Toxicology in Vitro 32 (2016) 347–361

Contents lists available at ScienceDirect

Toxicology in Vitro

journal homepage: www.elsevier.com/locate/toxinvit

Prevalidation of the ex-vivo model PCLS for prediction of respiratory toxicity

A. Hess [a,1], L. Wang-Lauenstein [b,c,1], A. Braun [b,c], S.N. Kolle [a], R. Landsiedel [a], M. Liebsch [e], L. Ma-Hock [a], R. Pirow [e], X. Schneider [d], M. Steinfath [e], S. Vogel [a], C. Martin [d,*,1], K. Sewald [b,c,**,1]

[a] BASF SE, Experimental Toxicology and Ecology, 67056 Ludwigshafen, Germany
[b] Department of Pre-clinical Pharmacology and In Vitro Toxicology, Fraunhofer Institute for Toxicology and Experimental Medicine (ITEM), Nikolai-Fuchs-Str. 1, 30625 Hannover, Germany
[c] German Center for Lung Research, Biomedical Research in Endstage and Obstructive Lung Disease Hannover (BREATH), Germany
[d] Institute of Pharmacology and Toxicology, Medical Faculty of RWTH Aachen University, Wendlingweg 2, Aachen, Germany
[e] German Federal Institute for Risk Assessment (BfR), Max-Dohrn-Str. 8-10, 10589 Berlin, Germany

ARTICLE INFO

Article history:
Received 1 July 2015
Received in revised form 5 January 2016
Accepted 7 January 2016
Available online 9 January 2016

Keywords:
Acute inhalation toxicity
Respiratory toxicity
Prevalidation
Reproducibility
Industrial chemicals
Two-group classification model
Precision-cut lung slices
Transferability
Submersed culture conditions

ABSTRACT

In acute inhalation toxicity studies, animals inhale substances at given concentrations. Without additional information, however, appropriate starting concentrations for in-vivo inhalation studies are difficult to estimate. The goal of this project was the prevalidation of precision-cut lung slices (PCLS) as an ex-vivo alternative to reduce the number of animals used in inhalation toxicity studies. According to internationally agreed principles for Prevalidation Studies, the project was conducted in three independent laboratories. The German BfR provided consultancy in validation principles and independent support with biostatistics.

In all laboratories, rat PCLS were prepared and exposed to 5 concentrations of 20 industrial chemicals under submerged culture conditions for 1 h. After 23 h post-incubation, toxicity was assessed by measurement of released lactate dehydrogenase and mitochondrial activity. In addition, protein content and pro-inflammatory cytokine IL-1α were measured. For all endpoints IC_{50} values were calculated if feasible. For each endpoint test acceptance criteria were established.

This report provides the final results for all 20 chemicals. More than 900 concentration–response curves were analyzed. $Log_{10}[IC_{50}\ (\mu M)]$, obtained for all assay endpoints, showed best intra- and inter-laboratory consistency for the data obtained by WST-1 and BCA assays. While WST-1 and LDH indicated toxic effects for the majority of substances, only some of the substances induced an increase in extracellular IL-1α. Two prediction models (two-group classification model, prediction of LC_{50} by IC_{50}) were developed and showed promising results.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

For many substances, inhalation is the most relevant route of occupational exposure. Regulatory application of alternative methods to animal testing, for both acute and repeated-dose inhalation toxicity studies, however, has lagged behind compared to other routes of administration, due to the complexity of the respiratory system and the diversity of local and systemic responses (Sullivan et al., 2014). It is further complicated by the fact that for some substances, the lungs are the main route of exposure but not the main target. Toxicity is then observed in other organs such as the liver, spleen, and kidney (Hope and Hope, 2012; Kennedy, 2012; Vandebriel and De Jong, 2012). Nevertheless, the respiratory tract is frequently the most sensitive and thus most important target in inhalation studies (Escher et al., 2010).

