## Toxicity Testing: search for an *in-vitro* alternative


## Introduction

Prior to introduction to the clinical setting, new pharmaceuticals must undergo rigorous toxicity testing during development to ensure their safety. Various toxicity tests are conducted depending on the length of potential drug usage in humans. Acute oral toxicity tests involve the administration of a single large dose to two different mammalian species and the short terms effects in organs and tissues observed over several weeks to determine specific toxic effects and the mode of toxicity. Subchronic and chronic toxicity tests are performed for pharmaceuticals intended to be used in humans for periods of less than, or greater than, 3 months respectively. Therefore, daily administration of the agent in two mammalian species enables longer term toxicity and pathology to be evaluated, including for example the occurrence of neurological and haematological abnormalities.  In addition, if an agent is intended for use over periods of greater than 6 months carcinogenicity studies and mutagenicity studies may be necessary, as well as teratogenicity evaluation if appropriate (OECD Summary Guideline Report, 1979).


## Traditional animal testing

Traditionally drug safety testing has been achieved through both *in vitro* tests and testing involving large numbers of animals and is an extremely costly, time-consuming process. Indeed, a large backlog of chemicals on the Western market are still awaiting testing to properly characterize toxicity. A recent study by Ukelis *et al* estimated that only 5,000 of 100,000 available chemicals have undergone appropriate testing (Ukelis *et al,* 08).

The use of animals in drug safety testing is, however, a sensitive area and subject to an ever increasing demand to reduce the numbers used. This was first suggested by Russell and Burch (1959) as the 3 R's principle, to reduce, refine and replace use of laboratory animals.

Indeed, according to Home Office Statistics of Scientific Procedures on Living Animals (2007) toxicology testing and safety and efficacy evaluation accounted for 416,400 procedures, 39% lower than in 1995. However, although widely publicised, toxicology studies account for only 13% of animal testing, with large numbers of animals also used in areas such as genetic studies, production of biological materials and pharmaceutical research. Whilst numbers of animals involved in toxicity testing appear high, almost 80% of these are rodents and less than 1% non-human primates.

In addition to reducing the numbers of animals used in individual trials, the need for improved communication to reduce unnecessary duplication of tests was recognised, and the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) subsequently formed in 1990 with the aim of minimising animal testing

and duplication of human trials without compromising safety. Greater similarity between animal testing guidelines in different countries is also necessary. Current EU guidelines for acute oral toxicity testing require tests to be conducted in 2 or more species, preferably rodents. However, in Japan 2 species are required (including one non-rodent) and in the US 3 species are necessary, one of which must be non-rodent (Ukelis *et al,* 08).

**Progress of animal testing**

In recent years, progress has been made within toxicity testing *in vivo* in terms of reducing numbers of animals used but also in the extent of tests used. Traditionally the LD50 assay has been used, a measure of the dose required to result in death in 50% of the animals in the study. An alternative to this assay was first suggested by the British Toxicology Society in 1984 and was based on administering a series of fixed doses and relying on the observation of clear signs of toxicity rather than the endpoint of the assay being death. This was introduced as a test guideline in 1992 and found to use fewer animals and cause less suffering than the LD50 test. The LD50 test for acute oral toxicity has since been abolished following an OECD Joint Meeting in 2001 (OECD Guideline 420, 2001).

A recent initiative between 18 European countries has also reviewed the use of acute toxicity tests in pharmaceutical development. The analysis indicates that acute toxicity test results are in practice not used to determine doses to be administered in phase I clinical trials, nor are they useful in terminating production of drugs in development. Conclusions reached by the group determined that acute toxicity tests are no longer often the first test carried out and are generally less useful than data generated from other tests routinely incorporated into toxicity testing (Robinson *et al,* 2008).

