed in GEMM are in an immunocompetent environment. Although they might be a better approach for the reliability of preclinical studies compared with the immunodeficient models, it still has restricted validity due to the fact that the limited genetic alterations introduced do not reflect the frequently occurring multiple abnormalities in human cancer [45,74,75].

The development of humanised mouse models is considered to be another important tool for improving the clinical translation process and develop personalised medicines [42]. Over the last 10 years, remarkable progress has been achieved in this field causing a variety of humanised mouse models. Mutations in already highly immunodeficient strains like the introduction of the mutant IL2ry gene into NOD/SCID mice leads to defects in special immune cells like T and B-lymphocytes and therefore enhances the potential differentiation of human hematopoietic cells after transplantation of human hematopoietic stem cells (HSC). The presence of human immune cells facilitates the investigation of potential mechanisms concerning human pathogenesis [76]. Some of the humanised models, also known as patient-derived tumour xenograft models (PDTX) serve as cancer models and can be utilised in order to engraft human primary tumour cell derived as well as patient derived xenografts. Cytotoxic drugs can be applied in clinically equivalent doses and may therefore be more predictive for efficacy as it would be the case for mouse-derived tumours [42,76]. Possible disadvantages can be the inadequate imitation of the human immune components, the long latency of tumour development or the low engraftment rates of different tumour types. Nevertheless, humanised mouse models constitute an important element in cancer research and may pave the way to personalised medicine for cancer patients and even other diseases.

In addition to well thought-out and validated animal models, biomarkers may be another interesting feature which deserves special attention. The National Institute of Health defines biomarkers as follows: *'A characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacological responses to a therapeutic intervention.'* One important goal of toxicity studies is the identification of potential target organs which are toxicologically influenced by the administered substance. Apart from identifying possible adverse responses of the drug, the safety expert has also to define the exposure at which an effect is observed in the targeted organs. Due to limited auxiliaries in precisely monitoring toxicity in preclinical studies as well as in subsequent clinical trials, extensive research was carried out in the last years to find suitable tools for this issue. The identification and characterisation of biomarkers for detecting organ-specific toxicity has marked a decisive step forward in the reliability of preclinical studies [77]. Biomarkers can have miscellaneous functions. Predictive toxicological biomarkers in cancer research may, for instance, be able to classify tumour responsiveness by determining the regression or progression of a tumour. In addition to the mentioned toxicological biomarkers, there is a range of different other biomarkers or commercial biomarkers assay kits. Those worth particular mention here are biomarkers of the immune system (e.g. cytokines like IL-1 or TNF-α), allowing determining endpoints involving leukocytosis or leukopenia, biomarkers for inflammation (e.g. acute phase proteins like fibrinogen or albumin) or renal biomarkers measuring functional or structural characteristics. These biomarkers and further pharmacogenomic, proteomic, pharmacological or imaging biomarkers can be helpful tools intended to determine the safety or efficacy of a drug and to promote the translational research from preclinical studies to subsequent clinical trials in humans.

Summary

The aim of this master thesis is to give an overview of the drug development process with a particular focus on preclinical studies and their reliability concerning the translation from animal testing to clinical trials in humans. The thesis addresses the potential obstacles and biases which can occur during the translational process from preclinical research to clinical trials and, in addition, discusses possible improvement strategies.

Drug development can commonly be divided into four main pillars: Drug discovery, preclinical development, clinical studies and marketing authorisation of a newly found chemical entity. During the phase of drug discovery potential new targets for special diseases are identified and validated. After finding initial hit substances they are optimised in order to obtain so-called lead compounds. The most successful lead compounds found during the drug discovery program are then investigated in preclinical in vitro and in vivo studies. The main objective of the considerable number of preclinical studies is getting a better understanding of the toxicological profile of an individually tested substance. With the knowledge of toxic effects and pharmacological properties of the respective compound, it is attempted to establish a safe initial dose for first in man exposure in clinical Phase I trials. Preclinical studies are conducted before and even during clinical studies in order to ensure continuous monitoring. The most promising candidates determined during preclinical research are tested in clinical trials in order to analyse their human compatibility. The conduction of Phase I to Phase III clinical trials is a mandatory tool in order to have the opportunity to apply for a marketing authorisation for new medicinal products. Obtaining a marketing authorisation is necessary to place a medicinal product on the market and thereby make it accessible to the general public.

