Toxicology and Adverse Drug Reactions

D. J. Snodin

Introduction

Toxicology has two main goals in relation to adverse drug reactions (ADRs). The first is to identify and characterize the potential for harmful effects that can be produced in biological systems, particularly laboratory animals, by a drug, and to suggest therapeutic circumstances in which toxic responses may occur and/or are unlikely to occur. The second is necessary when unexpected adverse clinical reactions are detected, i.e. those not predicted by conventional animal and clinical studies, and to investigate the mechanism in additional toxicological studies, often of nonstandard design, in order to understand these reactions and possibly how to avoid or ameliorate them.

Toxicity testing

Pharmacotoxicological tests

The term ‘pharmacotoxicology’ is often used to describe the experimental study of pharmacodynamic and toxicological effects of the ingredients of medicinal products. Toxicological testing of pharmaceuticals employs basic concepts, laboratory animal species, study types and designs that are similar to those used in other industrial sectors, such as chemicals and food ingredients, but there are several special features.

The intended biological activity of test materials has a number of consequences, such as:

- Selection of an appropriate species that is pharmacologically responsive, but in which responses reducing the effectiveness of a particular model (such as the production of vaccines (may contain adjuvants and preservatives), topical products and modified-release formulations, where excipients may alter the pharmacotoxicological responses.

---

1 Most tests are undertaken on active ingredients rather than the formulated medicinal product. Exceptions include vaccines (may contain adjuvants and preservatives), topical products and modified-release formulations, where excipients may alter the pharmacotoxicological responses.
neutralizing antibodies following administration of human-specific proteins) are minimized.

- Occurrence of both pharmacological effects and toxicological effects; the detection of toxicological effects can often be confounded by exaggerated pharmacodynamic responses.

Human data from clinical trials, not normally available for non-pharmaceuticals, for the most part supersede the results of animal studies, except in the case of those endpoints (such as genotoxicity, carcinogenicity and reproductive toxicity) where it is impractical and/or unethical to undertake human studies. An important role of toxicological studies during drug development is to provide sufficient safety data to evaluate the risk to patients participating in a particular clinical trial. Thus, the timing of studies is closely related to the key clinical elements (e.g. phase 1, 2 and 3 trials) of the clinical development programme. Toxicokinetic measurements are performed to enable comparison of systemic exposure in animals with that in patients who are exposed to the drug. This may not always be practicable, e.g. with topically applied drugs where systemic exposure is often negligible.

The nature and purposes of the principal non-clinical studies normally required for a new conventional (chemical) pharmaceutical active ingredient (new chemical entity; NCE) are described in Tables 3.1 and 3.2.

**Good laboratory practice**

All safety studies must be performed in compliance with GLP – a guiding set of principles with the aim of ensuring that laboratories design, perform and report all safety studies on pharmaceuticals (and other materials, such as industrial chemicals) carefully and document all activities in such a way that studies can be reconstructed at any time afterwards (Könemann, 1990; Hawkins, 1993; Department of Health, 2000; CFR, 2003). Many aspects of laboratory activities can influence the results produced and their subsequent interpretation, and so competent authorities in the major industrial countries (as well as organizations such as OECD: www.sourceoecd.org/content/templates/co/co_main_oecdguid.htm?comm=oecdguid) have promulgated GLP regulatory guidance documents. The role of GLP regulations (in the UK these are ‘The Good Laboratory Practice Regulations, 1999’) is to codify the components of GLP, the principal ones being:

- test facility organisation and personnel
- quality assurance programme
- facilities
- apparatus, materials and reagents
- test systems
- test and reference items
- standard operating procedures
- performance of the regulatory study
• reporting of regulatory study results
• storage and retention of records and materials
• inspections
• study audits.

Animal welfare

Strict controls exist over the use and welfare of experimental animals. The principal UK legislation is ‘The Animals (Scientific Procedures) Act 1986’. This controls experimental and other scientific work carried out on living animals that may cause pain, suffering or other distress to the animals. Both project and personal (i.e. the experimenter) licences, as well as a certificate of designation relating to the place where the work is undertaken, are required under the act and a variety of codes of practice regarding housing and care and humane killing, etc. have been published (Home Office, 2002). Similar provisions for the maintenance of animal welfare apply in other countries (European Biomedical Research Association www.ebra.org/; FRAME www.frame.org.uk/). GLP requirements also impact on animal welfare in respect of ‘support facilities and conditions for their (test animals) care, housing and containment which are adequate to prevent stress and other problems which could affect the test system and hence the quality of the data’ (Department of Health, 2000).

Determination of toxic potential

The overall aim of the non-clinical tests is to determine the potential for toxic reactions by examining a variety of endpoints that may be affected:

• Functional or dynamic. For example, a potentially adverse change in blood pressure or cardiac function (typically evaluated in safety pharmacology studies).

• Biochemical. For instance, a change in the concentration of a serum enzyme such as aspartyl aminotransferase (AST) indicating liver damage. On the other hand, direct (and possibly intended) enzyme inhibition (e.g. acetyl cholinesterase inhibition) may also occur.

• Haematological. For example, a treatment-related reduction in haematocrit indicating anaemia, or changes in lymphocytes suggesting immunological effects or a response to inflammation or infection.

• Structural. For example, pathological changes in organ weight and/or structure, such as liver hypertrophy and/or necrosis.

• Behavioural. Drug-related behavioural dysfunction, in most cases not obviously correlated with specific deficits in nervous structure or function.

• Developmental. For example, reduced body weight gain not accompanied by reduced food consumption often indicates a toxic response; impaired foetal development may be associated with foetal abnormalities.
<table>
<thead>
<tr>
<th>Test type</th>
<th>Test system</th>
<th>Results/evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safety pharmacology</td>
<td>Standard pharmacological procedures on organ system(s) such as cardiovascular, respiratory, renal, urinary and central nervous system (CNS)</td>
<td>Evaluation of any functional effect at a range of single doses</td>
</tr>
<tr>
<td>Acute toxicity</td>
<td>Limit test using single high dose (or a range of doses) by therapeutic route (plus parenteral route for orally administered drugs) in rodent and non-rodent; observation period normally 14 days</td>
<td>Clinical responses to treatment (e.g. lethargy, prostration), mortality and necropsy (on decedants and survivors), providing an early gross indication of organ systems likely to be affected by toxicity</td>
</tr>
<tr>
<td>Repeated-dose toxicity (sub-acute to chronic)</td>
<td>Administration of drug at three dose levels (plus control group) using intended therapeutic route usually over 2–26 weeks in rat and 2–39 weeks in appropriate non-rodent species (normally the dog)</td>
<td>Variety of clinical observations, body weight, food consumption, haematology, clinical chemistry, macro- and micro-pathology. ECG monitoring sometimes included in non-rodent. Data evaluated to assess target organ(s) for toxicity, dose-exposure-response relationships and reversibility</td>
</tr>
<tr>
<td>Genetic toxicity</td>
<td>Standard three-test battery for gene mutation in bacteria (Ames test), chromosome aberrations or mutations in mammalian cells in vitro and in vivo cytogenetic test (normally rodent bone-marrow micronucleus assay). In vitro assays performed in the absence and presence of induced rat liver S9 microsomal fraction as exogenous metabolizing system</td>
<td>Evaluation of DNA damage producing effects at the level of the gene or chromosome (clastogenicity)</td>
</tr>
</tbody>
</table>
Carcinogenicity

Control group plus three dose levels of drug using intended therapeutic route in rat for 24 months. Second study in mouse, either conventional 24-month bioassay or study in an acceptable alternative model (e.g. 6-month study in p53+/− mouse). Blood samples taken for measurement of drug plasma concentration. Adequate survival and demonstration that >25-fold human exposure or maximum tolerated dose (MTD) was achieved to ensure a valid test. Incidence, dose–response and statistical analysis of organ-specific tumours assessed, distinguishing different tumour types and metastases. Distinction between genotoxic and nongenotoxic carcinogens (tumour profile, threshold, mechanistic studies, etc.).

Toxicity to reproduction

Segment I study for fertility and general reproductive performance in the rat. Segment II studies in rat and rabbit for embryotoxicity/teratogenicity and Segment III study in the rat for peri-/post-natal toxicity. Tests can be combined (e.g. I/II) if appropriate. Three dose levels and control group in each test. Blood samples taken from pregnant or non-pregnant animals for measurement of drug plasma concentration. Determination of mating, reproductive parameters, fetal development, pup skeletal and soft-tissue abnormalities, nursing behaviour, postnatal development and pup survival. Evaluation of dose-response relationships and maternal and foetal no-effect levels.

Local tolerance

Particularly for drugs administered parenterally (e.g. i.v., s.c.) or topically, tests (e.g. in rabbit ear model) to assess reactions of adjacent tissues. Determination of nature, severity, dose–response of any local tolerance effects.

Special studies

Sensitization/immunogenicity, phototoxicity. Depends on specific test.

---

*In ICH S5A, Reproductive Toxicology: Detection of Toxicity to Reproduction for Medicinal Products, the sequence of events from premating and conception in one generation to conception in the next generation is divided into six stages, A–F inclusive, rather than three segments.
Any effects or abnormalities noted in standard *in vivo* tests are considered in relation to the drug dose administered, usually expressed in mg/kg/day or mg/m²/day, as well as to the systemic exposure (conventional metrics being $C_{\text{max}}$ and/or AUC). There is a reasonable expectation that toxic responses, other than those at the site of application, are likely to be dose related. If this is not the case then it is possible that the findings may be of doubtful significance for man (or there may be an effect only at the high dose). Toxicokinetic monitoring is now established as an essential component of repeated-dose toxicity studies and can help identify a variety of factors that can affect dose–response relationships and data interpretation (Dahlem *et al*., 1995; Schwartz, 2001). These include: first-pass metabolism and possible saturation thereof, enzyme induction, saturation of metabolic clearance and plasma accumulation of the drug and/or its metabolites.

It is often important to assess the reversibility of toxicological findings, and so a recovery
phase is frequently included in repeated-dose toxicity studies. For example, in a 3-month toxicity study a 4-week post-treatment reversibility phase would be typical and would normally involve the inclusion of additional animals, at least in the control and high-dose groups.

Testing in special types of animal (e.g. ovariectomized animals for osteoporosis products or juvenile animals for paediatric products) is sometimes advocated to assess pharmacodynamic and/or toxic responses. Although the former is quite well established (e.g. Vahle et al., 2002), use of juvenile animals is still being evaluated in the EU (CPMP Safety Working Party, 2001).