**Concordance of animal data with human toxicity**

Besides ethical reasons for reducing animal testing, lack of predictability of human toxicity from animal trials has been cited as a concern. Whilst generally animal trials can give good indications as to human toxicity, many studies have reported lack of concordance, termed interspecies uncertainty (Price *et al,* 08). False positive results from animal trials may result in withdrawal of a drug from trials or the use of subclinical doses. Conversely, and possibly a worse scenario involves false negative results which may lead to unexpected human toxicity. This was dramatically illustrated in the TGN 1412 phase I clinical trial in 2006. TGN 1412 is a monoclonal antibody thought to have an anti-inflammatory effect via activation of T regulatory cells and potential use in the treatment of leukaemia and autoimmune diseases. However, the opposite effect was seen in human trials producing massive systemic inflammatory responses (Suntharalingam *et al,* 2006). A subsequent investigation concluded that these serious adverse effects were not predicted in humans following apparently adequate preclinical animal tests (MHRA Report, 2006).

Whilst TGN 1412 is a particularly dramatic example, lack of concordance between human and animal trials has been previously reported in other studies. A report conducted in 2000 examined data from 12 pharmaceutical companies concerning 150 compounds (including 221 reported human toxicity events) and aimed to better understand the concordance between human toxicity and that observed in laboratory animals. Concordance overall was found to be 71% when using rodent and non-rodent test species. However, using non-rodents alone reduced concordance to 63% and just 43% when using rodents only, possibly highlighting potential issues where guidelines do not enforce the use of non-rodent species in combination with rodents, as is the case for example in the EU (Ukelis *et al,* 2008).

When extrapolating safety tests on animals to set human doses, this is often achieved by taking a dose associated with a particular toxicity and dividing by a series of uncertainty factors. This is designed to take into account species differences and the correlation of toxicity generally increasing with increasing body weight (Price *et al,* 08). In the United States for example, interspecies uncertainty is generally assigned a ratio of 10:1 (to allow for humans potentially being up to 10 times more sensitive to the agent than test animals). However, a recent study by Price *et al,* (2008) investigated the validity of this ratio in terms of anti-neoplastic agents and concluded that it may be inappropriate for a number of drugs. Particularly when comparing human and mouse toxicity ratios, the mean ratio of 54 agents was 20:1. In contrast, human/dog toxicity ratios were most closely correlated, with a mean value of 3.5:1, well within the value of 10 typically used in safety tests. Additionally, the report illustrates that human toxicity data is derived from individuals with cancer, who are often also elderly and may therefore be compromised in terms of health and resistance to toxicity compared with the general population.

**Toxicity testing of prodrugs**

Safety testing of clinical agents is increasingly difficult if the drug in question requires metabolism to the active form for example by cytochrome P450 enzymes. Whilst only a limited number of pharmaceuticals are administered in a prodrug form, the vast majority of drugs still require the presence of P450 enzymes to aid detoxification and elimination form the body, illustrated in

Table **1**.

Table 1– Important anti-cancer agents metabolised by the cytochrome P450 enzyme system

| Agent | Activating enzymes | Deactivating enzymes | Source |
|---|---|---|---|
| Cyclophosphamide | 2B6, 2C19 | 3A4, 2A6 | [a, b] |
| Dacarbazine | | 1A1, 1A2, 2E1 | [a] |
| Etoposide | | 3A4, 1A2, 2E1 | [a] |
| Flutamide | 1A2 | | [b] |
| Paclitaxel | | 2C8, 3A | [a] |
| Tamoxifen | 2D6, | 3A4, 3A5, 2B6, 2C9, 2C19 | [b] |
| Tegafur (prodrug of 5-Flurouracil | 2A6, 2C9 | | [b] |
| Vincristine | | 3A | [a] |

[a] McFayden et al., 2003.
[b] van Schaik., 2005
Agents are used in the treatment of the following tumours:
Cyclophosphamide (sarcoma, breast, ovarian, leukaemias), Dacarbazine (metastatic melanoma),Etoposide (lymphoma, osteosarcoma, testicular, small cell lung), Flutamide (prostate), Paclitaxel (ovarian, breast, non-small cell lung), Tamoxifen (breast), Tegafur (metastatic colorectal), Vincristine (lymphoma).

