EDITORIAL

In vitro prediction of organ toxicity: the challenges of scaling and secondary mechanisms of toxicity

Jan G. Hengstler[1] · Anna-Karin Sjögren[2] · Daniele Zink[3] · Jorrit J. Hornberg[4]

Recently, a study about in vitro prediction of nephrotoxicity has been published (Sjögren et al. 2018). The authors aimed to develop a first-line screening assay that can be applied to guide chemical design in early drug discovery (Johansson et al. 2019). They used a proximal tubular epithelial cell line expressing the organic anion transporter 1 (ciPTEC-OAT1). Renal proximal tubular epithelial cells (PTEC) are frequently affected by compound-induced nephrotoxic effects and are, therefore, widely applied in kidney-specific in vitro methods (Tiong et al. 2014). ciPTEC-OAT1 were incubated with 38 compounds known to be nephrotoxic in humans while 24 were non-nephrotoxic (Sjögren et al. 2018). The assay was performed in a 96-well plate format to facilitate the screening throughput required in drug discovery programs. After an incubation period with test substances for 48 h several fluorophores were added: Hoechst to stain nuclei, LysoTrackerGreen for visualization of lysosomes, MitoTracker Orange to image mitochondria, TOTO-3 as a dead cell indicator and Phalloidin to stain actin cytoskeleton. Subsequently, the cells were fixed by paraformaldehyde and, after high-content imaging, 200 cellular phenotypic parameters were analyzed. The most predictive phenotypic parameters were identified using a machine learning approach by mathematical feature selection algorithms and related to changes of the actin cytoskeleton, mitochondria, nuclear morphology and

integrity of the cell membrane (Sjögren et al. 2018). The 62 compounds were tested over a range of six concentrations covering the maximal plasma concentration (C_{max}) of the individual drugs. The authors evaluated at what concentration range, in relation to the C_{max}, the assay delivered optimal predictivity. They found that whereas an increasing number of toxic compounds were identified as the test concentration increased, there were no false positives in the assay up to a test concentration of 200-fold the C_{max}. Overall, a sensitivity of 66% and a specificity of 100% were reported.

The study (Sjögren et al. 2018) initiated a discussion with further scientists working with in vitro tests to predict organ toxicity (Zink 2019; Sjögren and Hornberg 2019), which led to a consensus concerning two aspects that may be of general interest: first, the necessity to use higher concentrations in the culture medium compared to blood concentrations in vivo; second, the need to improve our possibilities to predict toxicity caused by 'secondary mechanisms'.

The use of high in vitro concentrations compared to human C_{max} in plasma

A frequently used strategy to predict organ toxicity in vitro is to use a validation set of compounds known to cause an increased risk of organ toxicity for a certain dosing schedule (positive reference compounds) and compounds that do not lead to an increased risk (negative reference compounds). A compound tested in vitro at a specific concentration either leads to a positive or negative test result. The in vitro test is typically performed at a concentration range around and above the plasma peak concentrations (C_{max}) in humans. By comparison of the in vitro test result (positive, negative) to the clinical safety profile (toxic, non-toxic), statistical performance metrics, such as sensitivity, specificity or accuracy can be calculated. Importantly, the use of relatively high concentrations, often 20- to 200-fold higher than the therapeutic C_{max}, leads to a better accuracy than lower

✉ Jan G. Hengstler
hengstler@ifado.de

1 IfADo at TU Dortmund, Dortmund, Germany

2 Cardiovascular, Renal and Metabolism Safety, Clinical Pharmacology and Safety Sciences, R&D, AstraZeneca, Gothenburg, Sweden

3 NanoBio Lab and Innovations in Food and Chemical Safety Programme, Agency for Science, Technology and Research, Singapore, Singapore

4 Respiratory, Inflammation and Autoimmunity Safety, Clinical Pharmacology and Safety Sciences, R&D, AstraZeneca, Gothenburg, Sweden