Alternative test methods accepted by the regulatory authorities have been published for in-vivo studies with topical and oral administration of substances (BeruBe et al., 2009). For acute oral toxicity, for example, several validated in-vivo alternatives are internationally accepted (e.g. acute toxic class method, fixed-dose procedure, up-and-down procedure), optionally in combination with in-vitro testing of starting concentrations (OECD TG423, 2001; OECD TG425, 2002; Guidance Document, 2001; OECD, 2010). Similar approaches – merely based on refined and reduced animal testing – have been developed for acute inhalation toxicity testing (acute toxic class method, TG436) (OECD TG433, 2004; OECD TG436, 2009). There is, however, no validated in-vitro alternative currently available for the respiratory system that is accepted by the regulatory authorities.

Inhalation of substances can induce various local and systemic effects such as respiratory irritation, acute and chronic inflammation,

* Corresponding author.
** Correspondence to: K. Sewald, Fraunhofer Institute for Toxicology and Experimental Medicine ITEM, Department of Pre-clinical Pharmacology and In Vitro Toxicology, Nikolai-Fuchs-Str. 1, 30625 Hannover, Germany.
E-mail addresses: chmartin@ukaachen.de (C. Martin), katherina.sewald@item.fraunhofer.de (K. Sewald).
[1] Contributed equally.

http://dx.doi.org/10.1016/j.tiv.2016.01.006
0887-2333/© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

and sensitization (Calzetta et al., 2011; Wanner et al., 2010). Possible health outcomes include impaired respiratory function, severe organ injury, hyperplasia, fibrosis, and respiratory allergy (Hayes and Bakand, 2010). In view of this large diversity of effects, a single (non-animal testing) alternative predicting the entire diversity of biological responses of the respiratory tract is very unlikely to be found. On the other hand, many in-vitro models resembling different parts of the respiratory tract have been reported for scientific and industrial purposes (Hansen et al., 2005; Hulette et al., 2002; Lalko et al., 2012; Larsson et al., 2009; Megherbi et al., 2012; Mitjans et al., 2008; Nelissen et al., 2009; Patlewicz et al., 2007; Python et al., 2009). Among them are single cell lines forming monolayers, such as A549, BEAS-2B, and Calu-3, as well as three-dimensional (3D) models of human-derived epithelium, such as EpiAirway™ (MatTek Corporation, Ashland, MA, USA) and MucilAir™ (Epithelix Sarl, Geneva, Switzerland) (Hirakata et al., 2010; Huang et al., 2013; Ren and Daines, 2011; Reus et al., 2013). Moreover, research has remarkably changed the general perception of organotypic tissue models such as precision-cut lung slices (PCLS), parenchymal strips, and isolated vessels and bronchi (Kroigaard et al., 2012; Trifilieff et al., 2009). These models are considered to be of great importance, since the microanatomy of the respiratory tract comprises widely varying cell types that may respond differently to the same substance (Guilliams et al., 2013). Some in-vitro and ex-vivo models have been established for prediction of organ injury, respiratory sensitization and inflammation (Huang et al., 2009; Huang et al., 2011).

The goal of this BMBF-funded project was the scientific prevalidation of rat PCLS (rPCLS) as an alternative test method for in-vivo dose range finding experiments in acute inhalation toxicity studies. The project was conducted in three independent laboratories (Fraunhofer ITEM, BASF SE, and RWTH Aachen) according to internationally agreed principles for Prevalidation Studies (Curren et al., 2006; OECD GD34, 2005) in consecutive phases over three years, aiming at a first assessment whether the rPCLS method is reliable and relevant for the intended purpose. To achieve this, the study phases comprised (i) successful transfer of methods from experienced to naïve laboratories, (ii) refinement of methods, (iii) development of robust test acceptance criteria, (iv) assessment of reproducibility within and between laboratories, and (v) proposal of a preliminary prediction model (PM) to allow prediction of expected in-vivo toxicities from the in-vitro data (Archer et al., 1997; Worth and Balls, 2001). To develop the PM, a training set of twenty chemicals was used representing different chemistry and different modes of toxicological action. The German Federal Institute for Risk Assessment (BfR) provided advice in conformity with agreed validation principles and performed independent biostatistical analyses.