**Toxicity testing guidelines**

National guidelines on non-clinical testing requirements for pharmaceuticals have been available for the last 20–30 years. Such documents are intended to serve a number of purposes in terms of providing guidance to companies involved in drug development, such as:

- selection of an appropriate study package
- suitable study designs
- interpretation of results.

**International Conference on Harmonization**

Guidance in EU countries has been subject to two types of harmonization. Firstly, EU Notes for Guidance (NfGs) replaced all national guidelines (occurring mainly in the 1980s), and more recently a global harmonization procedure has been undertaken under the auspices of the International Conference on Harmonization (ICH; www.ich.org/). The ICH process (International Harmonization of Technical Requirements of Pharmaceuticals for Human Use) was initiated in October 1989 and was hosted by the International Federation of Pharmaceutical Manufacturers Associations (IFPMA; www.ifpma.org/). The principal objectives of the initial and subsequent conferences have been:

- To identify and eliminate the differing requirements in the three participating states/regions (USA, Japan and EU).
- To avoid repetition of all types of tests.
- To accelerate development of medicinal products, thus giving patients quicker access to new medicinal products without negatively affecting quality, safety and efficacy.

The technical discussions and drafting of guidelines is undertaken by expert working groups in quality, safety and efficacy supervised by the ICH Steering Committee. (Multidisciplinary working groups are also involved in some topics.) The working groups consist of representatives of each participating authority (Japan’s Ministry of Health and Welfare, the United States’ Food and Drug Administration (FDA) and the European Commission) and the pharmaceutical trade association from each region (Japan Pharmaceutical Manufac-
turers Association, Pharmaceutical Research and Manufacturers of America and European Federation of Pharmaceutical Industries’ Associations). Two observers (WHO and Canada) are part of the ICH global cooperation group.

Five steps, ranging from initiation to implementation, are involved in the ICH guideline process:

- **Step 1.** Preparation of concept paper and draft guideline by an expert working group.

- **Step 2.** A draft consensus guideline is signed by all six parties (authorities and associations) and released for 6 months’ consultation. A unanimous decision is required at this stage in order for the particular draft guideline to progress to the consultation stage. For example, a proposal for a guideline on immunotoxicity testing failed to secure unanimous agreement at the 2002 steering committee meeting in Brussels.

- **Step 3.** Comments received during the consultation period are evaluated by the competent authorities in the three regions and incorporated, as appropriate, into the existing draft. This draft is signed by the authorities and submitted to the steering committee for approval.

- **Step 4.** The final draft is approved by the steering committee and confirmed by the signatures of the regulatory authorities.

- **Step 5.** The process is concluded by implementation of the guideline in the three regions in compliance with legal and regulatory procedures.

Guideline maintenance has now become an issue after over 10 years of the ICH process, and some guidelines have been subjected to revision (denoted by R in parentheses) in recognition of scientific progress and/or the need for clarification. Some others have been updated using a maintenance procedure (M). ICH guidelines relevant to nonclinical testing are listed in Table 3.3.

**EU- and US-specific guidelines**

Although the ICH process has led to guideline harmonization for the major areas of nonclinical investigation, there remains scope (and probably this will always be the case) for national guidelines covering more specialized topics. In the EU, the CPMP, via its expert preclinical group, the Safety Working Party, has released guidance on a variety of topics either in the form of NfGs, Points to Consider (PTC) or Position Paper (PP) documents (Table 3.4).

US-specific guidelines may be located under ‘Guidance for Industry’ on the website of the FDA (www.fda.com).

**Biotechnology-derived and biological drugs**

During the last 10–15 years there has been a dramatic rise in the development and therapeutic use of biotechnological and biological products, sometimes called new biological entities (NBEs) or ‘biologics’. These drugs comprise a wide range of product types, including
vaccines, blood cells, rDNA versions of endogenous hormones, deliberately modified versions of natural hormones, cofactors and a variety of antibody-related entities (e.g. antibody fragments – mainly for diagnostic applications, humanized monoclonal antibodies). Non-clinical testing strategies for NBEs are highly case-specific, depending on product type, clinical indication and availability of suitable animal models (especially for human proteins in nonhomologous animal species) for both pharmacological and safety evaluation, leading to the frequent use of nonhuman primate species (Terrell and Green, 1994; Dayan, 1995; Griffiths and Lumley, 1998; Pilling, 1999; Serabian and Pilaro, 1999; Black et al., 2000; Dempster, 2000; Green and Black, 2000; Galluppi et al., 2001; Descotes et al., 2002; Verdier, 2002). Toxicity tests on the murine version of a recombinant product may provide useful

### Table 3.3 ICH guidelines relating to nonclinical testing

<table>
<thead>
<tr>
<th>Field</th>
<th>ICH topic</th>
<th>Guideline title</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safety</td>
<td>S1A</td>
<td>Guideline on the need for carcinogenicity studies of pharmaceuticals</td>
<td>Step 5</td>
</tr>
<tr>
<td></td>
<td>S1B</td>
<td>Carcinogenicity: testing for carcinogenicity of pharmaceuticals</td>
<td>Step 5</td>
</tr>
<tr>
<td></td>
<td>S1C</td>
<td>Carcinogenicity: dose selection for carcinogenicity studies of pharmaceuticals</td>
<td>Step 5</td>
</tr>
<tr>
<td></td>
<td>S1C(R)</td>
<td>Addendum: addition of a limited dose and related notes</td>
<td>Step 5</td>
</tr>
<tr>
<td></td>
<td>S2A</td>
<td>Genotoxicity: specific aspects of regulatory genotoxicity tests for pharmaceuticals</td>
<td>Step 5</td>
</tr>
<tr>
<td></td>
<td>S2B</td>
<td>Genotoxicity: a standard battery for genotoxicity testing of pharmaceuticals</td>
<td>Step 5</td>
</tr>
<tr>
<td></td>
<td>S3A</td>
<td>Toxicokinetics: the assessment of systemic exposure in toxicity studies</td>
<td>Step 5</td>
</tr>
<tr>
<td></td>
<td>S3B</td>
<td>Pharmacokinetics: guidance for repeated-dose tissue distribution studies</td>
<td>Step 5</td>
</tr>
<tr>
<td></td>
<td>S4A</td>
<td>Duration of chronic toxicity testing in animals (rodent and non-rodent)</td>
<td>Step 5</td>
</tr>
<tr>
<td></td>
<td>S5A</td>
<td>Reproductive toxicology: detection of toxicity to reproduction for medicinal products</td>
<td>Step 5</td>
</tr>
<tr>
<td></td>
<td>S5B(M)</td>
<td>Reproductive toxicology: toxicity to male fertility</td>
<td>Step 5</td>
</tr>
<tr>
<td></td>
<td>S6</td>
<td>Preclinical safety evaluation of biotechnology-derived pharmaceuticals</td>
<td>Step 5</td>
</tr>
<tr>
<td></td>
<td>S7A</td>
<td>Safety pharmacology studies for human pharmaceuticals</td>
<td>Step 5</td>
</tr>
<tr>
<td></td>
<td>S7B</td>
<td>Safety pharmacology studies for assessing the potential for delayed ventricular repolarization (QT interval prolongation) by human pharmaceuticals</td>
<td>Step 5</td>
</tr>
<tr>
<td>Quality*</td>
<td>Q3A(R)</td>
<td>Impurities testing: impurities in new active substances</td>
<td>Step 3</td>
</tr>
<tr>
<td></td>
<td>Q3B(R)</td>
<td>Impurities in new medicinal products</td>
<td>Step 3</td>
</tr>
<tr>
<td>Multidisciplinary</td>
<td>M3(M)</td>
<td>Nonclinical safety studies for the conduct of human clinical trials for pharmaceuticals</td>
<td>Step 5</td>
</tr>
</tbody>
</table>

*ICH Q3C(M). Note for guidance on impurities: residual solvents, is not included since, although the permitted residues of specified solvents are derived on the basis of toxicological data, there is no obvious opportunity for applicants to make independent safety-based assessments that will be acceptable to all regulatory authorities.
information, as reported for recombinant interferon gamma (Terrell and Green, 1993). Nevertheless, whenever possible, NBEs should be tested using the same range of study types as for NCEs with appropriate modifications in respect of dose levels, species and endpoints (e.g. determination of serum neutralizing antibodies). Where an evaluation of carcinogenic potential is feasible and appropriate (e.g. for a chronic indication such as diabetes or osteoporosis), use of only one rodent species (normally the rat) is acceptable (ICH S6 Guideline). Most biological products are of a proteinaceous nature and so would not be expected to exhibit genotoxic properties, unless, for example, residues of linker molecules were unexpectedly present. But nongenotoxic proteins, especially those with higher potency trophic activity in rodents compared with man, can still produce a neoplastic response, as has been reported for recombinant human parathyroid hormone (1–34) when evaluated in a conventional rat bioassay (Vahle et al., 2002). The latter finding seems unlikely to represent a hazard for patients when factors such as relative potency and duration and extent of exposure are considered, but data from additional studies will be required to confirm this.

**Dossier compilation and Common Technical Document**

For reasons of consistency and ease of assessment, regulatory authorities have established detailed requirements for dossier content and order of presentation. As well as copies of

---

**Table 3.4 EU-specific guidance on nonclinical testing**

| Reference    | Title                                                                 | Status    |
|--------------|                                                                      |           |
| CPMP/SWP/465/95 | Preclinical pharmacological and toxicological testing of vaccines | Adopted NfG |
| CPMP/SWP/728/95 | Replacement of animal studies by in vitro models                     | Adopted NfG |
| CPMP/SWP/997/96 | Pre-clinical evaluation of anticancer medicinal products             | Adopted NfG |
| CPMP/SWP/1042/99 | Repeated dose toxicity                                                | Adopted NfG |
| CPMP/SWP/2145/00 | Non-clinical local tolerance testing of medicinal products           | Adopted NfG |
| CPMP/SWP/112/98 | Safety studies for gene therapy products                              | Draft NfG |
| CPMP/SWP/2877/00 | Carcinogenic potential                                               | Adopted NfG |
| CPMP/SWP/4446/00 | Specification limits for residues of metal catalysts                | Draft NfG |
| CPMP/SWP/398/01 | Photosafety testing                                                  | Draft NfG |
| 3CC29a       | Investigation of chiral active substances                             | Adopted NfG |
| CPMP/986/96  | Assessment of the potential for QT interval prolongation by non-cardiovascular medicinal products | PTC       |
| CPMP/SWP/372/01 | Non-clinical assessment of the carcinogenic potential of insulin analogues | PTC       |
| CPMP/SWP/2600/01 | Need for the assessment of reproductive toxicity of human insulin analogues | PTC       |
| CPMP/SWP/2592/02 | CPMP SWP conclusions and recommendations with regard to the use of genetically modified animal models for carcinogenicity assessment | Recommendations |
actual study reports, different types of summary, overview and critical assessment have been specified by authorities as part of the application dossiers for new drugs, called new drug applications (NDAs) in the USA and marketing authorization applications (MAAs) in the EU. An initiative by ICH (www.ich.org/) has produced the Common Technical Document (CTD) – a global guideline on dossier format (Table 3.5). The nonclinical modules are 2.4, 2.6 and 4. This format, accompanied by an optional electronic version (eCTD), is expected to become mandatory in the next 1–2 years on a worldwide basis; its introduction should lead to considerable time and cost savings in dossier preparation. It seems unlikely, however, that the ultimate goal of the global dossier (i.e. one identical dossier acceptable to all regulatory authorities worldwide) will be achieved in the foreseeable future owing to inter-regional differences in legal systems, medical practice and testing guidelines.