Well-conceived preclinical studies are a critical element of translational research and the subsequent successful marketing authorisation of a new medicinal product, thus making new, urgently needed medicinal products available to patients. However, this translational process may frequently be riddled with various hurdles which need to be overcome. Shortcomings in in vitro analyses of a compound like intra-laboratory cell line heterogeneity or a lack of quality control and quality assurance may be major issues. Furthermore, investigator bias and a potential lack of scientific rigor play important roles concerning the poor predictability of a lot of preclinical studies. Special guidance like the Good Laboratory Practice (GLP), adequate training and the supervision of scientists should therefore be essential requirements in order to receive validity of the study results. Potential weaknesses in the design, conduction and analysis of preclinical studies may be overcome by adopting standards similar to those applied in clinical trials, e.g. realising tools like randomisation and blinding or having placebo

controls. Another important point all participants (investigators, publisher, clinical staff, etc.) should be aware of, is the fact that the animals used in preclinical studies are only models for special human diseases which in most cases cannot reflect the complexity of the particular disease. Special validity scoring systems allow compilation of a combination of animal models and may reach maximal validity, therefore coming closer to the clinical situation. Apart from weaknesses in the design, conduction and analysis of preclinical studies and reduced external validity due to disparities between animal models and humans, a systematic problem exists which is assumed to play a major role for the reliability of preclinical studies. Nowadays, the number and the impact of publications are considered to be key drivers for a successful career in science. By looking for excellent results, publication bias is encouraged and inevitably leads to an overestimation of treatment findings. Tolerating and giving incentives for fast publications of excellent stories is frequently combined with poor quality and non-validated methods or analyses. The system has to change in a way whereas only publications of validated preclinical studies conducted and analysed with high quality standards as used in clinical trials should be supported. Furthermore, attention should also be given to negative outcomes as learning from these findings can possibly minimise unnecessary in vitro and animal studies.

References

1. DiMasi J.A. Success rates for new drugs entering clinical testing in the United States. *Clin Pharmacol Ther* **1995**, *58*, 1–14.
2. Caldwell, J. The Role of Drug Metabolism in Drug Discovery and Development: Opportunities to Enhance Time- and Cost-efficiency. *Pharmaceutical Science* **1996**, *2*, 117–119.
3. Hughes, J.P.; Rees, S.; Kalindjian, S.B.; Philpott, K.L. Principles of early drug discovery. *British journal of pharmacology* **2011**, *162*, 1239–1249.
4. Forbes. *The truly staggering cost of inventing new drugs*. Available online: www.forbes.com/sites/matthewherper/2012/02/10/the-truly-staggering-cost-of-inventing-new-drugs/ (accessed on 7 May, 2015).
5. Arrowsmith, J. Trial watch: Phase II failures: 2008-2010. *Nature reviews. Drug discovery* **2011**, *10*, 328–329.
6. Hodges, M.; Applebe G.E. Control of the safety of drugs, 1868-1968 (Part 1). *The Pharmaceutical Journal* **1987**, *1*, 119–122.
7. Ballentine, C. Sulfanilamide Distaster - Taste of Raspberries, Taste of Death - The 1937 Elixir Sulfanilamide Incident. *FDA Consumer Magazine* **1981**.
8. U.S. Food and Drug Administration. *FDA History - Part II, The 1938 Food, Drug, and Cosmetic Act*. Available online: http://www.fda.gov/AboutFDA/WhatWeDo/History/Origin/ucm054826.htm (accessed on 4 February, 2015).
9. Ray, W.A.; Stein, C.M. Reform of drug regulation--beyond an independent drug-safety board. *The New England journal of medicine* **2006**, *354*, 194–201.
10. European Parliament and Council of the European Union. *Directive 2001/83/EC*, 2001.
11. Brodniewicz, T.; Grynkiewicz, G. Preclinical drug development. *Acta poloniae pharmaceutica* **2010**, *67*, 578–585.
12. Simon, N.G. Drug Discovery and Preclinical Development.
13. Ruppert, T. *Wie entsteht ein neues Medikament?*; 2010. Available online: https://www.vfa.de/download/wie-entsteht-ein-neues-medikament.pdf (accessed on 30 January, 2015).
14. Durham, T.B.; Blanco, M.-J. Target Engagement in Lead Generation. *Bioorganic & medicinal chemistry letters* **2015**.
15. Keseru, G.M.; Makara, G.M. Hit discovery and hit-to-lead approaches. *Drug discovery today* **2006**, *11*, 741–748.