Springer

concentrations. Also in the present study (Sjögren et al. 2018), while most toxic drugs were identified at concentrations around or below 20-fold the therapeutic C_{max}, testing at higher concentrations (up to 200-fold the C_{max}) achieved a very good sensitivity and specificity. It is relevant to note that, for screening assays that are to be applied during early drug discovery, it is essential to maintain high specificity, to avoid deselection of potentially promising compound series (whereas high sensitivity becomes more relevant towards selection of actual candidate drugs). Therefore, Sjögren et al. suggested a threshold of 200-fold, since beyond that threshold false positives were observed. Still, this leads to the question, whether such high concentrations are still relevant for the human in vivo situation. From a cell biology point of view this objection is more than justified. Nevertheless, the ultimately critical criterion is if the test condition leads to a correct prediction of the clinical situation or not. Also in previous studies on hepatotoxicity (O'Brien et al. 2006; Persson et al. 2013; Albrecht et al. 2019), a similar observation was made that higher concentrations than C_{max} better differentiated between positive and negative control compounds. There are several potential explanations for this. For example, toxicity is clinically often observed after chronic exposure and accumulation of subtle insults, while tests in the lab usually rely on more dramatic insults upon short-term incubation. While this is certainly correct, it may not be sufficient to explain the situation. A recent study with primary human hepatocytes incubated 30 test compounds for 1, 2 and 7 days and analyzed cytotoxicity (Gu et al. 2018). As expected, longer incubation periods led to lower EC_{50} values. However, this decrease of EC_{50} values after 7 days compared to shorter incubations occurred also for negative test compounds and did not allow a better differentiation between positive and negative controls by just using longer incubation periods (Albrecht et al. 2019). This further exemplifies that it is relevant to maintain a low false positive rate.

Another potential reason for the high concentrations required in vitro may be that target organs may in reality experience much higher concentrations in certain cell compartments than indicated by measurements of the plasma C_{max}, or compounds require active transport and/or metabolism for the toxicity to become apparent (Will and Dykens 2014). Down-regulation of transporters for drugs and other xenobiotics and of drug-metabolizing enzymes applies to PTEC (Jenkinson et al. 2012; Lash et al. 2008; Tiong et al. 2014) as well as to hepatocytes (Gomez-Lechon et al. 2014; Vildhede et al. 2015; Vinken and Hengstler 2018). In fact, although lots of efforts are spent currently on the development of improved cell models, there is no PTEC or hepatocyte model available that would in vitro fully recapitulate the in vivo performance, in particular with respect to the quantitative expression levels of transporters and drug-metabolizing enzymes. In light of these differences it would

be rather surprising if cells in vitro would show the same concentration response as in vivo. Apart from developing improved cell models, the exploration of scaling factors that account for different transporter and enzyme expression levels in vitro and in vivo may be useful. Such scaling factors have been successfully applied recently in physiologically based modelling of drug clearance (Chan et al. 2019; Vildhede et al. 2015).

Down-regulation of drug-metabolizing enzymes in vitro may in part explain why ifosfamide was not detected as toxic in the present study (Sjögren et al. 2018), because ifosfamide requires metabolic activation by cytochrome P450 enzymes. However, a more likely explanation is that this false negative is a result of the fact that the set of endpoints analyzed in the high content imaging assay did not reflect the mechanism of toxicity for ifosfamide: the compound does induce gene expression of the oxidative stress marker HMOX1 in ciPTEC-OAT1. Apart from other potential solutions as discussed above, this limitation may be overcome by the additional inclusion of a drug metabolizing system into the test in future (Godoy et al. 2013).

The challenge of 'secondary mechanisms' of organ toxicity

Some cell types are particularly susceptible to the toxic effects of chemicals, such as PTEC of the kidney or hepatocytes of the liver. A relatively large fraction of nephrotoxic or hepatotoxic compounds act by causing damage directly to these cell types. However, adverse effects on the kidney or liver can also be induced by mechanisms that initially do not involve PTEC or hepatocytes. Examples are altered hemodynamics or crystal formation (kidney) as well as compromised cholangiocytes or obstructions of the biliary tract (liver). Here we refer to that as mechanisms of secondary toxicity. Of course, identification of compounds acting by primary mechanisms represents an important milestone. Nevertheless, the final goal is to identify all toxic compounds, independent of their mechanism of action. A practical implication for in vitro testing is that tests for primary mechanisms, e.g. based on cultivated PTEC or hepatocytes, should be accompanied by a battery of tests that detect the secondary mechanisms. However, relatively little is known about how accurately these secondary toxicities can be identified.

In the field of nephrotoxicity, a categorization with three sub-groups of compounds has been introduced (Li et al. 2013, 2014; Kandasamy et al. 2015; Su et al. 2016).

Group 1 Nephrotoxicants that damage PTEC.