*In vivo,* enormous variation may be seen due to species differences as well as within species due to the highly polymorphic nature of the enzymes. Dietary and environmental factors can also significantly affect P450 enzymes, for example CYP1A2 is significantly higher in smokers than non-smokers, and only 2 weeks of a high protein/low carbohydrate diet can significantly increase levels of CYP1A1 (Lin *et al,* 2006). In addition, gene copy number variation (CNV) where large amounts of DNA (>1kb) are duplicated or deleted can also greatly affect gene expression and inter-individual variation (Freeman *et al,* 2006). Consequently, only 30-60% of common drug therapy is effective, at least in part due to genetic variation between individuals (Ingelman-Sundberg and Rodriguez-Antona, 2005). Both providing a source of enzymes such as P450 within an *in vitro* test system and allowing for large inter-individual variability, present challenges to drug safety testing in the laboratory.

Cytochrome P450 enzymes have traditionally been sourced from preparations such as S9 liver extracts where following processing, a specific liver homogenate fraction is a rich source of P450 enzymes. Sources include rat liver, often induced with agents such as phenobarbital to increase protein and activity levels (Waxman *et al,* 1990). However, phenobarbital itself has been reported to be mutagenic (Deutsch *et al,* 2001) and again, species differences occur within the P450 system. S9 preparations are also available from human liver sources, removing the challenge of species variation; however, cost and availability of such preparations may limit their use in *in vitro* safety testing. Lack of inducibility may also limit use of S9 preparations as a representative P450 system, as *in vivo* some agents including the commonly used chemotherapeutic agent cyclophosphamide, are able to upregulate relevant P450 enzymes, thereby inducing their own metabolism in subsequent exposures (Afsharian *et al,* 2007).

Primary liver tissue as a source of enzymes for *in vitro* testing is a highly attractive option, however, difficulties with ethics, cost and obtaining sufficient quantities greatly limit use. In addition, primary tissue in culture has a low proliferative rate, which in toxicity studies may be a disadvantage as proliferation can be used as a sensitive marker of adverse events such as hepatotoxicity (O'Brien *et al,* 2006).

Numerous hepatocyte cell lines have also been used, one of the best characterised of which is HepG2 (Aden *et al,* 1979). However, differences have been reported between sources of HepG2 (Hewitt and Hewitt, 2004), culture conditions and passage number (Wilkening and Bader, 2003). Additionally, the presence of one genotype only, which has been suggested to lead to low or even absent levels of P450's may be a limiting factor, however, other reports conclude HepG2 to be a suitable model for biotransformation studies (Brandon *et al,* 2006). More recently, new hepatocyte cell lines have shown promise in drug toxicity testing, including HBG BC2, with levels of some important metabolic enzymes such as CYP3A4 considerably higher than HepG2 (Fabre *et al,* 2003). Advantages of utilising cell lines such as HepG2 include their availability, low cost, lack of ethical issues and inducibility (Brandon *et al.,* 2006).

**New *in-vitro* testing models**

Recently, a method has been developed within our lab to enable culture of HepG2 cells into three-dimensional spheroids (Figure 1). These spheroids are able to survive significantly longer in culture than monolayer cells and are more representative of human liver in terms of liver specific functions (Xu *et al,* 2003).

Figure 1– Mature HepG2 liver spheroids after 6 days in culture
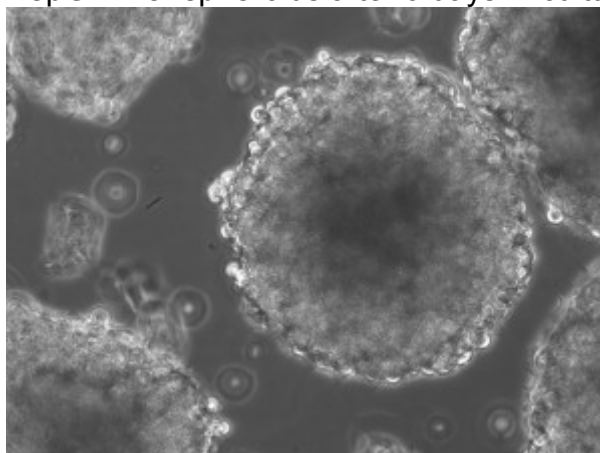


Following monolayer culture until confluency, HepG2 cells can be cultured on a gyrotatory shaker to enable the formation of spheroids. (x10 magnification)

Co-culture of these liver spheroids with other cell types has enabled chemotherapeutic damage from pro-drugs such as cyclophosphamide to be studied, with results obtained resembling damage seen in patients who have received chemotherapeutic treatment *in vivo* (Figure 2).