2. Materials and methods

2.1. Participants and organization

The three participating laboratories of the project were the Fraunhofer Institute for Toxicology and Experimental Medicine ITEM (Germany, coordinating laboratory), BASF SE (Germany), and RWTH Aachen (Germany). Coordinator of the project was A. Braun (Fraunhofer ITEM, Germany). The German Federal Institute for Risk Assessment (BfR) provided support with biostatistics (Fig. 1). Team meetings were organized twice per year during the term of the project (2010–2013) in order to share relevant information, develop details of the protocols, discuss results, and decide on the next steps. Practical training meetings for the laboratory staff were performed to ensure correct use of the protocols. Laboratory names were anonymized for this paper.

2.2. Media, reagents and chemicals

Dulbecco's Modified Eagle's Medium Nutrient Mixture F-12 Ham (DMEM) with L-glutamine, 15 mM HEPES without phenol red, pH 7.2–7.4, penicillin/streptomycin (P/S), low melting agarose, Earle's Balanced Salt Solution (EBSS), triton X-100 (TX-100), ammonium hexachloroplatinate (AHCP), lipopolysaccharide (LPS, from *Escherichia coli* 0111:B4) and protease inhibitor cocktail were supplied by Sigma Aldrich (Munich, Germany). Phosphate buffered saline (PBS, without Ca^{2+} and Mg^{2+}) was obtained from Lonza (Wuppertal, Germany). Ammonium hexachloroplatinate, aniline, ethanol, formaldehyde, n-hexyl chloroformate, methyl methacrylate, paracetamol, paraquat, sodium dodecylsulfate, trimellitic anhydride, triton-X 100, toluene diisocyanate, zinc oxide were purchased from Sigma Aldrich (Munich, Germany). Isophorone diisocyanate and octanoyl chloride were obtained from ABCR GmbH (Karlsruhe, Germany). Acetic anhydride was supplied by Bernd Kraft GmbH (Duisburg, Germany). Glutaraldehyde and lactose were purchased from Fisher Scientific GmbH (Schwerte, Germany). Acetone and N,N-dimethylformamide were obtained from Honeywell Riedel de Haën (Seelze, Germany). Enzyme-linked immunosorbent assay (ELISA) kits were supplied by R&D Systems (DuoSets, Wiesbaden-Nordenstadt,