Group 2 Nephrotoxicants that damage the kidney by other mechanisms, e.g. by altering renal hemodynamics, or by

crystal formation; examples are indinavir, which exerts nephrotoxicity because of crystallization and stone formation; lisinopril, an angiotensin-converting enzyme inhibitor which causes nephrotoxicity by hemodynamic effects.

Group 3 Non-nephrotoxic compounds, i.e. compounds that up to a known C_{max} do not cause nephrotoxicity in humans.

To test the hypothesis that the assay would be more sensitive to drugs that directly damage PTEC, the present study of Sjögren and colleagues (2018) also analyzed the predictive performance for those drugs separately (i.e. compared group 1 and group 3). This resulted in a performance metrics of 75% (15 of 20 drugs) for sensitivity and 100% for specificity (24 of 24 drugs).

The situation becomes more complex for the group 2 compounds: of the 18 compounds in that category, 10 tested positive and 8 negative (Sjögren et al. 2018). Previous studies observed a relatively high percentage of negative results for group 2 compounds up to very high concentrations of 1000 µg/ml (Li et al. 2013, 2014; Su et al. 2016). Differences in outcomes between the studies may relate to the analyzed compounds and the endpoints. In the future, it will be important to identify and study larger numbers of group 2 compounds. These studies should include the currently available PTEC-based tests to allow a direct comparison; moreover, in vitro tests of secondary mechanisms of nephrotoxicity, e.g. precipitation tests and tests for vascular dysfunction should be included. These studies should show with which sensitivity and specificity group 2 compounds can be identified.

Currently, much effort is invested in the development of in vitro systems (Sachinidis et al. 2019; Leist et al. 2017) particularly to predict nephrotoxicity (Adler et al. 2016; Li et al. 2013, 2014; Su et al. 2016; Kandasamy et al. 2015), hepatotoxicity (Godoy et al. 2015, 2016; Proctor et al. 2017; Grinberg et al. 2014), cardiotoxicity (Archer et al. 2018; Sampaio et al. 2016; Reis-Mendes et al. 2017; Chaudhuri et al. 2016), developmental toxicity (Waldmann et al. 2014, 2017), and the improvement of physiologically based pharmacokinetic modeling and other modeling strategies (Ghallab et al. 2016; Chan et al. 2019). The results of the here discussed high content screening assay (Sjögren et al. 2018) show that it is possible to implement a high-throughput screening assay in early drug discovery to guide chemical design away from nephrotoxicity. Compounds acting via damaging PTEC can be identified with a relatively high accuracy, and further studies are needed to clarify whether addition of further readouts to a test battery improves the identification or mechanistic investigation of compounds acting by secondary mechanisms. In vitro research in different areas, nephrotoxicity, hepatotoxicity, cardiotoxicity, neurotoxicity and developmental toxicity, seems to be confronted with similar challenges.

Compliance with ethical standards

References

Adler M, Ramm S, Hafner M, Muhlich JL, Gottwald EM, Weber E, Jaklic A, Ajay AK, Svoboda D, Auerbach S, Kelly EJ, Himmelfarb J, Vaidya VS (2016) A quantitative approach to screen for nephrotoxic compounds in vitro. J Am Soc Nephrol 27(4):1015–1028. https://doi.org/10.1681/ASN.2015010060

Albrecht W, Kappenberg F, Brecklinghaus T et al (2019) Prediction of human drug-induced liver injury (DILI) in relation to oral doses and blood concentrations. Arch Toxicol 93(6):1609–1637. https://doi.org/10.1007/s00204-019-02492-9

Archer CR, Sargeant R, Basak J, Pilling J, Barnes JR, Pointon A (2018) Characterization and validation of a human 3D cardiac microtissue for the assessment of changes in cardiac pathology. Sci Rep 8(1):10160. https://doi.org/10.1038/s41598-018-28393-y

Chan JCY, Tan SPF, Upton Z, Chan ECY (2019) Bottom-up physiologically-based biokinetic modelling as an alternative to animal testing. Altex 36(4):597–612. https://doi.org/10.14573/altex.1812051

Chaudhuri D, Artiga DJ, Abiria SA, Clapham DE (2016) Mitochondrial calcium uniporter regulator 1 (MCUR1) regulates the calcium threshold for the mitochondrial permeability transition. Proc Natl Acad Sci USA 113(13):E1872–E1880. https://doi.org/10.1073/pnas.1602264113