Figure 2 – Comparison of *in vitro* modelling of cyclophosphamide treatment with effects seen in patients who have received *in vivo* therapy.
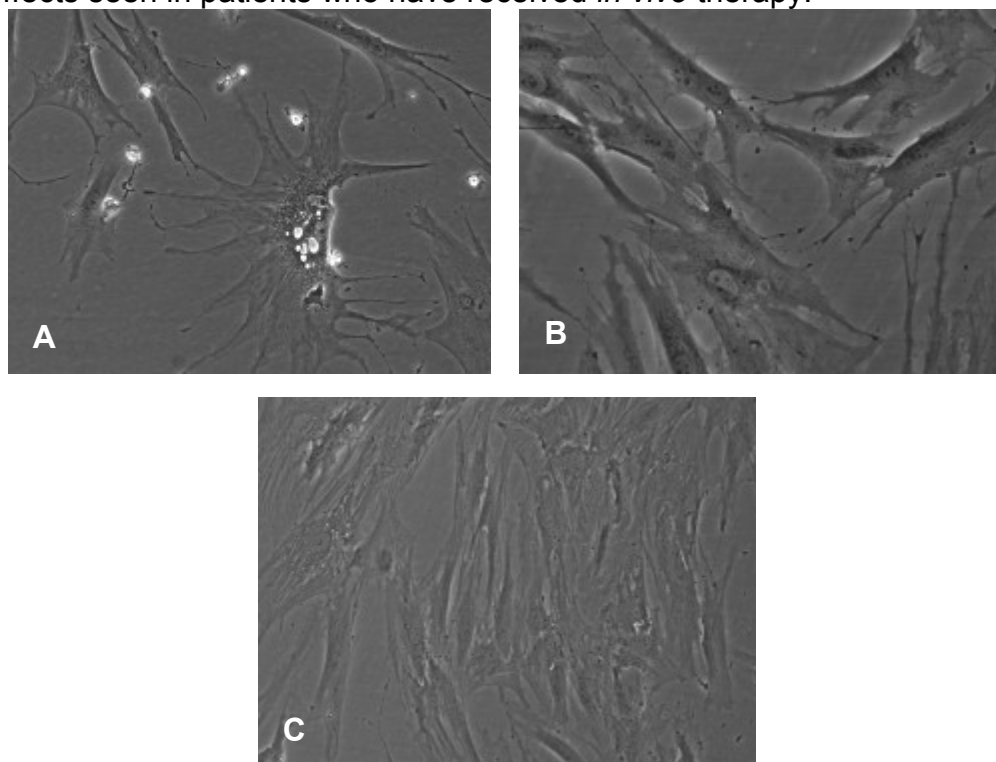


Mesenchymal Stem Cells (MSC) from a patient previously treated with CY and fludarabine *in vivo* (A), MSC exposed to CY *in vitro* in the presence of HepG2 liver spheroids (B), control untreated MSC (C). (Images A and C x10 magnification, Image B x20 magnification)

Similarly a novel *in vitro* system utilising normal primary human cells from 5 different organs together with a breast cancer cell line was developed by Li *et al,* (2004) to allow study of the effects of an agent on more than one cell type simultaneously. The model uses a 'well within a well' system where 6 cell types can be cultured within separate wells in specialised media, but all 6 wells are within one larger well, allowing it to be flooded with an agent able to access all internal wells. Consequently the effect of the agent on several organ systems can be studied, as well as interactions between the different systems, resulting in a much more realistic model of the *in vivo* situation.