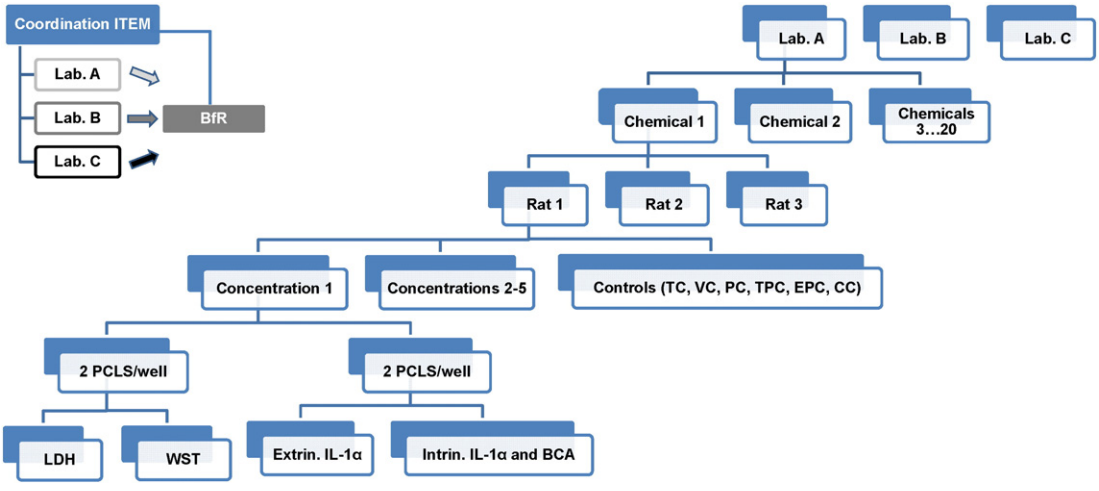


Fig. 1. The figure in the top left corner shows the management of the study. Each laboratory submitted its results to BfR for biostatistics. The study was coordinated by Fraunhofer ITEM. The big figure shows the experimental design of inter-laboratory prevalidation of rat PCLS. Each participating laboratory performed identical experiments. Each substance was tested in three independent biological runs. Twenty substances were tested. Each substance was applied at five concentrations. Experiments included controls for every endpoint method. Each sample was measured in duplicate (technical replicates). Cytotoxicity was measured by LDH and WST-1 assay. Extrinsic and intrinsic IL-1α and protein content were determined by ELISA and BCA assays. TC: tissue control with culture medium only; VC: vehicle control, if necessary; CC: chemical control of the highest tested concentration without PCLS for chemical interference test; PC: positive control for WST-1 assay with AHCP; TPC: technical positive control for LDH assay with TX-100; EPC: positive control for ELISA with LPS; extr.: extrinsic; intr.: intrinsic.

Germany). Cell proliferation reagent WST-1 and LDH cytotoxicity detection kit were purchased from Roche (Mannheim, Germany). BCA protein assay kit was obtained from Pierce (Rockford, IL, USA).

2.3. Laboratory equipment

Commonly used equipment was different in each laboratory. Major equipment was as follows: laboratory A — Krumdieck tissue slicer (Alabama Research & Development, USA), incubator (HeraCell, Thermo Scientific Heraeus, Germany), ELISA Reader (MRX, Dynatech Laboratories, USA); laboratory B — Krumdieck tissue slicer (Alabama Research and Development, USA), incubator (BDD 6220, Thermo Scientific Heraeus, Germany), Multilabel counter (Wallac 1420, Perkin Elmer, USA), ELISA reader (Sunrise, Tecan, Switzerland); laboratory C — Krumdieck tissue slicer (Alabama Research and Development, USA) or microtome (#06.12.07, Wissenschaftliche Werkstatt, RWTH Aachen, Germany), incubator (Innova CO-170, New Brunswick Scientific Co., Inc., USA), GENios ELISA Reader (Tecan, Switzerland).

2.4. Animals

Female rats [Wistar, Crl:WI (Han) nulliparous and non-pregnant] were housed under conventional and certified laboratory conditions in a regular 12-hour dark/light cycle at ambient temperature of 22 ± 2 °C and a relative air humidity of 55 ± 15%. Diet and drinking water were available ad libitum. Animals were acclimated for at least one week and sacrificed at the age of 8–10 weeks.

2.5. Preparation of rat precision-cut lung slices (rPCLS)

Animals were sacrificed by an i.p. overdose (~100 mg/kg body weight) of pentobarbital-Na and the lung was immediately dissected without damaging the lung tissue. The trachea was cannulated and the lung was filled with 10 mL/200 g body weight 37 °C-warm 1.5% low-melting agarose medium solution, either before or after the lung was separated from the animal. After polymerization of agarose to gel, lung lobes were separated, and tissue cores (8 mm in diameter) were prepared and cut into about 200–300 μm thick slices in 4 °C-cold EBSS using a Krumdieck microtome (Alabama Research and Development, Munford, AL, USA). Subsequently, tissue slices were incubated for 2 h in DMEM, supplemented with 100 units/mL penicillin (P) and 100 μg/mL streptomycin (S) in a petri dish under standard cell culture conditions (37 °C, 5% CO_2). During this incubation period the medium was exchanged every 30 min in order to remove cell debris.

2.6. Test substance selection

In this study, twenty chemicals were chosen for determination of their cytotoxic and pro-inflammatory potential using rPCLS: paraquat, isophorone diisocyanate (IPDI), triton X-100 (TX-100), toluene diisocyanate (TDI), AHCP, glutaraldehyde (GA), sodium dodecyl sulfate (SDS), octanoyl chloride (OC), n-hexyl chloroformate (HCF), formaldehyde (FA), trimellitic anhydride (TMA), acetic anhydride (Ac2O), aniline, zinc oxide (ZnO), N,N-dimethylformamide (DMF), methyl methacrylate (MMA), ethanol, acetone, paracetamol, and lactose. Out of these substances, lactose, which is well tolerated by inhalation and shows no toxic effect in-vivo, and paracetamol, which is pharmacologically active in liver, were proposed as negative controls. TX-100 was chosen as positive control substance. All other substances have known modes of toxicological action via inhalation and can chemically be divided into acrylates, aldehydes, amines, anhydrides, chloroformates, chlorides of organic acids, isocyanates, solvents, and detergents. Two metal compounds were also included. All investigated chemicals including CAS numbers, molecular weights, and tested concentrations are listed in Table 1. All compounds were purchased by one laboratory and distributed to the others. Supplement Table 1 presents – as far as known – inhalation median lethal concentration (LC_{50}) and oral median lethal dose (LD_{50}) values and the corresponding GHS classification for all chemicals. No in-vivo experiment had to be performed for the present study.