### Table 3.5 Outline of Common Technical Document (CTD)

<table>
<thead>
<tr>
<th>Module</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Administrative information and prescribing information</td>
</tr>
<tr>
<td>2</td>
<td>Common Technical Document summaries</td>
</tr>
<tr>
<td>2.1</td>
<td>CTD table of contents</td>
</tr>
<tr>
<td>2.2</td>
<td>CTD introduction</td>
</tr>
<tr>
<td>2.3</td>
<td>Quality overall summary</td>
</tr>
<tr>
<td>2.4</td>
<td>Nonclinical overview</td>
</tr>
<tr>
<td>2.5</td>
<td>Clinical overview</td>
</tr>
<tr>
<td>2.6</td>
<td>Nonclinical written and tabulated summaries</td>
</tr>
<tr>
<td></td>
<td>Pharmacology</td>
</tr>
<tr>
<td></td>
<td>Pharmacokinetics</td>
</tr>
<tr>
<td></td>
<td>Toxicology</td>
</tr>
<tr>
<td>2.7</td>
<td>Clinical summary</td>
</tr>
<tr>
<td>3</td>
<td>Quality</td>
</tr>
<tr>
<td>4</td>
<td>Nonclinical study reports</td>
</tr>
<tr>
<td>4.1</td>
<td>Module 4 table of contents</td>
</tr>
<tr>
<td>4.2</td>
<td>Study reports</td>
</tr>
<tr>
<td>4.3</td>
<td>Literature references</td>
</tr>
<tr>
<td>5</td>
<td>Clinical study reports</td>
</tr>
</tbody>
</table>

### Drug development

#### Introduction

Drug development is an exceedingly complex, high-risk and costly process involving scientists from many disciplines. At the time of writing (late 2002), the development programme from discovery to authorization for a typical NCE would be expected to take at least 6 years and cost not less than £500M (House of Lords, 2002). A recent UK report indicates that the UK drug failure rate is still high; of 320 compounds reported to be in development in 1998, only 47 are now available as approved medicines; about 150 of these were discontinued and the remainder are still in development (Anonymous, 2002). This pattern is being replicated worldwide. It is claimed that of half a million chemical structures/compounds initially considered, computational and other (*in vitro*) preclinical lead
optimization screening reduces the number tested in animals to ten, and the results of animal studies cause seven to be rejected; three compounds go into clinical trials and just one is eventually authorized for human use (House of Lords, 2002). Higher success rates have been reported (e.g. 0.01–0.02 per cent; Dorato and Vodicnik, 2001), but such estimates may go back some years, when more targets amenable to simple rational approaches remained available for commercial exploitation, and may predate the advent of high-throughput screening techniques.

Increasing attention in the pharmaceutical industry is being focused on the declining numbers of new molecular entities (NMEs; includes NCEs and biologics) going into clinical development and achieving regulatory approval in spite of an increase in NMEs entering preclinical development (Jones, 2002). Reasons suggested to explain this include the increasing complexity of the clinical workup and the rarity of unmet therapeutic need in ‘easy’ disorders (with clear clinical endpoints and well-understood pathophysiology). Even modest clinical gains in diseases that tend to be chronic and the result of aging, and which are not readily treatable using current drug therapies, can be quite difficult to achieve (Cohen, 2002).

**Toxicological requirements for conventional clinical trial programmes**

Clinical trials in drug development are normally divided into three phases: phase 1, phase 2 and phase 3. Table 3.6 summarizes the essential elements of such trials, including toxicological data requirements for a typical NCE. Some authors add one or two further phases: phase 0 (nonclinical discovery phase) and phase 4 (post-marketing studies for further evaluation of safety in normal clinical use and/or assessment of comparative risk–benefit). The division of phase 2 into two parts (e.g. 2A and 2B), and sometimes similarly with phase 1, can provide a more stepwise and cautious approach (Lesko et al., 2000).

<table>
<thead>
<tr>
<th>Phase</th>
<th>Description</th>
<th>Number of patients</th>
<th>Normal toxicological requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Initial studies normally in (male) volunteers, but sometimes in patients, to determine tolerance, safe dosage range and basic kinetics and metabolism</td>
<td>30–50</td>
<td>2–4 weeks in rodent and non-rodent basic genotoxicity and pharmacokinetics</td>
</tr>
<tr>
<td></td>
<td>Early controlled trials in a limited number of patients under closely monitored conditions to determine preliminary efficacy and short-term safety at a range of doses</td>
<td>250–500</td>
<td>3–6 months in rodent and non-rodent extensive genotoxicity and pharmacokinetics rat and rabbit teratology</td>
</tr>
<tr>
<td>2</td>
<td>Extended large-scale controlled trials to obtain definitive evidence of efficacy and safety, and to characterize the adverse-event profile. Studies on drug interactions and in special patient groups (e.g. elderly, hepatic/renal impairment)</td>
<td>300–3000</td>
<td>6 months in rodent and 9 months in non-rodent segments I and III reproductive toxicity carcinogenicity</td>
</tr>
</tbody>
</table>

*See text comments on single-dose toxicity studies.*
Drug discovery phase

‘Conventional’ drug discovery involving ‘rational’ design of small organic molecules based on structure–activity considerations relating to drug target (e.g. inhibition or augmentation of a particular enzyme, cytokine or neurotransmitter) is still undertaken. A range of techniques, such as chemoinformatics (e.g. to assemble virtual compound libraries), combinatorial chemistry, genomics, proteomics and high-throughput screening are now employed to complement the traditional approaches (Atterwill and Wing, 2000; Johnson et al., 2001; Schmid et al., 2001; Alaoui-Ismaili et al., 2002; Augen, 2002; Törnell and Snaith, 2002; Valler, 2002; van Dongen et al., 2002).

Following candidate selection and the application of various screening procedures (for potential pharmacotoxicological activity), the more promising compounds would be further evaluated using in vitro and in vivo pharmacological models (Caldwell, 2001). Surviving candidates would then be eligible for an initial toxicity evaluation, usually involving in vitro assays for genotoxicity and single- and/or repeated-dose toxicity studies in one or two animal species. A preliminary ADME assessment in vitro and in animals would also be undertaken in normal circumstances. Although a drug may show activity in vitro and also possibly in some in vivo (rodent) models, kinetic and metabolic factors may alter the nature and magnitude of this response in man and in animal species that provide better models for man. Therefore, it is important at this early stage, and throughout the development programme, to integrate the available pharmacodynamic and pharmacokinetic information. Incorporating initial evaluations of toxicity and pharmacokinetics within the drug discovery phase may provide sufficient information to enable, where appropriate, some ‘re-engineering’ of the chemical structure of a promising candidate in order, for example, to improve bioavailability or minimize toxicity. A major objective of the initial toxicological assessment is to provide sufficient reassuring safety information to proceed with a first-dose-in-man (sometimes called first-into-man, FIM) study, subject of course to ethics committee review (Johnson and Wolfgang, 2001).

Parallel programmes of chemical (or biotechnological) and pharmaceutical development, carefully coordinated with the non-clinical and clinical programmes, need to be undertaken, and are crucial to the eventual authorization and commercialization of any new drug. The specification of the test material used in non-clinical studies is likely to change as its synthesis progresses from a bench- to pilot-plant- and eventually to commercial-scale process. Detailed analytical information should be available on batches of test material used in non-clinical (and clinical) studies to evaluate the consistency of the impurity profile and whether the material toxicologically tested is representative in terms of chemical composition of the proposed commercial active ingredient. Owing to difficulties in achieving adequate physicochemical characterization of biotechnology and biological products, extremely close attention to process control is required in order to manufacture a consistent active ingredient.

FIM studies to assess tolerability and kinetics are usually undertaken in healthy (male) volunteers (Meulenbelt et al., 1998). Such a study (studies) would be the first in the phase 1 programme. Other studies in patients in order to obtain an indication of pharmacodynamic effects, potential efficacy and dose–response relationships, possibly using a surrogate marker, would follow as soon as possible in clinical development. Increasingly, companies are attempting to accelerate evaluation in humans, one approach being to employ low-dose proof-of-concept (POC) studies in phase 1 rather than phase 2 (Lesko et al., 2000).
Toxicological requirements for FIM and most other phase 1 studies in the EU include two 14-day repeated-dose studies in a rodent and non-rodent species (normally the rat and dog – ICH M3 guideline). However, in the USA, the FDA can authorize a single-dose clinical trial on the basis of data from extended single-dose toxicity studies in the rodent and non-rodent. A recent draft position paper issued by the EU Committee on Proprietary Medicinal Products (CPMP) proposes a somewhat similar but more restrictive approach of using an extended single-dose toxicity study in one species to support single low-dose exploratory screening trials in humans, e.g. in order to characterize pharmacokinetic properties or receptor selectivity using positron emission tomography (PET) imaging or other sensitive analytical techniques (Table 3.4).

A high proportion of early development projects, especially in small companies, involve anticancer drugs, since there remain significant unmet therapeutic needs and development times tend to be somewhat shorter than those in most other therapeutic areas. For anticancer drugs in general, and cytotoxics in particular, initial development has a number of special features (DeGeorge et al., 1998; Den Otter et al., 2002):

- All studies have to be undertaken in patients, for ethical reasons.
- As all cytotoxics exhibit relatively similar toxicity profiles (targeting cell populations with a rapid turnover), toxicological evaluation tends to be focused on revealing any important deviations from the expected toxicity profile, establishing a no-observed-adverse-effect level (NOAEL) and providing basic information on pharmacokinetics. In the past, virtually all of this work was successfully undertaken in one species, usually the rat, but use of additional species such as the dog is being increasingly recommended (Clark et al., 1999).
- A safe starting dose for entry into patients can be derived from the toxicological data; one-third of the rodent LD_{10} or one-third of the dog toxic high dose, both in mg/m², is suggested by Clark et al. (1999) for platinum-based anticancers. Pharmacokinetically guided approaches can also be applied (Reigner and Blesch, 2002).