16. Wikipedia, The Free Encyclopedia. *Hit to Lead*. Available online: http://en.wikipedia.org/wiki/Hit_to_lead (accessed on 6 February, 2015).
17. Anderson, A.C. Structure-based functional design of drugs: from target to lead compound. *Methods in molecular biology (Clifton, N.J.)* **2012**, *823*, 359–366.
18. Front Matter. In *Pharmacokinetic Optimization in Drug Research*; Testa, B., van de Waterbeemd, H., Folkers, G., Guy, R., Eds.: Verlag Helvetica Chimica Acta: Zürich, 2001, pp. I.
19. Testa, B.; van de Waterbeemd, H.; Folkers, G.; Guy, R., Eds. *Pharmacokinetic Optimization in Drug Research*; Verlag Helvetica Chimica Acta: Zürich, 2001.
20. Bode, G. *ICH M3 (R) Guideline*: *Preclinical Support of Clinical Trials and Marketing*.
21. Keutz, E. von. *Pharmacological-toxicological Inverstigations*: *General Principles of Toxicological Studies*. DGRA, Master of Drug Regulatory Affairs (MDRA 15).
22. International Conference on Harmonisation (ICH). *ICH M3 (R2) Guidance on Nonclinical Safety Studies for the Conduct of Human Clinical Trials ans Marketing Authorization for Pharmaceuticals*, 2nd ed, 2009.
23. Kramer, P.J. *Pharmacological-Toxicological Studies - Single Dose/Repeated Dose Toxicity*.
24. International Conference on Harmonisation (ICH). *ICH S2(R1) Guidance on Genotoxic Testing and Data Interpretion for Pharmaceuticals intended for Human Use*, 2011.
25. Wikipedia, The Free Encyclopedia. *Ames test*. Available online: http://en.wikipedia.org/wiki/Ames_test (accessed on 13 April, 2015).
26. International Conference on Harmonisation (ICH). *ICH S1A Guideline on the Need for Carcinogenicity Studies of Pharmaceuticals*, 1995.
27. bode, g.d. *DGRA / Toxicology, Cancerogenicity*, 2014.
28. International Conference on Harmonisation (ICH). *ICH S1B Testing for Carcinogenicity of Pharmaceuticals*, 1997.
29. International Conference on Harmonisation (ICH). *S5(R2) Detection of Toxicity to Reproduction for Medicinal Products & Toxicity to Male Fertility*, 1993 (accessed on Nov. 2005).
30. Keutz, E. von. *Pharmacological-toxicological Investigations*: *Local Tolerance/Phototoxicity*. DGRA, Master of Drug Regulatory Affairs (MDRA 15).
31. Keutz, E. von. *Pharmacological-toxicological Investigations*: *Pharmacology*. DGRA, Master of Drug Regulatory Affairs (MDRA 15).