Ghallab A, Cellière G, Henkel SG et al (2016) Model-guided identification of a therapeutic strategy to reduce hyperammonemia in liver diseases. J Hepatol 64(4):860–871. https://doi.org/10.1016/j.jhep.2015.11.018

Godoy P, Hewitt NJ, Albrecht U et al (2013) Recent advances in 2D and 3D in vitro systems using primary hepatocytes, alternative hepatocyte sources and non-parenchymal liver cells and their use in investigating mechanisms of hepatotoxicity, cell signaling and ADME. Arch Toxicol 87(8):1315–1530. https://doi.org/10.1007/s00204-013-1078-5

Godoy P, Schmidt-Heck W, Natarajan K et al (2015) Gene networks and transcription factor motifs defining the differentiation of stem cells into hepatocyte-like cells. J Hepatol 63(4):934–942. https://doi.org/10.1016/j.jhep.2015.05.013 (**Erratum in: J Hepatol 2016 Feb;64(2):525–526**)

Godoy P, Widera A, Schmidt-Heck W et al (2016) Gene network activity in cultivated primary hepatocytes is highly similar to diseased

mammalian liver tissue. Arch Toxicol 90(10):2513–2529. https://doi.org/10.1007/s00204-016-1761-4

Gomez-Lechon MJ, Tolosa L, Conde I, Donato MT (2014) Competency of different cell models to predict human hepatotoxic drugs. Expert Opin Drug Metab Toxicol 10(11):1553–1568. https://doi.org/10.1517/17425255.2014.967680

Grinberg M, Stöber RM, Edlund K et al (2014) Toxicogenomics directory of chemically exposed human hepatocytes. Arch Toxicol 88(12):2261–2287. https://doi.org/10.1007/s00204-014-1400-x

Gu X, Albrecht W, Edlund K et al (2018) Relevance of the incubation period in cytotoxicity testing with primary human hepatocytes. Arch Toxicol 92(12):3505–3515. https://doi.org/10.1007/s00204-018-2302-0

Jenkinson SE, Chung GW, van Loon E, Bakar NS, Dalzell AM, Brown CD (2012) The limitations of renal epithelial cell line HK-2 as a model of drug transporter expression and function in the proximal tubule. Pflugers Arch 464(6):601–611. https://doi.org/10.1007/s00424-012-1163-2

Johansson J, Larsson MH, Hornberg JJ (2019) Predictive in vitro toxicology screening to guide chemical design in drug discovery. Curr Opin Toxicol 15:99–108. https://doi.org/10.1016/j.cotox.2019.08.005

Kandasamy K, Chuah JKC, Su R et al (2015) Prediction of drug-induced nephrotoxicity and injury mechanisms with human induced pluripotent stem cell-derived cells and machine learning methods. Sci Rep 5:12337. https://doi.org/10.1038/srep12337

Lash LH, Putt DA, Cai H (2008) Drug metabolism enzyme expression and activity in primary cultures of human proximal tubular cells. Toxicology 244(1):56–65. https://doi.org/10.1016/j.tox.2007.10.022

Leist M, Ghallab A, Graepel R et al (2017) Adverse outcome pathways: opportunities, limitations and open questions. Arch Toxicol 91(11):3477–3505. https://doi.org/10.1007/s00204-017-2045-3

Li Y, Oo ZY, Chang SY et al (2013) An in vitro method for the prediction of renal proximal tubular toxicity in humans. Toxicol Res 2(5):352–362. https://doi.org/10.1039/c3tx50042j

Li Y, Kandasamy K, Chuah JK et al (2014) Identification of nephrotoxic compounds with embryonic stem-cell-derived human renal proximal tubular-like cells. Mol Pharm 11(7):1982–1990. https://doi.org/10.1021/mp400637s

O'Brien PJ, Irwin W, Diaz D, Howard-Cofield E, Krejsa CM, Slaughter MR, Gao B, Kaludercic N, Angeline A, Bernardi P, Brain P, Hougham C (2006) High concordance of drug-induced human hepatotoxicity with in vitro cytotoxicity measured in a novel cell-based model using high content screening. Arch Toxicol 80(9):580–604. https://doi.org/10.1007/s00204-006-0091-3

Persson M, Løye AF, Mow T, Hornberg JJ (2013) A high content screening assay to predict human drug-induced liver injury during drug discovery. J Pharmacol Toxicol Methods 68(3):302–313. https://doi.org/10.1016/j.vascn.2013.08.001