**Drug development phase**

**Strategic considerations**

The selection of which candidate drug(s) to take from discovery into development is frequently rated as the most important decision in the drug-development process (Parkinson et al., 1996). All toxicological information available at this early stage plays a major role at this go/no-go decision point, and further toxicological studies play an enabling role in providing key safety data to support phase 2 and phase 3 trials. The timing of toxicological studies in relation to important milestones in the clinical programme is an important strategic issue that has to be decided by individual companies on a case-by-case basis. Some companies may adopt a cautious, cost-effective approach and undertake just enough studies to support the next clinical trial, whereas others may be less risk averse and decide to commission carcinogenicity studies on the basis of early results from phase 2A trials, essentially taking a calculated risk on a positive outcome to the phase 2 programme and associated toxicological studies.
**Statistical aspects**

Of necessity, owing to a range of considerations including animal usage, consumption of test material, cost and time, many compromises are involved in the design of toxicity studies, particularly in respect of the number of dose groups and the number animals per group. In repeated-dose toxicity studies, group sizes of 10–20 rodents and four to six non-rodents/sex are normally employed. In conventional carcinogenicity studies, the usual group size is 50 animals/sex. Thus, toxicity studies tend to have less statistical power than most phase 3 studies, typically involving hundreds of patients. For example, consider a case where the desired $\alpha$-level (type I error) is set at 0.05 (single-sided test) and the $\beta$-level (type II error) is set at 0.1 (i.e. 90% chance of detecting a unidirectional effect at the 95 per cent confidence level. For two groups of animals (control and test), if the historical control response is 50 units with a standard deviation of 20 units and one wishes to detect a difference in response of 10 units (i.e. $\delta = 10/20 = 0.5$), then group sizes of 36 animals would be required. For a two-sided test with $\alpha = 0.05$, 44 animals per group would be required (Lee, 1993). However, one is able to administer multiples of the (anticipated) therapeutic dose in animals which may partially compensate for the low number of animals. Dosage selection in toxicological studies has traditionally been based on a high degree of empiricism, but the use of toxicokinetic data from dose-ranging studies can provide a more rational approach (Bus and Reitz, 1992; Spurling and Carey, 1992; Morgan et al., 1994; Swenberg, 1995).

**Species selection**

Species selection in drug development is usually based on pilot toxicity and general pharmacology studies in rodent and non-rodent species together with supporting kinetic and metabolic information. In practical terms, choices are limited to species that are available from laboratory animal suppliers, which are of a suitable size and for which there is an adequate database for parameters such as survival, haematology and clinical chemistry on control animals and in which appropriate investigations and measurements are feasible. Within the limited range of options, selection of the most appropriate species can be crucial, particularly in respect of toxicity studies in non-rodents and embryotoxicity studies (Morton, 1998; Dixit, 2000; Smith et al., 2001).

**Effectiveness of standard nonclinical studies**

Toxicological studies are considered by those involved in drug development to be indispensable in respect of highlighting the principal target organs that are at risk of exhibiting toxic responses; development of a significant number of drug candidates is curtailed based on the results of nonclinical studies (Broadhead et al., 2000). Unfortunately, little information on this valuable function of identifying drug candidates that are toxic and/or appear to have an inferior benefit–risk profile are in the public domain, presumably because commercial pressures force pharmaceutical companies to concentrate on developing leading candidates without diverting resources to compile and publish data on those that have been discarded. On the other hand, standard toxicological studies have, at the margin, inherent limitations associated mainly with statistical considerations (see above) and the less-than-perfect nature of animal models. Consequently, it must be accepted that not all human
adverse events will be predicted, particularly those that occur with low frequency (the latter generally not being well predicted in clinical trials).

Zbinden (1991) highlights a number of examples of human drug tragedies (such as those associated with chloramphenicol, hexachloraphene, gossypol and methoxyflurane) that, with hindsight, could have been prevented by undertaking appropriate animal studies combined with taking suitable precautionary measures. Although such examples of drug toxicity that would have been detected with current testing strategies are of largely historical interest, they illustrate the continuing improvements in pharmaceutical toxicology and the potential dangers of moving prematurely to alternative (non-animal) methods (House of Lords, 2002).

As a drug development programme proceeds, pre-clinical data are, to a large extent, progressively superseded by clinical safety data gathered in clinical trials, provided that the safety monitoring is of sufficient breadth and depth to evaluate appropriate target organs and toxic responses found in animals. Some non-clinical toxicity endpoints, such as reproductive toxicity, genotoxicity and carcinogenicity, are not amenable to clinical experimentation owing to a variety of practical and ethical factors. For newly authorized drugs there is almost total reliance on animal and in vitro data for safety assessments in these three areas.

At the clinical trial stage, and immediately following authorization, companies and regulatory agencies generally take a cautious approach regarding the use of a new drug in pregnant women. Even when a drug causes no adverse effects in the standard battery of reproductive toxicity tests (Table 3.1), pregnant women would normally be advised to avoid the drug unless its use is absolutely essential (European Commission, 1999). The availability of reassuring information on accidental or deliberate foetal exposure during clinical trials may lead to less restrictive labelling. These would be unusual circumstances, however, since pregnant women are normally excluded from clinical trials. Some relaxation in pregnancy labelling is possible if evidence can be presented to regulatory authorities demonstrating that there is no association with adverse reproductive effects after several years of clinical use.

**Nonstandard studies**

Nonstandard and/or special investigative studies are often commissioned during drug development in an attempt to answer specific questions. For example, non-clinical studies may predict a particular toxic response that is, in fact, not observed in clinical trials. It is more convincing to be able to explain why a toxic effect occurred in animals but not in clinical-trial patients rather than just relying on the absence of clinical adverse events. Special studies could be undertaken in vitro and/or in animals in order to understand the causative factors of the effect and whether the mechanism involved excludes the likelihood of a human response. Some examples are:

- Animal toxicity attributable to species-specific kinetic, metabolic or pharmacodynamic effects (Warrington et al., 2002; Tabacova and Kimmel, 2001; Soars et al., 2001; Van Gelder et al., 2000; Molderings et al., 2000; Honma et al., 2001).
- Validation and use of the rhesus monkey as a suitable model for testing the effects of finasteride, a type 2 5-α-reductase inhibitor, on external genital differentiation of the male foetus (Prahalada et al., 1997).
- Various studies on tamoxifen demonstrating that the formation of liver tumours in the rat is not relevant to the use of the drug in women (De Matteis et al., 1998; White et al., 2001).
**Validity and relevance of non-clinical testing programme**

Through a process of continual improvement and capture of best practice in regulatory guidelines, non-clinical testing programmes in drug development tend to be similar across a broad range of therapeutic categories. But there still remains a degree of flexibility that enables individual programmes to be improved, for instance by ‘backtracking’ and undertaking more detailed investigations in problem areas and/or using special studies. In this way it is possible to produce a robust non-clinical database of optimum relevance to human safety assessment. Some of the more important factors that need to be considered in this optimization process are summarized in Table 3.7.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Main considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope</td>
<td>• All appropriate endpoints incorporated, e.g. potential carcinogenic effects of growth factors</td>
</tr>
<tr>
<td>Species suitability</td>
<td>• Pharmacodynamically responsive to drug</td>
</tr>
<tr>
<td></td>
<td>• Similar metabolic profile to that in man</td>
</tr>
<tr>
<td>Receptor profiles</td>
<td>• Adequate separation in targeting of desired and undesired receptors in terms of pharmacodynamic responses in animal models and humans</td>
</tr>
<tr>
<td></td>
<td>• Applicability of highly specific receptor targeting in animal models to clinical situation</td>
</tr>
<tr>
<td>Kinetics</td>
<td>• Linear or non-linear kinetics?</td>
</tr>
<tr>
<td></td>
<td>• Reasons for non-linearity (e.g. saturation of excretion mechanisms, liver enzyme induction)</td>
</tr>
<tr>
<td></td>
<td>• Plasma accumulation on repeated dosing?</td>
</tr>
<tr>
<td>Metabolism</td>
<td>• Safety evaluation of any significant human metabolites not detected in animal models</td>
</tr>
<tr>
<td>Study design and implementation</td>
<td>• Adequate number of dose groups and numbers of animals</td>
</tr>
<tr>
<td></td>
<td>• Use of one animal gender or two as appropriate</td>
</tr>
<tr>
<td></td>
<td>• Suitable and validated endpoints and test methods</td>
</tr>
<tr>
<td></td>
<td>• Duration of dosing relevant to indication</td>
</tr>
<tr>
<td></td>
<td>• Use of recovery groups</td>
</tr>
<tr>
<td></td>
<td>• GLP considerations</td>
</tr>
<tr>
<td>Dose levels and exposure</td>
<td>• Dose levels justified by range-finding studies</td>
</tr>
<tr>
<td></td>
<td>• Adequate toxicokinetic monitoring incorporated into toxicity studies</td>
</tr>
<tr>
<td>Genotoxicity</td>
<td>• Confirmation that metabolic activating system employed in <em>in vitro</em> assays is capable of simulating <em>in vivo</em> metabolism</td>
</tr>
<tr>
<td></td>
<td>• If not confirmed, separate studies on the principal metabolites may be necessary</td>
</tr>
<tr>
<td></td>
<td>• Relationship between concentration of test material used in <em>in vitro</em> assays and human plasma $C_{\text{max}}$</td>
</tr>
<tr>
<td></td>
<td>• Dose- or concentration-related changes in metabolism</td>
</tr>
<tr>
<td>Mechanistic studies</td>
<td>• Direct or indirect consequence of exaggerated pharmacology?</td>
</tr>
<tr>
<td></td>
<td>• Species specificity and reasons for this</td>
</tr>
<tr>
<td></td>
<td>• Disruption of homeostatic mechanisms, e.g. by modification of endocrine system</td>
</tr>
</tbody>
</table>
Data interpretation and risk assessment

Interpretation of nonclinical toxicity data

Routine toxicity tests deliberately incorporate a plethora of different endpoints (Table 3.1) selected on the basis of experience to be effective at detecting toxic responses. It is common for statistically significant differences to be observed between parameters for test and concurrent control animals, but such differences may occur by chance or through normal biological variation, not reflecting a genuine treatment-related effect. The greater the number of different endpoints, the higher will be the probability that some differences will occur by chance. Careful scrutiny of the data from individual studies is required in order to assess whether the effects observed fit a logical pattern indicative of a toxic response. Subsequently, data from several studies should be evaluated to ascertain consistency of response. Any major inter-study inconsistencies should be thoroughly investigated in an attempt to establish the reasons for the variable response.