32. International Conference on Harmonisation (ICH). *Safety pharmacology studies for human pharmaceuticals S7A*, 2000.
33. Bode, G. *Pharmaco-Kinetics, Toxico-Kinetics, Metabolism*.
34. International Conference on Harmonisation (ICH). *ICH S3B Guidance for Repeated Dose Tissue Distribution Studies*, 1994.
35. International Conference on Harmonisation (ICH). *ICH S3A Note for Guidance on Toxicokinetics: The Assessment of Systemic Exposure in Toxicity Studies*, 1994.
36. European Parliament and Council of the European Union. *Clinical Trials Directive 2001/20/EC*, 2001.
37. European Commission. *NtA: Volume 2A Procedures for marketing authorisation Chapter 1 Marketing authorisation*, 3rd ed.
38. European Commission. *NtA Volume 2B Medicinal Products for human use, Presentation and format of the dossier, Common Technical Document (CTD)*.
39. European Commission. *Guideline on the details of the various categories of variations to the terms of marketing authorisations for medicinal products for human use and veterinary medicinal products*.
40. Woolf, S.H. The meaning of translational research and why it matters. *JAMA* **2008**, *299*, 211–213.
41. Duke University School of Medicine. *Clinical and Translational Research*. Available online: http://medschool.duke.edu/research/clinical-and-translational-research (accessed on 2 May, 2015).
42. Denayer, T.; Stöhr, T.; van Roy, M. Animal models in translational medicine: Validation and prediction. *New Horizons in Translational Medicine* **2014**, *2*, 5–11.
43. Kaniklidis, C. The Boundaries and Limitations of Preclinical Research: A Mini-Review. *Research Gate Journal Club* **2012**, 1–7.
44. Hartung, T. Look back in anger - what clinical studies tell us about preclinical work. *ALTEX* **2013**, *30*, 275–291.
45. Loong, H.H.; Siu, L.L. Selecting the best drugs for phase I clinical development and beyond. *American Society of Clinical Oncology educational book / ASCO. American Society of Clinical Oncology. Meeting* **2013**, 469–473.
46. Nelson-Rees, W.; Daniels, D.; Flandermeyer, R. Cross-contamination of cells in culture. *Science* **1981**, *212*, 446–452.
47. Drexler, H.G.; Uphoff, C.C. Mycoplasma contamination of cell cultures: Incidence, sources, effects, detection, elimination, prevention. *Cytotechnology* **2002**, *39*, 75–90.

48. Rottem, S.; Barile, M.F. Beware of mycoplasmas. *Trends Biotechnol* **1993**, *11*, 143–151.
49. Nagaoka, R.; Iwasaki, T.; Rokutanda, N.; Takeshita, A.; Koibuchi, Y.; Horiguchi, J.; Shimokawa, N.; Iino, Y.; Morishita, Y.; Koibuchi, N. Tamoxifen activates CYP3A4 and MDR1 genes through steroid and xenobiotic receptor in breast cancer cells. *Endocrine* **2006**, *30*, 261–268.
50. Mitra, R.; Guo, Z.; Milani, M.; Mesaros, C.; Rodriguez, M.; Nguyen, J.; Luo, X.; Clarke, D.; Lamba, J.; Schuetz, E.*; et al.* CYP3A4 mediates growth of estrogen receptor-positive breast cancer cells in part by inducing nuclear translocation of phospho-Stat3 through biosynthesis of (+/-)-14,15-epoxyeicosatrienoic acid (EET). *J Biol Chem* **2011**, *286*, 17543–17559.
51. Prinz, F.; Schlange, T.; Asadullah, K. Believe it or not: how much can we rely on published data on potential drug targets? *Nat Rev Drug Discov* **2011**, *10*, 712.
52. Unger, E.F. All is not well in the world of translational research. *Journal of the American College of Cardiology* **2007**, *50*, 738–740.
53. van der Worp, H Bart; Howells, D.W.; Sena, E.S.; Porritt, M.J.; Rewell, S.; O'Collins, V.; Macleod, M.R. Can animal models of disease reliably inform human studies? *PLoS medicine* **2010**, *7*, e1000245.
54. Juni, P.; Altman, D.G.; Egger, M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *BMJ* **2001**, *323*, 42–46.
55. Altman, D.G.; Schulz, K.F.; Moher, D.; Egger, M.; Davidoff, F.; Elbourne, D.; Gotzsche, P.C.; Lang, T. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* **2001**, *134*, 663–694.
56. Sams-Dodd, F. Strategies to optimize the validity of disease models in the drug discovery process. *Drug Discov Today* **2006**, *11*, 355–363.
57. Lai, D.Y. Rodent carcinogenicity of peroxisome proliferators and issues on human relevance. *J Environ Sci Health C Environ Carcinog Ecotoxicol Rev* **2004**, *22*, 37–55.
58. Hartung, T. Food for Thought … on Animal Tests. *ALTEX* **2008**, *25 1/08*, 3–9.
59. Pound, P.; Ebrahim, S.; Sandercock, P.; Bracken, M.B.; Roberts, I. Where is the evidence that animal research benefits humans? *BMJ* **2004**, *328*, 514–517.
60. Begley, C.G.; Ellis, L.M. Drug development: Raise standards for preclinical cancer research, pp. 531–533.
61. Franks, M.E.; Macpherson, G.R.; Figg, W.D. Thalidomide. *The Lancet* **2004**, *363*, 1802–1811.
62. Attarwala, H. TGN1412: From Discovery to Disaster. *J Young Pharm* **2010**, *2*, 332–336.