In some cases only in-vivo studies with more or less than 4 h inhalation were found. To classify these substances into GHS categories LC_{50} values were converted according to following formulas: LC_{50} value B for A hours was converted into LC_{50} estimate value D for C hours using for Gas/vapor: $D = B\sqrt{A} / \sqrt{C}$; and for Dust/mist: $D = BA / C$. In the case of performing GHS classification, enter 4 (hours) for C (GHS Guidance, 2005).

Regarding acute inhalation toxicity and according to the Globally Harmonized System (GHS), 15% of the substances (3/20) fall into category 1 ($LC_{50} \leq 0.05$ mg/L for aerosols, $LC_{50} \leq 0.5$ mg/L for vapors, and $LC_{50} \leq 100$ ppm for gases), 25% (5/20) into category 2 ($0.05 < LC_{50} \leq 0.5$ mg/L for aerosols, $0.5 < LC_{50} \leq 2.0$ mg/L for vapors, and $100 < LC_{50} \leq 500$ ppm for gases), 15% (3/20) into category 3

Table 1
Twenty test substances were selected for prevalidation of rPCLS. Substances were tested in rPCLS in all three participating laboratories. Table shows for each substance: number according to known order in inhalation toxicity, chemical name, abbreviation, chemical class, CAS-number, molecular weight, used vehicle, tested concentrations, and dilution factor. Abbr.: abbreviation; DMSO: dimethyl sulfoxide; EGDE: ethylene glycol dimethylether.

Substances	Abbr.	Class	CAS #	M [g/mol]	Vehicle	Tested concentrations [mM]	Dilution factor	Purity
Acetic anhydride	Ac2O	Anhydrides	108-24-7	102.1	–	2.3, 5.0, 10.8, 23.3, 50.0	2.15	≥99.5%
Acetone	Acetone	Solvents	67-64-1	58.1	–	458, 732, 1172, 1875, 3000	1.6	≥99.9%
Ammonium hexachloroplatinate	AHCP	Metal compounds	16919-58-7	443.9	–	0.05, 0.1, 0.2, 0.5, 1.0	2.15	≥99.9%
Aniline	Aniline	Amines	62-53-3	93.1	–	4.7, 10.1, 21.6, 46.5, 100.0	2.15	≥99.5%
Ethanol	Ethanol	Solvents	64-17-5	46.1	–	534, 854, 1367, 2188, 3500	1.6	≥99.8%
Formaldehyde	FA	Solvents	50-00-0	30.3	–	0.8, 2.0, 5.0, 12.4, 31.0	2.5	37 wt.%
Glutaraldehyde	GA	Aldehydes	111-30-8	100.1	–	0.2, 0.5, 1.1, 2.3, 5.0	2.15	25 wt.%
Isophorone diisocyanate	IPDI	Isocyanates	4098-71-9	222.2	0.5% DMSO	0.1, 0.3, 0.5, 1.2, 2.5	2.15	≥97.5%
Lactose	Lactose	Carbohydrate	63-42-3	342.3	–	4.7, 10.1, 21.6, 46.5, 100.0	2.15	≥99.0%
Methyl methacrylate	MMA	Acrylates	80-62-6	100.1	–	4.7, 10.1, 21.6, 46.5, 100.0	2.15	≥98.5%
n-Hexyl chloroformate	HCF	Chloroformates	6092-54-2	164.6	0.5% DMSO	4.7, 10.1, 21.6, 46.5, 100.0	2.15	≥96.5%
N,N-dimethylformamide	DMF	Solvents	68-12-2	73.1	–	381, 610, 977, 1563, 2500	1.6	≥99.9%
Octanoyl chloride	OC	Chlorides of organic acids	111-64-8	162.7	1% ethanol	4.6, 7.3, 11.7, 18.8, 30.0	1.6	≥98.5%
Paracetamol	Paracetamol	Analgetics	103-90-2	151.2	–	4.2, 9.1, 19.5, 41.9, 90.0	2.15	
Paraquat	Paraquat	Pesticides	1910-42-5	257.2	–	0.4, 1.3, 4.0, 12.7, 40.0	3.16	≥97.5%
Sodium dodecyl sulfate	SDS	Detergents	151-21-3	288.4	–	0.05, 0.1, 0.2, 0.5, 1.0	2.15	≥99.0%
Trimellitic anhydride	TMA	Anhydrides	552-30-7	192.1	0.5% DMSO	3.1, 4.9, 7.9, 12.5, 20.0	1.6	97%
Triton X-100	TX-100	Detergents	9002-93-1	647.0	–	0.02, 0.05, 0.11, 0.23, 0.50	2.15	≥99.0%
Toluene diisocyanate	TDI	Isocyanates	584-84-9 & 26471-62-5	174.2	1% EGDE	4.7, 10.1, 21.6, 46.5, 100.0	2.15	≥94.5%
Zinc oxide	ZnO	Metal compounds	1314-13-2	81.4	0.1% acetic acid	0.1, 0.3, 0.5, 1.2, 2.5	2.15	≥99.0%