The processes involved in non-clinical data interpretation include:

- Establishment of pattern of toxic response
  - endpoints affected
  - route/dose/exposure/time relationships
  - reversibility
  - inter-species variation
- Determination of target organ(s)
  - confirmed by multiple endpoints
  - inter-species differences explained (e.g. by kinetic, metabolic and/or known species sensitivities)
  - pharmacodynamic and/or toxicological effects
- Assessment of no-adverse-effect and lowest-effect doses and the systemic exposures corresponding to these doses.
- Proposed toxicological mechanism(s)
  - use of established precedents and information on class effects
  - special studies commissioned as appropriate.

A toxic effect in a particular organ will normally be associated with changes in a variety of parameters. For example, a toxic response in the liver would be expected to be associated with a change in bodyweight-related organ weight, increases in serum enzymes associated with hepatotoxicity (e.g. alanine aminotransferase and aspartate aminotransferase) and histopathological changes. Although slight alterations in one parameter (e.g. in serum enzymes) may be suggestive of liver damage, without confirmatory evidence from other sources, hepatotoxicity would not in this case normally be considered as a major concern.

Thus, a variety of separate effects may often be consequences of the same toxicological
process. Gaining an understanding of the pattern of toxicity may suggest a causal mechanism that is often a critical prerequisite for effective extrapolation of non-clinical toxicological findings to man.

**Risk assessment: extrapolation of toxicological data to man**

**Introduction**

Risk assessment is the process of determining the types and likelihoods of adverse reactions in humans that may result from exposure to chemical, biological or physical hazards (Brecher, 1997). In the context of drug development, particularly at the early stages, the essence of risk assessment is the extrapolation of the non-clinical data to man. Most of the information is derived from *in vivo* experiments in animals using high doses of the drug substance. The relevance of responses in animals to patients using the intended therapeutic dose will be assessed by a number of interested parties (company, regulatory agency/ethics committee for clinical trials, regulatory agency at the MAA stage) at various time points during development.

**Allometry**

In the early days of drug toxicology there was a search, ultimately futile, to find a test species that metabolized drugs in the same way as humans. Eventually, it became apparent that drug metabolism in animals hardly ever proceeds at the same rate as in humans. Several investigators noted that drug clearance per unit of body weight was nearly always considerably higher in small animals compared with larger animals such as man. Determination of biological half-life and clearance of drugs in several species suggested that these and other pharmacokinetic parameters were proportional to some power of the body mass. In other words, several fundamental pharmacokinetic parameters appeared to obey allometric principles (i.e. the study of size and its consequences).

The general allometric equation linking morphological and biological functions $Y$ and body weight $W$ is

$$Y = aW^b$$

(3.1)

where $a$ is the allometric coefficient and $b$ is the allometric exponent.

A corollary of this equation is that the traditional bases for extrapolation of data, i.e. $W^{1.0}$ (mg/kg body weight) and $W^{0.67}$ (mg/m$^2$ body surface area) have no unique justification; they are just two examples of the general case and provide quantitatively different scaling factors.

Logarithmic transformation of equation (3.1) yields:

$$\log Y = \log a + b \log W$$

(3.2)

Equation (3.2) is of the form $y = mx + c$, and so it is possible to plot $\log Y$ versus $\log W$ for different animal species and from the linear relationship to predict values of $Y$ for man. Alternatively, the data can be analysed by linear regression using a statistical software package.

Unfortunately, the application of allometric interspecies scaling to animal toxicity and
pharmacokinetic data has led to somewhat disappointing results (Voci and Farber, 1988; Bachmann, 1989; Chappell and Mordenti, 1991; Mordenti et al., 1992; Ritschel et al., 1992; Lin, 1998; Mahmood, 1998, 1999a,b, 2000a, 2001a, 2002; Lave et al., 1999; Mahmood and Balian, 1999; Mahmood and Sahajwalla, 2002).

Scaling of toxicity data has been reasonably successful (e.g. up to 80% of compounds) for single-dose studies. Acute toxicity data (e.g. LD10, MTD) for chemotherapeutic drugs have been extensively evaluated with exponents ranging from 0.6 to 0.9 (Chappell and Mordenti, 1991); an exponent of 0.75 appears to be more effective than use of surface-area relationships ($b = 0.67$) (Travis and White, 1988; Mahmood, 2001b). Travis (1991) also recommends in the more general case the use of 0.75 as exponent in preference to the conventional parameters (0.67 or 1.0). In a survey of 26 NCEs evaluated by the Medicines Evaluation Board in the Netherlands in the early 1990s, extrapolation on the basis of ‘metabolism equivalents’ using 0.75 as exponent was found to produce the closest fit to extrapolations based on pharmacokinetic data (Peters-Volleberg et al., 1994).

There are many examples of successful pharmacokinetic interspecies scaling, particularly for drugs that are completely or largely renally excreted (Chappell and Mordenti, 1991; Lin, 1998). Determination of the most important pharmacokinetic parameters (e.g. distribution volumes, half-life, clearance, AUC) from experiments using young adults of four animal species followed by linear regression is recommended (Chappell and Mordenti, 1991). Drugs cleared by hepatic extraction are more difficult to evaluate by interspecies scaling, though some success has been achieved by correlating half-life with liver blood flow rather than body weight (Aviles et al., 2001). The need to use brain weight as well as body weight in these situations brings a multivariate aspect to the problem (Lin, 1998).

Extrapolation to humans is often not straightforward, owing to the many intrinsic minor biochemical and physiological differences between animals and humans. Various minor modifications can be made to the basic allometric pharmacokinetic model, such as employing free-fraction drug concentration [rather than the total (free + bound) concentration] and lifespan potential (physiological time). However, the value of using free-fraction concentrations has been disputed (Mahmood and Balian, 1999; Mahmood, 2000b).

A more sophisticated approach, physiologically based pharmacokinetics (PB-PK), involves mass balance models in which it is generally assumed that organs and tissues with similar behaviour can be combined into compartments and connected by the fluid motion through the compartments. This is a reductionist paradigm in contrast to the predominantly empirical approach in allometric scaling (Chappell and Mordenti, 1991). Setting up PB-PK models is time consuming, data intensive and costly (Campbell, 1996), and has not achieved a significant uptake in pharmaceutical toxicology.

**NOAEL safety factor (or margin of exposure) approach combined with conventional dosimetry**

This classical approach to risk assessment relies on applying a safety factor to the NOAEL obtained in the ‘most sensitive species’. The conventional safety factor is 100 (10 each for intra- and inter-species variation), as is still employed in the assessment of food additives and other materials (Gaylor, 1983; Feron et al., 1990; Newman et al., 1993; Walker, 1998). For drug substances, where the (intended) maximum human dose (MHD) is known, individual safety factors (sometimes called margin of exposure, MoE, when used in this
context) can be calculated for each toxicological effect by dividing the NOAEL by the MHD (both in mg/kg/day).

Many deficiencies in this approach have been identified (Garattini, 1985; Berry, 1988). These include:

- Absence of mechanistic/pharmacokinetic considerations, particularly in respect of why one species is more sensitive than another, and the relevance of this to man.
- Reliance on the applied dose, leading for example to exaggeration of safety factors, particularly those based on rodent data.
- Frequent nonlinearity of dose–exposure relationships, especially in animals at higher doses.

Body-surface-area-based doses are employed in interspecies scaling for some drug classes (e.g. anticancers and antivirals), and this metric has been described as a more accurate and conservative method (compared with using mg/kg doses) for general use in regulatory toxicology (Voisin et al., 1990). However, scaling on the basis of mg/m$^2$ doses is probably better described as over-conservative since animal:human safety margins tend to be underestimated compared with those based on kinetic data (Table 3.8; Peters-Volleberg et al., 1994). This is further illustrated in the phenolphthalein example (Table 3.9).

Given the above criticisms, the NOAEL safety factor approach is generally avoided in pharmaceutical toxicology, although its use has been recommended in risk assessment of reproductive toxicity data (Newman et al., 1993). The technique may be employed with reluctance in situations where no comparative kinetic data are available. The risk assessment of impurities, such as solvents based on sub-acute (rodent) NOAELs, is an important example (see ICH Q3C). Special considerations apply in respect of the comparative dosimetry of inhaled drugs (Dahl et al., 1991; Bide et al., 2000; Mahmood, 2001c).

In spite of its many deficiencies, the NOAEL safety factor approach has been used since the 1950s by the FAO/WHO Joint Expert Committee on Food Additives (JECFA), and appears to have been successful in terms of preventing adverse effects in consumers. This probably reflects the intrinsic conservatism of the JECFA procedure, which generally employs high safety factors ($\geq 100$).

### Table 3.8 Body weights and scaling factors based on exponents of 0.67 and 0.75$^a$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Man</th>
<th>Beagle dog (cynomolgus)</th>
<th>Monkey</th>
<th>Rabbit</th>
<th>Rat</th>
<th>Mouse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body weight (kg)</td>
<td>70</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>0.20</td>
<td>0.025</td>
</tr>
<tr>
<td>Exponent 0.67</td>
<td>1</td>
<td>2.0</td>
<td>2.2</td>
<td>3.2</td>
<td>6.9</td>
<td>14</td>
</tr>
<tr>
<td>Exponent 0.75</td>
<td>1</td>
<td>1.7</td>
<td>1.8</td>
<td>2.4</td>
<td>4.3</td>
<td>7.3</td>
</tr>
</tbody>
</table>

$^a$ Example: to scale on the basis of exponent 0.75, a dose of 4.3 mg/kg/day in the rat corresponds to 1 mg/kg/day in man.

Scaling factors are calculated from $(W_a/W_h)^b$, where $W_a$ is the body weight of animal, $W_h$ is the human body weight, $b$ is an exponent. For derivation see Rodricks et al., (2001).
Use of NOAEL and relative systemic exposure

In this commonly used technique, safety factors are based on systemic exposure of the drug in animals at the NOAEL relative to that in man at the maximum human dose (MHD):

\[
\text{Safety Factor} = \frac{\text{Animal AUC at NOAEL}}{\text{Human AUC at MHD}}
\]

Exposure margins in carcinogenicity studies can also be calculated in a similar fashion using the animal AUC at the high dose (or other dose) for the numerator.