Developments in detection of subtle indicators of toxicology may also enable progress in *in vitro* toxicity testing. Many traditional *in vitro* tests rely upon detecting parameters involved in lethal toxicity events such as apoptosis or necrosis and often lack sensitivity. A study by O'Brien *et al,* (2006) investigated the possibility of using new technology to detect earlier subtle, sub-lethal indicators of toxicity *in vitro*. The High Content Screening (HCS) assay involves automated quantitative epifluorescence microscopy to monitor live cells *in vitro* in real time, examining parameters such as plasma membrane permeability, nuclear size, cellular mitochondrial membrane potential, concentration of intracellular free calcium and cell number. Cell number was the first parameter to be affected in 56% of drugs testing positive for hepatotoxicity. Altered nuclear size was found to be the most precise, with changes caused by 70% of hepatotoxic drugs. In many cases, this was seen

as a significant decrease in nuclear size (up to 50%); however, a small increase was seen with some drugs, and even a 50% increase following acetaminophen exposure (O'Brien *et al,* 2006).

The assay was found to have greater than 90% concordance with human toxicity and offer the potential for high throughput screening. Although the assay failed to detect some toxicities it does demonstrate the need and benefits of continued research into *in vitro* toxicity testing methods as new technologies are developed and improved.

In addition to methods for toxicity testing arguments for adapting the level of testing depending on the setting in which the pharmaceutical will be used could also be put forward. For example, in patients with a terminal illness it may be appropriate to minimise safety testing as the risks of toxicity events, particularly longer term are greatly reduced and may not outweigh possible clinical benefits.


**Conclusions**

Much progress has been made both in *in vitro* and *in vivo* toxicity testing. Following the 3R's principle, numbers of animals involved in testing have been greatly reduced and as increased knowledge of toxicity has been gained, methods of testing have been improved. Additionally, developments in technology have and are currently enabling the introduction of increasingly specific and sensitive *in vitro* alternatives.

In conclusion, currently there is still a place for animal testing within the toxicity setting, with a well documented history and providing opportunity to study the entire organism. However, many alternative *in vitro* methods are now available and in development and whilst not currently a complete replacement for animal testing, can be used prior to and in some cases to complement. With growing developments in knowledge and technology, *in vitro* tests should become more predictive of the *in vivo* situation and should be used wherever possible, although care must be taken to consider possible limitations of models used.

**References:**

Aden, D.P., Fogel, A., Plotkin, S., Damjanov, I. and Knowles, B.B., (1979). Controlled synthesis of HBsAg in a differentiated human liver carcinoma-derived cell line. *Nature.* **282**:615-616

Afsharian, P., Terelius, Y., Hassan, Z*., et al.,* (2007). The Effect of Repeated Administration of Cyclophosphamide P450 2B in Rats. *Clinical Cancer Research.* **13**(14):4218-4224

Brandon, E.F.A., Bosch, T.M., Deene, M.J., *et al.,* (2006). Validation of *in vitro* cell models used in drug metabolism and transport studies; genotyping of

cytochrome P450, phase II enzymes and drug transporter polymorphisms in the human hepatoma (HepG2), ovarian carcinoma (IGROV-1) and colon carcinoma (CaCo-2, LS180) cell lines. *Toxicology and Applied Pharmacology*. **211**:1-10

Deutsch, W.A., Kukreja, A., Shane, B., *et al.,* (2001). Phenobarbital, oxazepam and Wyeth 14,643 cause DNA damage as measured by the Comet assay. *Mutagenesis.* **16**(5):439-442

Fabre, N., Arrivet, E., Trancard, J., *et al.,* (2003). A new hepatoma cell line for toxicity testing at repeated doses. *Cell Biology and Toxicology.* **19**:71-82

Freeman, J.L., Perry, G.H., Feuk, L., *et al.,* (2006). Copy number variation: New insights in genome diversity. *Genome Research.* **16**:949-961

Hewitt, N.J. and Hewitt, P., (2004). Phase I and II enzyme characterisation of two sources of HepG2 cell lines. *Xenobiotica.* **34**(3):243-256

Home Office Statistics of Scientific Procedures on Living Animals (UK) (2007). Available from http://scienceandresearch.homeoffice.gov.uk/animal-research/publications-and-reference/statistics/. [Accessed 22/09/08]

Ingelman-Sundberg, M. and Rodriguez-Antona, C., (2005). Pharmacogenetics of drug-metabolizing enzymes:implications for a safer and more effective drug therapy. *Philosophical Transactions of the Royal Society B*. **360**:1563-1570