Proctor WR, Foster AJ, Vogt J, Summers C, Middleton B, Pilling MA, Shienson D, Kijanska M, Ströbel S, Kelm JM, Morgan P, Messner S, Williams D (2017) Utility of spherical human liver microtissues for prediction of clinical drug-induced liver injury. Arch Toxicol 91(8):2849–2863. https://doi.org/10.1007/s00204-017-2002-1

Reis-Mendes A, Gomes AS, Carvalho RA, Carvalho F, Remião F, Pinto M, Bastos ML, Sousa E, Costa VM (2017) Naphthoquinoxaline metabolite of mitoxantrone is less cardiotoxic than the parent compound and it can be a more cardiosafe drug in anticancer therapy. Arch Toxicol 91(4):1871–1890. https://doi.org/10.1007/s00204-016-1839-z

Sachinidis A, Albrecht W, Nell P et al (2019) Road map for development of stem cell-based alternative test methods. Trends Mol Med 25(6):470–481. https://doi.org/10.1016/j.molmed.2019.04.003

Sampaio SF, Branco AF, Wojtala A, Vega-Naredo I, Wieckowski MR, Oliveira PJ (2016) p66Shc signaling is involved in stress responses elicited by anthracycline treatment of rat cardiomyoblasts. Arch Toxicol 90(7):1669–1684. https://doi.org/10.1007/s00204-015-1583-9

Sjögren AK, Hornberg JJ (2019) Compound selection and annotation to validate the predictivity of in vitro toxicity assays for use in drug discovery, in response to Commentary by Dr. Zink (Zink D. Arch Toxicol (2019)). Arch Toxicol 93(1):225–226. https://doi.org/10.1007/s00204-018-2359-9

Sjögren AK, Breitholtz K, Ahlberg E, Milton L, Forsgard M, Persson M, Stahl SH, Wilmer MJ, Hornberg JJ (2018) A novel multi-parametric high content screening assay in ciPTEC-OAT1 to predict drug-induced nephrotoxicity during drug discovery. Arch Toxicol 92(10):3175–3190. https://doi.org/10.1007/s00204-018-2284-y

Su R, Xiong S, Zink D, Loo L-H (2016) High-throughput imaging-based nephrotoxicity prediction for xenobiotics with diverse chemical structures. Arch Toxicol 90(11):2793–2808. https://doi.org/10.1007/s00204-015-1638-y

Tiong HY, Huang P, Xiong S, Li Y, Vathsala A, Zink D (2014) Drug-induced nephrotoxicity: clinical impact and preclinical in vitro models. Mol Pharm 11(7):1933–1948. https://doi.org/10.1021/mp400720w

Vildhede A, Wisniewski JR, Noren A, Karlgren M, Artursson P (2015) Comparative proteomic analysis of human liver tissue and isolated hepatocytes with a focus on proteins determining drug exposure. J Proteome Res 14(8):3305–3314. https://doi.org/10.1021/acs.jproteome.5b00334

Vinken M, Hengstler JG (2018) Characterization of hepatocyte-based in vitro systems for reliable toxicity testing. Arch Toxicol 92(10):2981–2986. https://doi.org/10.1007/s00204-018-2297-6

Waldmann T, Rempel E, Balmer NV et al (2014) Design principles of concentration-dependent transcriptome deviations in drug-exposed differentiating stem cells. Chem Res Toxicol 27(3):408–420. https://doi.org/10.1021/tx400402j

Waldmann T, Grinberg M, König A et al (2017) Stem cell transcriptome responses and corresponding biomarkers that indicate the transition from adaptive responses to cytotoxicity. Chem Res Toxicol 30(4):905–922. https://doi.org/10.1021/acs.chemrestox.6b00259

Will Y, Dykens J (2014) Mitochondrial toxicity assessment in industry—a decade of technology development and insight. Expert Opin Drug Metab Toxicol 10(8):1061–1067. https://doi.org/10.1517/17425255.2014.939628

Zink D (2019) Comment on Sjögren et al. (2018) A novel multi-parametric high-content screening assay in ciPTEC-OAT1 to predict drug-induced nephrotoxicity in drug discovery. Arch Toxicol 92(10):3175-3190. Arch Toxicol 93(1):221-223. https://doi.org/10.1007/s00204-018-2327-4