AUC tends to be the default, but \( C_{\text{max}} \) or other appropriate metrics related to systemic exposure may also be employed. The measure of systemic exposure can be based on parent drug substance alone and/or important (active) metabolites. Stereochemical preferences in the disposition of racemic drugs often differ among species, e.g. in relation to the nature and extent of chiral inversion. Consequently, exposure extrapolations for chiral drugs from one species to another should be made with caution (Ruelius, 1987).

Scaling on the basis of relative systemic exposure avoids a variety of problems associated with other approaches (Voisin et al., 1990):

- Drugs that are extensively metabolized cannot be compared across species using models that rely on body weight.
- Although numerous mammalian physiological parameters are related to body surface area \( (W^{0.67}) \) rather than body weight, the specific metabolic profile of many drugs does not correlate with overall metabolic rate, and thus with surface area.

### Table 3.9 Phenolphthalein: animal: man exposure multiples (EMs) based on three different metrics in rat and mouse dietary carcinogenicity studies

<table>
<thead>
<tr>
<th>Species</th>
<th>Dose in mg/kg/day (EM)</th>
<th>Dose in mg/m²/day (EM)</th>
<th>Animal AUC$_{24}$ (h.μmol/l)</th>
<th>EM based on AUC$_{24}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rat, F344 (male)</td>
<td>24.9 (12.5)</td>
<td>129 (1.74)</td>
<td>1090</td>
<td>4.84</td>
</tr>
<tr>
<td></td>
<td>58.8 (24.4)</td>
<td>306 (4.14)</td>
<td>1550</td>
<td>6.84</td>
</tr>
<tr>
<td></td>
<td>176 (88)</td>
<td>915 (12.4)</td>
<td>4290</td>
<td>19.0</td>
</tr>
<tr>
<td></td>
<td>606 (303)</td>
<td>3150 (42.6)</td>
<td>9620</td>
<td>42.6</td>
</tr>
<tr>
<td></td>
<td>2780 (1390)</td>
<td>14500 (196)</td>
<td>9100</td>
<td>40.2</td>
</tr>
<tr>
<td>Mouse, B6C3F1 (male)</td>
<td>38.2 (19.1)</td>
<td>115 (1.55)</td>
<td>983</td>
<td>4.35</td>
</tr>
<tr>
<td></td>
<td>65.9 (33.0)</td>
<td>198 (2.68)</td>
<td>2350</td>
<td>10.4</td>
</tr>
<tr>
<td></td>
<td>143 (72)</td>
<td>429 (5.80)</td>
<td>3970</td>
<td>17.5</td>
</tr>
<tr>
<td></td>
<td>551 (276)</td>
<td>1650 (22.3)</td>
<td>9110</td>
<td>40.3</td>
</tr>
<tr>
<td></td>
<td>2140 (1070)</td>
<td>6410 (86.6)</td>
<td>15200</td>
<td>67.2</td>
</tr>
</tbody>
</table>

*Human data from study in male volunteers: dose 2.0 mg/kg, 74 mg/m²; AUC$_{24}$ 226 h.μmol/l; data based on measurement of total (free plus glucuronide conjugate) plasma phenolphthalein. Data are taken from Collins et al., (2000). Note: (a) subproportional dose-related increase in animal AUC (due to saturation of absorption, possibly plus some enzyme induction in the rat); (b) over- and under-estimation of exposure margins by mg/kg and mg/m² metrics respectively (except for mg/m² at highest doses in the rat, where absorption is saturated).
Quantitative interspecies differences in ADME profiles are highly drug specific and often confound body-weight-related interspecies relationships.

As well as being drug specific, using relative systemic exposure for scaling purposes enables data obtained by different routes of administration to be compared. For example, oral long-term bioassays may be employed to evaluate the carcinogenic potential of drugs that are intended for parenteral or inhalation administration in patients.

The NOAEL has been criticized in relation to its inferior statistical properties, e.g. for its sensitivity to sample size and its high sampling variability from experiment to experiment (Leisenring and Ryan, 1992; Brand et al., 1999). Although alternative approaches have often been advocated, the NOAEL is still used quite extensively.

**Pre-authorization risk assessment: species susceptibility and mechanistic studies**

When a drug substance produces adverse effects in an animal model that are considered unlikely to be relevant to human safety, existing information on species sensitivity (e.g. exaggerated gastrointestinal toxicity in rodents to NSAIDs, gastric carcinoids produced by chronic administration of proton pump inhibitors, formation of liver tumours in mice) and/or new data from drug-specific mechanistic studies are often extremely helpful in terms of risk assessment (Williams, 1997).

A number of mechanisms have been proposed to account for carcinogenic responses to nongenotoxic drugs, the strength of the evidence being quite variable from one case to another (Alden, 2000; Silva Lima and van der Laan, 2000). Chronic prolactin stimulation has been identified as a promoter of carcinogenicity (Yokoro et al., 1977; Johnson et al., 1995; Cook et al., 1999; Suwa et al., 2001) but confirmatory evidence such as serum prolactin data is important in regulatory decision making in individual cases. Interspecies differences in metabolism are known to account for differences in cancer susceptibility and toxicity (Hengstler et al., 1999).

Unexpected adverse events detected in clinical trials may sometimes be amenable to investigation in animal models. HER-2 (human epidermal growth factor receptor 2 protein) is a member of the c-erbB family of receptor tyrosine kinases and is overexpressed by 20–30 per cent of human breast cancers. HER-2 overexpression is an independent adverse prognostic factor. Trastuzumab, a humanized monoclonal antibody that binds with high affinity to the extracellular domain of HER-2, is effective when used in combination with cytotoxics in the second-line treatment of advanced metastatic breast cancer (Harries and Smith, 2002; Ligibel and Winer, 2002). However, in combination with anthracyclines, in patients there is a marked increase in cardiotoxicity: 24.5 per cent heart failure versus 7.4 per cent with anthracyclines alone (Garattini and Bertele, 2002; Keefe, 2002; Page et al., 2002; Tham et al., 2002). Attempts were made to develop an animal model for this interaction in order to investigate the mechanism with the aim of eliminating or minimizing the cardiotoxic response. But, at the time of marketing authorization in the EU, all such attempts had been unsuccessful (European Medicines Evaluation Agency (EMEA: www.eudra.org/humandocs/humans/epar.htm; European Public Assessment Report (EPAR) for Herceptin).

In summary, species susceptibility and mechanisms of toxicity play critical (generally qualitative) roles in risk assessment. Many useful drugs would have failed to gain registration without this type of evidence.
Predictivity of nonclinical studies

The effectiveness of non-clinical studies, particularly animal studies, at predicting human toxic responses to pharmaceuticals is difficult to assess because, as already noted, virtually all of the relevant data are owned by pharmaceutical companies and there are distinct commercial barriers to the release of this information. Even though it is possible in theory to make an assessment of data available in the public domain, or available to regulatory authorities, both of these would be highly biased datasets since they would fail to include data on the significant number of drugs whose development was terminated based solely on internal company decisions.

The problems inherent in the above situation have been addressed by the Health and Environmental Safety Institute/International Life Sciences Institute who compiled relevant data via a multinational survey of pharmaceutical companies (Olson et al., 2000). The survey included data from 12 companies on 150 compounds with 221 different human toxicities (HTs) being reported. Multiple HTs were reported in 47 cases. The results showed the true positive HT concordance rate of 71 per cent for rodent and non-rodent species, with non-rodents alone being predictive for 63 per cent of HTs and rodents alone for 43 per cent. The highest incidence of overall concordance was seen in haematological, gastrointestinal and cardiovascular HTs, and the least was seen in cutaneous HT. Where animal models, in one or more species, identified concordant HT, 94 per cent were first observed in studies of 1 month or less in duration. Of the 29 per cent of effects not detected in animal tests, the majority were of a type that animal tests were not designed to detect, or were intrinsically undetectable in this type of test, e.g. headache, dizziness and certain skin reactions.

Although the concept of undertaking repeated-dose toxicity studies in both rodents and non-rodents considerably predates the concordance survey described above, it is noteworthy that inclusion of the non-rodent markedly improves predictivity. In addition, techniques such as receptor and ADME profiling can help assess the degree of relevance to man of a particular species, and thus improve predictivity (Zbinden, 1993).

Adverse drug reactions detected after authorization

Adverse reactions and drug withdrawals

With long-term monitoring all drugs can be expected to exhibit side effects (i.e. unwanted effects of no therapeutic value) in some patients. Collection of safety data during clinical trials (especially phase 3 trials) enables detailed adverse event profiles to be compiled for a closely defined patient population. Such information is essential to the assessment of the benefit–risk of a particular drug. Rare adverse reactions, unlikely to be apparent in clinical trials, are detected only after the drug has been marketed and used by large numbers of patients, possibly including some more sick or less sick than those in the clinical trial population. Although a more comprehensive risk profile may begin to emerge only after widespread use, possibly leading to drug withdrawal, various confounding ‘lifestyle’ and other factors can often make determination of causation a tricky and complex process (Corrigan, 2002).

Since few relevant data are in the public domain, it is generally not possible to assess whether any inadequacies in non-clinical safety evaluation have contributed to the withdrawal of drugs on grounds of safety (see Chapter 1 and Appendix 1). However, quite a few
drugs developed during the era of extensive pre-authorization toxicological evaluation have been withdrawn, suggesting that the observed human toxic responses were not clearly predicted by animal studies (or by clinical trials). Retrospective analysis might, in some cases, have shown weak signals. Newly introduced toxicological tests (for QT prolongation, ICH S7B) might have been helpful in the case of fluoroquinolone antibiotics, e.g. grepafloxacin (Owens and Ambrose, 2002).

Types of adverse drug reactions and their toxicological investigation

Type A adverse reactions are dose dependent and predictable based on the pharmacology (and kinetics) of the drug; about 80 per cent of all reported ADRs are type A. On the other hand, according to Knowles et al. (2000), idiosyncratic drug reactions (type B) cannot be explained on the basis of the conventional pharmacology of the drug, and although they may be dose dependent in susceptible individuals, they do not occur at any dose in most patients. Type B adverse reactions can affect any organ system; they include IgE-mediated anaphylaxis and allergy. In addition, there can be reactive metabolite effects, such as:

- Hypersensitivity-syndrome reactions. These are usually defined by the triad of fever, skin eruption and internal organ involvement; such reactions have been associated with anticonvulsants (phenytoin, phenobarbital, carbamazepine and lamotrigine), sulphonamide antibiotics, dapsone, minocycline and allopurinol.