Li, A.P., Bode, C. and Sakai, Y., (2004). A novel in vitro system, the integrated discrete multiple organ cell culture system, for the evaluation of human drug toxicity: comparative cytotoxicity of tamoxifen towards normal human cells from five major organs and MCF-7 adenocarcinoma breast cancer cells. *Chemico-Biological Interactions.* **150**:129-136

Lin H.J., (2006). CYP Induction-Mediated Drug Interactions: *in Vitro* Assessment and Clinical Implications. *Pharmaceutical Research*. **23**(6):1089-1116

Medicines and Healthcare products Regulatory Agency document (MHRA) - 'Investigations into adverse incidents during clinical trials of TGN1412.' (2006) Available from http://www.mhra.gov.uk/NewsCentre/Pressreleases/CON2023822. [Accessed 22/09/08]

O'Brien, P.J., Irwin, W., Diaz, D., *et al.,* (2006). High concordance of drug-induced human hepatotoxicity with in vitro cytotoxicity measured in a novel cell-based model using high content screening. *Archives of Toxicology.* **80**:580-604

OECD/OCDE 420 Guideline Document. OECD Guideline for Testing of Chemicals. Acute Oral Toxicity – Fixed Dose Procedure. (2001)

http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/OECD/OECD_GL420.pdf
[Accessed 21/01/09]

OECD Guidelines for the Testing of Chemicals .
Summary of Considerations in the Report from the OECD Expert Groups on Short Term and Long Term Toxicology (1979). Available from
http://oberon.sourceoecd.org/vl=335656/cl=12/nw=1/rpsv/ij/oecdjournals/1607310x/v1n4/s1/p1 [Accessed 21/01/09]

Olson, H., Betton, G., Robinson, D., *et al*., (2000). Concordance of the toxicty of Pharmaceuticals in Humans and Animals. *Regulatory Toxicology and Pharmacology*. **32**:56-67

Price, P.S., Keenan, R.E. and Swartout, J.C., (2008). Characterising interspecies uncertainty using data from studies of anti-neoplastic agents in animals and humans. *Toxicology and Applied Pharmacology. Article in Press*

Robinson, S., Delongeas, J., Donald, E., *et al.,* (2008). A European pharmaceutical company initiative challenging the regulatory requirement for acute toxicity studies in pharmaceutical drug development. *Regulatory Toxicology and Pharmacology.* **50**:345-352

Russell, W.M.S. and Burch, R.L., (1959). The principle of Humane Experimental Technique. London, UK. Available online from http://altweb.jhsph.edu/publications/humane_exp/chap2a.htm. [Accessed 21/01/09]

Suntharalingam, G., Perry, M.R., Ward, S., *et al.,* (2006). Cytokine Storm in a Phase 1 Trial of the Anti-CD28 Monoclonal Antibody TGN1412. *The New England Journal of Medicine.* **355**:1018-1028.

Ukelis, U., Kramer, P.J., Olejniczak, K. and Mueller, S.O., (2008). Replacement of in vivo acute oral toxicity studies by *in vitro* cytotoxicity methods: Opportunities, limits and regulatory status. *Regulatory Toxicology and Pharmacology.* **51**:108-118

Waxman, D.J., Morrissey, J.J., Naik, S. and Jauregui, H.O., (1990). Phenobarbital induction of cytochromes P-450. *Journal of Biochemistry.* **271**:113-119

Wilkening, S. and Bader, A., (2003). Influence of Culture Time on the Expression of Drug-metabolizing Enzymes in Primary Human Hepatocytes and Hepatoma Cell Line HepG2. *Journal of Biochemistry and Molecular Toxicology.* **17**(4):207-213

Xu, J., Ma, M. and Purcell, W.M., (2003). Charactrisation of some cytotoxic endpoints using rat liver and HepG2 spheroids as *in vitro* models and their application in hepatotoxicity studies. I. Glucose metabolism and enzyme release as cytotoxic markers. *Toxicology and Applied Pharmacology.* **189**:100-111