($0.5 < LC_{50} \leq 1.0$ mg/L for aerosols, $2.0 < LC_{50} \leq 10.0$ mg/L for vapors, and $500 < LC_{50} \leq 2500$ ppm for gases), 15% (3/20) into category 4 ($1.0 < LC_{50} \leq 5.0$ mg/L for aerosols, $10.0 < LC_{50} \leq 20.0$ mg/L for vapors, and $2500 < LC_{50} \leq 20{,}000$ ppm for gases), and 30% (6/20) into category 5. The criteria for category 5 are intended to enable identification of substances which are of relatively low acute toxicity hazard, but under certain circumstances may present a danger to vulnerable populations. These substances are anticipated to have an oral or dermal LD_{50} in the range of 2000–5000 mg/kg body weight and equivalent doses for inhalation.

2.7. Exposure of rPCLS to substances

Twenty substances were tested in all participating laboratories, each in three independent experiments minimum (biological replicates, $N \geq 3$). Endpoints were evaluation of membrane integrity via LDH assay, mitochondrial activity via WST-1 assay, protein content via BCA assay, and intrinsic and extrinsic IL-1α content via ELISA.

Substances were treated as "unknown" and tested according to the following procedure: (i) checking of solubility and dilution steps, and if necessary (ii) finding of appropriate solvents for heavily soluble or insoluble chemicals without additional effects of the solvent on tissue. Thus, substances were dissolved in DMEM + P/S, if necessary with an appropriate vehicle, in five different concentrations using dilution factors as given in Table 1. Then 24-well cell culture plates were filled with 500 μL of pre-warmed substance solution per required well. Two visually round and intact rPCLS per well were added. Four replicates were used for each substance concentration: two technical replicates for LDH and WST-1 assay and two technical replicates for measurement of protein and IL-1α content. The experimental design is presented in Fig. 1. After 1 h of exposure under standard cell culture conditions, a three-fold washing step was performed to remove the substances from the tissue (except for the technical positive control, see below), followed by a post-incubation period of 23 h with DMEM + P/S. Different exposure time periods have been compared before start of the study, including pulse exposure vs. permanent exposure. Finally, the combination of 1 h pulse exposure plus 23 post-incubation was selected. For this approach we obtained optimal effects on cytotoxicity using WST-1 and LDH assay without e.g. loss of LDH activity due to interferences with substances. Plates were sealed with an adhesive film to account for volatility of substances for the entire incubation time.