- Serum-sickness-like reactions. These are distinct from serum sickness and defined by fever, rash, usually urticaria, and arthralgias occurring 1–3 weeks after drug exposure, immune complexes, hypocomplementaemia, vasculitis and renal lesions being absent; drugs implicated in such reactions are antibiotics such as cefaclor, cefprozil and minocycline.

- Drug-induced lupus. This is characterized by frequent musculoskeletal complaints, fever and weight loss, pleuropulmonary involvement in more than half of the patients, with no cutaneous findings of lupus erythematosus, symptoms and serological changes generally occurring more than a year after starting therapy. Drugs implicated in the causation of drug-induced lupus include procainamide, isoniazid, hydralazine, chlorpromazine, methyldopa and penicillamine.

Type B ADRs are generally unpredictable and often result in the post-marketing failure of otherwise useful therapies. Examples of recent cases include zileuton, trovafloxacin, troglitazone and felbamate. Zileuton, a 5-lipoxygenase inhibitor authorized in the USA (but not in the EU) to prevent and relieve the symptoms of chronic asthma, has been shown to cause liver toxicity in some patients (Sorkness, 1997). It is not widely used, owing to the need for four times daily dosing and the requirement for liver function monitoring during the first few months of therapy. US post-marketing surveillance data on trovafloxacin, a fluoroquinolone antibiotic, indicated the possibility of serious hepatic reactions and pancreatitis, leading to significant restrictions in its use (Ball, 2000; Bertino and Fish, 2000). Troglitazone and felbamate are discussed below.

Deciding whether a particular ADR is likely to be type A or B may not be possible until data on symptomatology and patient characteristics (e.g. race, gender, genetic polymorph-
isms, concurrent disease status and medications) have been analysed on a cohort of affected patients. Key factors in clinical assessment are the temporal relationship between drug intake and the appearance of symptoms, skin tests and provocation tests (Ring and Brockow, 2002).

**Toxicological investigation of type A adverse drug reactions**

Toxicological evaluation of predictable type A ADRs would normally be targeted on pharmacodynamic–pharmacokinetic mechanisms. Causative factors could include one or more of:

- Excess systemic exposure leading to an exaggerated pharmacodynamic response
  - impaired elimination of active drug due to interactions (e.g. P450 inhibition)
  - slow active drug clearance due to genetic polymorphism.
- Formation of toxic metabolite specific to humans.
- Other biological human-specific mechanisms, such as different receptor specificity or sensitivity.

The starting point for toxicological evaluation would be the establishment of appropriate *in vitro* and/or *in vivo* models based on pharmacodynamic and kinetic considerations. Studies would then be focused on a case-by-case basis on appropriate endpoints, such as enzyme induction/inhibition, interactions and genetic variability.

Variations in genes coding for drug-metabolizing enzymes, drug receptors and drug transporters have been associated with individual variability in the efficacy and toxicity of drugs. It is difficult to disentangle the contribution of environmental and genetic factors in an individual patient. Genotyping can predict the extremes of phenotypes, but less definable factors (such as other variant genes) and environmental factors (such as smoking and diet) contribute to the patient’s phenotype.

Possibly the most actively researched area in genetic polymorphism relates to the contribution of genetically determined variability in drug metabolizing enzymes to inter-patient differences in response to drugs (Lu, 1998; Meyer, 2000). The most important clinically relevant drug-metabolizing enzyme polymorphisms relate to:

- CYP2C9, e.g. warfarin, tolbutamide, phenytoin, glipizide, losartan;
- CYP2D6, e.g. antiarrhythmics, antidepressants, antipsychotics, opioids;
- CYP2C19, e.g. omeprazole, diazepam;
- N-acetyltransferase, e.g. sulphonamides, procainamide, hydralazine, isoniazid;
- UDP-glucuronosyltransferase (UGT), e.g. irinotecan.

Toxicological studies in extensive- and poor-metabolizer animals (particularly non-rodents) may be helpful in assessing the possible safety impact of some human genetic polymorphisms.
Toxicological investigation of type B adverse drug reactions

Little is known with certainty about the mechanisms involved in the majority of idiopathic ADRs. Circumstantial, rather than direct, evidence suggests that drug reactive metabolites (DRMs) are responsible for most type B ADRs. The current hypothesis (Knowles et al., 2000) linking reactive metabolites to ADRs suggests that there are basically three possibilities for further reaction of a newly formed DRM:

- Deactivation by nucleophiles and radical scavengers, e.g. glutathione, epoxide hydrolases.
- Reaction with macromolecules leading to cytotoxicity.
- Hapten formation – covalent binding to proteins and altered protein triggers an immune response.

A more recent ‘multiple determinant hypothesis’ states that the low frequency (<1/5000) of idiosyncratic drug toxicity is due to the requirements for occurrence of multiple critical and discrete events. The principal determinants of these events are proposed to be: chemical properties (including potential for DRM production), patient exposure, environmental and genetic factors (Li, 2002).

The generation and fate of reactive metabolites are determined by activating, inactivating and precursor-sequestering enzymes. In turn, these enzymes are controlled by long-term induction and repression, as well as the short-term control of post-translational modification and low-molecular-weight activators and inhibitors. The effectiveness of such enzyme systems in preventing DRM-mediated toxicity relates principally to their subcellular compartmentalization and isoenzyme multiplicity. Susceptibility differences to DRM-related toxic challenges between species and individuals are frequently thought to be causally linked to differences in these control factors (Oesch et al., 1990).

The formation of an epoxide DRM has been postulated for some anticonvulsants (carbamazepine, phenobarbital and phenytoin), and enhanced individual susceptibility was thought to be related to a deficiency of epoxide hydrolase. More recent studies have thrown some doubt on this hypothesis and alternative DRMs have been postulated such as free radicals and an orthoquinone for phenytoin and an iminoquinone for carbamezepine (Knowles et al., 2000).

The potential of a drug to stimulate idiosyncratic reactions probably relates more to its chemical structure than its pharmacological mechanism. If biotransformation can yield products containing structural elements such as quinones, phenols, acyl halides, aromatic or hydroxyl amines, then the potential for type B ADRs is increased (Petersen, 2002). Examples supporting this hypothesis are tacrine (Figure 3.1; primary aromatic amine, hepatotoxic cholinesterase inhibitor for treatment of Alzheimer’s disease superseded by drugs such as donepezil and rivastigmine (Figure 3.1) with less potential to produce DRMs) and troglitazone (can be transformed into a quinone, a property not shared by its successors such as rosiglitazone and pioglitazone; see Figure 3.1). However, an alternative to this suggestion by Peterson to explain the hepatotoxicity of troglitazone is advanced by Haskins et al., (2001); see Figure 3.2.

Some researchers believe that it is currently impossible to predict which chemical species...
will cause idiosyncratic ADRs and advocate the need for a more thorough understanding of basic drug metabolism before attempting to relate chemical species formation to biological function (Williams and Naisbitt, 2002). Many drugs that are associated with idiosyncratic toxicity contain nitrogen which is relatively easy to oxidize and many nitrogen-containing compounds undergo redox cycling, which can generate active oxygen species. Moreover, several nitrogen-containing substances, including aromatic amines, nitro compounds, hydrazines and compounds that can be oxidized to iminoquinones and related substances, have been associated with adverse reactions. However, in addition to the presence of such structural elements, various other factors, such as dose, electron density and patient susceptibility, may play a role (Uetrecht, 2002).

Modern biochemical, molecular and immunochemical techniques have enabled identification of specific target proteins of xenobiotic covalent binding, and it is apparent that binding is not random but rather selective in its targeting. Selective protein binding may correlate better with target organ toxicity, and evidence on several compounds (e.g. paracetamol, halothane and 2,5-hexanedione) tends to support this proposal (Cohen et al., 1997).

Many idiosyncratic reactions appear to have an immunological aetiology; hapten formation followed by uptake, antigen processing and T-cell proliferation appear to be the critical parts of the mechanism (Naisbitt et al., 2000, 2001; Park et al., 2000, 2001; Ju and Uetrecht, 2002). Drugs associated with a high incidence of hypersensitivity appear to be capable of ready formation of DRMs, but this appears not to apply to all drugs that can form DRMs. One possible explanation is that orally administered drugs may lead to oral tolerance in most individuals through mechanisms similar to those found with orally administered antigens (i.e. interaction with gut-associated lymphoid tissue of the small intestine). Following oral administration of the NSAID diclofenac (Figure 3.1) to rats, a series of diclofenac protein adducts (55 to 142 kDa) were detected in small intestine homogenates. Two of the adducts were identified as aminopeptidase N (CD13) and sucrase-isomaltase and were localized primarily in the mid-villus and villus-tip enterocytes and also in the dome overlying Peyer’s patches. Similar adducts were detected in villus-tip enterocytes of rats treated with halothane or paracetamol. It is possible that such intestinal protein adducts of drugs formed in gut-associated lymphoid tissue may lead to down-regulation of drug-associated allergic reactions in many individuals (Ware et al., 1998).

The liver is the principal site of drug metabolism and it is a common target for idiosyncratic drug reactions (Jaeschke et al., 2002). In the case of immune reactions directly involving leukocytes, the enzyme system most likely to be responsible for the formation of reactive metabolites is the NADPH oxidase/myeloperoxidase system found in neutrophils and monocytes. In addition to the proposed hapten/T-lymphocyte pathway, other mechanisms may exist, such as molecular mimicry (caused by a common alteration in the processing and presentation of antigens due to non-drug stimuli such as viruses) and direct alteration of the class II major histocompatibility (MHC) molecule by a DRM leading to a graft versus host reaction (Uetrecht, 1997). Hepatitis of the type triggered by drugs such as halothane, tienilic acid and dihydralazine appears to have a range of immunological features, including dose independence, immune system manifestations such as fever and eosinophilia, delay between drug treatment and disease onset, shortened delay on rechallenge and occasional presence of serum autoantibodies (Beaune and Lecoeur, 1997; Dansette et al., 1998; Castell, 1998). Genetic imbalance between bioactivation and detoxification pathways, as well as reduced cellular defences against DRMs due to disease or concomitant drug therapy, may act as risk factors to the onset and severity of ADRs (Hess and Rieder, 1997).
Tolcapone, a catechol-\(O\)-methyltransferase inhibitor used for treatment of motor fluctuations in Parkinson's disease, had been associated with numerous cases of hepatotoxicity, including three cases of fatal fulminant hepatic failure. The structurally similar drug entacapone appears not to stimulate this type of toxic response. In vitro studies suggest that CYP P450-catalysed oxidation of amine and acetylamine human tolcapone metabolites leads to formation of reactive intermediates that may form covalent adducts to hepatic proteins, resulting in damage to liver tissues (Smith et al., 2003). Use of the broad-spectrum antiepileptic drug felbamate has been limited due to reports of treatment-related hepatotoxicity and aplastic anaemia. It has been proposed that bioactivation leads to formation of a highly reactive electrophilic metabolite, atropaldehyde (ATPAL), an \(\alpha,\beta\)-unsaturated aldehyde, that is capable of forming covalent protein adducts in vivo. In vitro studies on ATPAL support this hypothesis and suggest that both direct covalent binding with critical macromolecules and indirect interference with cellular detoxication mechanisms may be involved (Dieckhaus et al., 2002; Kapetanovic et al., 2002).