63. Eastwood, D.; Findlay, L.; Poole, S.; Bird, C.; Wadhwa, M.; Moore, M.; Burns, C.; Thorpe, R.; Stebbings, R. Monoclonal antibody TGN1412 trial failure explained by species differences in CD28 expression on CD4+ effector memory T-cells. *Br J Pharmacol* **2010**, *161*, 512–526.
64. Coecke, S.; Balls, M.; Bowe, G.; Davis, J.; Gstraunthaler, G.; Hartung, T.; Hay, R.; Merten, O.-W.; Price, A.; Schechtman, L.*; et al.* Guidance on good cell culture practice. a report of the second ECVAM task force on good cell culture practice. *Altern Lab Anim* **2005**, *33*, 261–287.
65. Baker, D.; Lidster, K.; Sottomayor, A.; Amor, S. Reproducibility: Research-reporting standards fall short. *Nature* **2012**, *492*, 41.
66. Bebarta, V.; Luyten, D.; Heard, K. Emergency medicine animal research: does use of randomization and blinding affect the results? *Acad Emerg Med* **2003**, *10*, 684–687.
67. Hackam, D.G.; Redelmeier, D.A. Translation of research evidence from animals to humans. *JAMA* **2006**, *296*, 1731–1732.
68. Begley, C.G. An unappreciated challenge to oncology drug discovery: pitfalls in preclinical research. *Am Soc Clin Oncol Educ Book* **2013**, 466–468.
69. FDA, Center for Drug Evaluation and Research (CDER). Guidance for Industry, Investigators , and Reviewers - Exploratory IND Studies **2006**.
70. Kinders, R.; Parchment, R.E.; Ji, J.; Kummar, S.; Murgo, A.J.; Gutierrez, M.; Collins, J.; Rubinstein, L.; Pickeral, O.; Steinberg, S.M.*; et al.* Phase 0 clinical trials in cancer drug development: from FDA guidance to clinical practice. *Molecular interventions* **2007**, *7*, 325–334.
71. Kummar, S.; Doroshow, J.H. Phase 0 trials: expediting the development of chemoprevention agents. *Cancer prevention research (Philadelphia, Pa.)* **2011**, *4*, 288–292.
72. Patek, S.D.; Bequette, B.W.; Breton, M.; Buckingham, B.A.; Dassau, E.; Doyle, Francis J 3rd; Lum, J.; Magni, L.; Zisser, H. In silico preclinical trials: methodology and engineering guide to closed-loop control in type 1 diabetes mellitus. *J Diabetes Sci Technol* **2009**, *3*, 269–282.
73. Kovatchev, B.P.; Breton, M.; Man, C.D.; Cobelli, C. In silico preclinical trials: a proof of concept in closed-loop control of type 1 diabetes. *J Diabetes Sci Technol* **2009**, *3*, 44–55.
74. Rangarajan, A.; Weinberg, R.A. Opinion: Comparative biology of mouse versus human cells: modelling human cancer in mice. *Nat Rev Cancer* **2003**, *3*, 952–959.

75. Sharpless, N.E.; Depinho, R.A. The mighty mouse: genetically engineered mouse models in cancer drug development. *Nat Rev Drug Discov* **2006**, *5*, 741–754.
76. Ito, R.; Takahashi, T.; Katano, I.; Ito, M. Current advances in humanized mouse models. *Cellular & molecular immunology* **2012**, *9*, 208–214.
77. Campion, S.; Aubrecht, J.; Boekelheide, K.; Brewster, D.W.; Vaidya, V.S.; Anderson, L.; Burt, D.; Dere, E.; Hwang, K.; Pacheco, S.*; et al.* The current status of biomarkers for predicting toxicity. *Expert Opin Drug Metab Toxicol* **2013**, *9*.

Declaration

Erklärung

Hiermit erkläre ich an Eides statt, die Arbeit selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel verwendet zu haben.

Dortmund, 06. August 2015

_____

Dr. Jennifer Anthöfer