In parallel with the chemically treated tissue samples, a tissue control (TC) cultivated only with culture medium, a chemical control (CC) without rPCLS (substance interference test for WST-1 and LDH assays), a positive control (PC) for WST-1 assay with 200 μM or 500 μM AHCP, a technical positive control (TPC) for LDH assay with 0.1% TX-100, a positive control for ELISA (EPC) with 100 ng/mL LPS and, when necessary, a vehicle control (VC) were prepared. VC tissue slices were treated with the vehicle selected for the test item using the same concentration as in the tissue slices exposed to the test item.

2.8. Cytotoxicity assays

Cytotoxicity of substances was determined using the cell proliferation kit WST-1 and LDH cytotoxicity detection kit. For the WST-1 assay the medium was removed from the rPCLS and 250 μL of the WST-1 working reagent per well were added. After 1 h of incubation under standard cell culture conditions, 100 μL of supernatant were transferred to a 96-well plate in duplicates. Absorbance was measured at 450 nm with a reference wavelength of 600–700 nm.

For the LDH assay 50 μL of tissue culture supernatant were transferred to a 96-well plate in duplicates. Fifty μL of LDH reaction mixture were added and plates were incubated for 20 min at room temperature, protected from light. Absorbance was measured at 490 nm with a reference wavelength of 690 nm.

2.9. Protein determination

The BCA assay is based on the colorimetric detection of a bicinchoninic acid/Cu^{+} complex which is influenced by the presence of protein bonds. It is widely used for the determination of protein concentrations.

At the end of culture period, supernatant was removed and the total protein content of the remaining tissue was subsequently measured using the BCA assay. The tissue was therefore lysed by incubation with 500 μL 1% TX-100 in PBS, pH 7.4 plus 0.2% protease inhibitor cocktail at standard cell culture conditions for 1 h. The extracted total protein of the tissue was measured by using the protocol of the Pierce® BCA protein assay kit. Bovine serum albumin (included in the assay kit) was used for the standard curve.

2.10. ELISA: quantification of IL-1α

Culture supernatant was removed at the end of culture period. The remaining tissue was lysed by incubation with 500 μL 1% TX-100 in PBS, pH 7.4 plus 0.2% protease inhibitor cocktail at standard cell culture conditions for 1 h. IL-1α was measured in supernatants and lysis extracts of rPCLS using commercially available ELISA kits according to the manufacturer's protocol. For the eight-point standard curve recombinant rat IL-1α, included in the kit, was used. The lower and upper limits of quantification were at 16 pg/mL and 1000 pg/mL, respectively.

2.11. Definition of acceptance criteria

At the beginning of the project initial protocols based on existing protocols and experience were used. According to Prevalidation principles, during the pilot phase six substances shown in Table 1 (aniline, GA, lactose, MMA, paracetamol, TX-100) were used to refine the rPCLS protocols and to develop robust test acceptance criteria (ACs) for each endpoint based on data from three different laboratories. Both, the final rPCLS endpoint protocols and the biostatistically developed ACs for each endpoint protocol were then included in final standard operating procedures (SOPs) and later applied to the remaining 14 test chemicals. The next section gives a detailed description of the ACs established for each of the individual endpoints in order to assure reliable and valid assay performance in each laboratory. Only if each acceptance criterion for each individual endpoint was met the data were accepted for further evaluation and curve fitting. An experiment had to be repeated, if one of the endpoint-specific controls did not meet the established acceptance criteria.

Depending on the toxicological endpoint measured three to four different ACs per endpoint assay were defined (Table 2). In general, the ACs for all endpoint assays were established to ensure a minimum technical quality, to define endpoint related values for the negative and positive reference samples and – for the BCA and IL-1α assay – to ensure a minimum quality of the standard curves.

In detail, AC1, controlling the magnitude of technical errors, was generally used for all endpoints (except IL-1α) and was set to ≤0.3. For the cytotoxicity assays ACs were defined so as to assure sufficiently high tissue viability of non-substance-exposed reference samples. For the WST-1 assay this was achieved by defining of a minimum optical density of TC and VC equal to or greater than 0.6 (WST-1 AC2). For the LDH assay AC2 assured cellular tissue integrity under untreated conditions, thus release of LDH in TC and VC had to be ≤30% in comparison with the TPC (LDH AC3). Further ACs (WST-1 AC3; LDH AC2) demonstrated responsiveness of the tissue to an active substance. Therefore, AHCP was used in laboratory-specific concentrations for the WST-1 assay. This PC had to reduce viability to 66% minimum in comparison with the TC. To demonstrate a maximum response of rPCLS in the LDH assay, TX-100 was used as TPC (Table 2). Here, TPC had to have a minimum optical density of 1.0. The standard curve of BCA was evaluated by analysis of eight calibration standards using a three-parameter logistic