Cutaneous reactions are the most frequently occurring adverse reactions to drugs, with the incidence amongst hospitalized patient normally ranging from 1 to 3 per cent, although the frequency of cutaneous reactions to specific drugs may exceed 10 per cent. Antifungals and anticonvulsants are most commonly associated with adverse skin reactions. The varied nature of cutaneous reactions, even with specific drugs, indicates a multiplicity
of mechanisms. Dividing cutaneous reactions into four mechanistically based categories has been proposed (Svensson et al., 2001):

- immediate-type immune-mediated reactions;
- delayed-type immune-mediated reactions;
- photosensitivity reactions;
- autoimmune syndromes.

Important predisposing factors are viral infection and female gender. To account for the latter, gender differences in T-cell activation and proliferation have been proposed, as well as the increased prevalence of skin diseases, such as systemic lupus erythematosus and photosensitivity (Rademaker, 2001). Although it has been well established for many years that sulphonamide drugs produce delayed-type cutaneous reactions that severely limit their therapeutic utility, in spite of much research, the mechanisms involved remain unclear (Reilly and Ju, 2002).

Gender differences in the metabolism of xenobiotics in both humans and laboratory animals may provide an additional confounding factor in the toxicological evaluation of ADRs. Xenobiotic metabolism by male rats can reflect human metabolism when CYP1A or CYP2E are involved, because there is strong regulatory conservation of these isoforms between rodents and humans. Unfortunately, the identification of sex-dependent differences in metabolism by rats does not generally translate to humans. The major confounding factor is that CYP2C, a major subfamily in the rat, which is expressed in a sex-specific manner, is not found in humans. In addition, sex-specific isoforms of cytochrome P450 appear to be absent in humans, indicating that the commonly used male rat is unlikely to be an accurate model for the prediction of sex-related differences in metabolism in humans (Mugford and Kedderis, 1998).

Rare, but serious, ADRs associated with some carboxylic-acid-containing drugs have been investigated in a variety of toxicological studies. Drugs containing a carboxylic acid moiety can be bioactivated by two distinct pathways: by UGT-catalysed conjugation with glucuronic acid, resulting in the formation of acyl glucuronides, or by acyl-CoA synthetase-catalysed formation of acyl-CoA thioesters. Both metabolites are electrophilic species that, if they escape inactivation by glutathione, can acylate target proteins. Although there is accumulating evidence that acyl glucuronides can alter cellular function by a variety of mechanisms, including haptenation of peptides, glycation or acylation of specific proteins, and direct stimulation of neutrophils and macrophages, the roles of acyl-CoA are less clear (Boelsterli, 2002).

In summary, the underlying immunological and other mechanisms of idiosyncratic drug toxicity are poorly understood, which greatly hampers the development of suitable animal models. Each type of immunopathology is thought to result from a specific cluster of immunologic and biochemical phenomena, and other factors such as genetic predisposition, metabolic variability and concomitant diseases. It may thus be difficult to find common mechanisms that lead to nonclinical models able to predict specific types of systemic hypersensitivity reaction. However, there are already adequate models for detecting drugs that induce contact sensitization, and it may be possible to develop screening tests for signs indicative of a general hazard for immune-based reactions (Hastings, 2001). New technolo-
gies, such as toxicogenomics, proteomics and metabolomics, offer the potential to identify human toxicants during drug development (Steiner and Anderson, 2000; Castle et al., 2002; Wilkins, 2002). Use of transgenic models, including mouse models that have been ‘humanized’ in various ways, e.g. by changing receptors that regulate cytochrome P450 enzymes, may be helpful in understanding mechanisms of toxicity involved in the causation of idiosyncratic ADRs (Wolf and Henderson, 1998; Rudmann and Durham, 1999; Xie and Evans, 2002). One approach may be to classify compounds in various ways, e.g. relating to reactivity/adduct formation, and to identify predictive markers for screening purposes (Uetrecht, 2000, 2001; Nelson, 2001). In respect of candidate drugs in the discovery phase, key questions (Baillie and Kassahun, 2001) might include:

- What level of covalent binding to proteins in liver (or other relevant target tissue) might be acceptable?
- Which protein targets are critical to the viability of the cell, and which are likely to be adducted by a specific reactive intermediate?
- Is there a dose threshold for toxicity, and what factors influence this?
- Is it possible to predict the human immunogenicity of potential drug–protein conjugates?

Sites of metabolic activation within a new drug candidate series have been rapidly identified by trapping the reactive intermediates formed on incubation with rat and human liver microsomes as their glutathione conjugates and mass-spectral characterization of these thiol adducts. A strategy of iterative structural modification of the chemical series in order to block bioactivation sites led to a significant reduction in the propensity to undergo metabolic activation, as evidenced by decreases in the irreversible binding of radioactivity to liver microsomal material on incubation of tritium-labelled compounds in vitro (Samuel et al., 2003).

Proteomics might also be employed as a diagnostic tool for recognizing a drug signature in a tissue exhibiting an adverse response (Hellmond et al., 2002). The development of these technologies in respect of predicting or diagnosing ADRs is still at an early stage, and major breakthroughs seem unlikely in the short term.

**Examples of toxicological investigation of adverse drug reactions**

Many drugs that produce ADRs have been investigated using nonclinical models; each drug will have a particular aetiology and the nonclinical investigations are likely to be tailored to this. Two cases are described below in order to provide examples of typical approaches.

**Hepatotoxicity of thiazolidinediones**

Thiazolidinediones (TZDs), particularly troglitazone (Figure 3.2), when used in the treatment of type 2 diabetics, are associated with sporadic clinical hepatotoxicity not predicted by conventional animal studies. In isolated rat and human hepatocytes, multiprobe fluores-
Scence analysis showed disruption of mitochondrial activity as an initiating event followed by increased membrane permeability, calcium influx and nuclear condensation. Other effects included treatment-related hepatic enzyme leakage, decreased reductive metabolism and cytoplasmic ATP depletion. The relative potency of TZDs for causing these effects was: troglitazone > pioglitazone > rosiglitazone. The authors conclude that hepatic alterations in vitro are characteristic of TZDs, with only quantitative differences in subcellular organelle dysfunction (Haskins et al., 2001).

**Indinavir-induced hyperbilirubinemia**

Up to 25 per cent of patients receiving indinavir (Figure 3.3) for the treatment of HIV infection develop unconjugated hyperbilirubinemia, prompting discontinuation of treatment in some patients.

![Figure 3.2 Structure of troglitazone](image1)

![Figure 3.3 Structure of indinavir](image2)

The hypothesis that the side effect occurred by indinavir-mediated inhibition of bilirubin UGT was tested in two ways: (a) evaluation of patients with Gilbert’s polymorphism (reduced hepatic UGT activity); (b) studies in the Gunn rat model of UGT deficiency. Serum bilirubin increased by a mean of 0.34 mg/dl in indinavir-treated patients lacking Gilbert’s polymorphism versus 1.45 mg/dl in those who were heterozygous or homozygous for the mutant allele.

Indinavir competitively inhibits UGT activity ($K_i = 183 \mu M$) and concomitantly induces hepatic bilirubin UGT mRNA and protein expression. Although saquinavir also competitively inhibits UGT activity, there is no association with hyperbilirubinemia, probably because of the higher $K_i$ (360 $\mu M$) and the lower therapeutic plasma levels compared with indinavir.

Oral indinavir increased plasma bilirubin in wild-type and heterozygous Gunn rats, the mean rise being markedly greater in the latter group.
The findings of the various studies are considered to indicate that clinical hyperbilirubinemia results from indinavir-mediated inhibition of bilirubin conjugation (Zucker et al., 2001).

Conclusions

Pharmacotoxicology has developed out of all recognition since the thalidomide tragedy (Dally, 1998); mandatory prelicensing nonclinical tests of largely standardized design perform a critical function in drug discovery and development. With careful study design, species selection, study performance and data interpretation, taking particular account of kinetic and metabolic differences between the animal models and man, the tests are generally highly predictive for most toxic responses encountered in clinical trials. A recent industry survey found that human gastrointestinal, cardiovascular and haematological toxicities were best predicted, whereas cutaneous effects were most difficult to detect in animal models.

Low-frequency ADRs of the type likely to be identified at the post-marketing stage after large numbers of patients have been exposed to a particular drug are, unsurprisingly, not well predicted by standard animal studies. The majority of these ADRs have been found to result from unanticipated pharmacodynamic and kinetic factors, and the mechanisms involved are often amenable to nonclinical investigation using in vitro and/or in vivo systems. On the other hand, a minority of ADRs are idiosyncratic and the processes leading to their causation are poorly understood. The formation of DRMs is thought to be involved; although this hypothesis is consistent with some of the evidence, a major anomaly is why more patients do not suffer such adverse reactions, since DRMs are thought to be produced by many drugs.

There is good evidence that most idiosyncratic ADRs have an immunological basis; the need for suitable animal models for mechanistic studies has been stressed by many investigators, but this may be difficult to achieve given the clinical variability in susceptibility, thought to be due to the genetic, metabolic and concomitant disease status of the individual patient. Use of genetically modified animals could lead to suitable models, and toxicogenomics and proteomics may also provide useful contributions. The process of screening new drug candidates for metabolic activation potential and introducing structural modifications in order to block bioactivation sites seems to show great potential for minimizing idiosyncratic ADRs. However, even with the benefit of these new technologies, developing reliable models for the prediction and evaluation of idiosyncratic ADRs possibly represents the greatest current challenge in pharmacotoxicology.

Acknowledgements

The valuable contribution of Dr C. J. Powell in terms of developing the outline for this chapter and reviewing an early draft is acknowledged. Thanks are also due to Professor A. D. Dayan for his most helpful comments.

References


REFERENCES


REFERENCES