Table 2
Endpoint specific acceptance criteria. AC: acceptance criteria; OD: optical density; TC: tissue control only with culture medium; VC: vehicle control; PC: positive control for WST-1 assay with AHCP; TPC: technical positive control for LDH assay with TX-100; EPC: positive control for ELISA with LPS; BSA: bovine serum albumin; $S_{y,x}$: residual standard error; CV_{adj}: offset adjusted coefficient of variation.

Assay	Acceptance criteria	Definition
WST-1	AC1	Difference between duplicate OD measurements has to be ≤0.3
	AC2	Minimum OD of the TC and VC has to be ≥0.6
	AC3	PC has to be ≤66% of the TC and VC
LDH	AC1	Difference between duplicate OD measurements has to be ≤0.3
	AC2	Minimum OD of the TPC has to be 1.0
	AC3	TC and VC have to be ≤30% of the TPC
BCA	AC1	Difference between duplicate OD measurements has to be ≤0.3
	AC2	Mean OD of the maximum BSA concentration of standard curve has to be ≥0.8
	AC3	$S_{y,x}$ of the fit of the standard curve has to be ≤0.1
	AC4	Mean OD of the TC and VC has to be ≥0.3
IL-1α	AC1	Restricting CV_{adj} has to be ≤20%
	AC2	Mean OD of the maximum concentration of standard curve has to be ≥1.0
	AC3	$S_{y,x}$ of the fit of the standard curve has to be ≤0.1
	AC4	EPC has to be ≥140% of the TC and VC for intrin. IL-1α

regression algorithm. The upper limit of detection was defined to be higher than 0.8, which provides a standard curve covering a sufficiently high dynamic range (BCA AC2). The quality of standard curve fit was controlled by the residual standard error ($S_{y,x}$), which had to be equal to or lower than 0.1 (BCA AC3). The total protein content of two untreated tissue sections (TC and VC) was defined at a minimum optical density of 0.3 (BCA AC4). Similar quality control checks were used for IL-1α (Table 2). AC4 here was used to assure a sufficiently high anti-inflammatory cytokine response of tissue exposed to LPS. It had to result in 140% or more intrinsic IL-1α for the TC or VC. All ACs are summarized in Table 2.

2.12. Data management

All laboratories used the same EXCEL spreadsheet template, developed by BfR and Fraunhofer ITEM, to capture all test results. The structure and functions of this template were write-protected. To allow *ad hoc* decisions about the test validity at the laboratory level, the results for all acceptance criteria were calculated by the implemented functions. All data entered were submitted to the project partner BfR responsible for further analysis across laboratories. Curve fitting and statistical analyses were performed as described below.

2.13. Curve fitting

For all endpoints a sigmoid concentration–response model was fitted to the data and IC_{50} values were calculated. Curve fitting and analysis were performed automatically using the statistical computing environment R (R Development Core Team, 2010). The functions "loess" for non-parametric fitting, "drm" (Ritz, 2005) for parametric fitting, and "mrdrm" for a combination of the two were employed. For curve fitting a model-robust approach was used (Nottingham and Birch, 2000). The response (y) is predicted by a weighted sum of a parametric fitting ($\hat{y}_p$) and a non-parametric fitting ($\hat{y}_{np}$):

$$\hat{y} = (1-\lambda)\hat{y}_p + \lambda\hat{y}_{np}.$$

The non-parametric fitting applied was a local linear regression procedure. The mixing parameter λ is determined in the generalized cross-validation procedure PRESS (Nottingham and Birch, 2000). For the endpoints WST-1, BCA, and intrinsic IL-1α a parametric model was used, either a three-parametric log-logistic model or the Brain–Cousens model (Ritz, 2005). The Brain–Cousens model is given by the following equation:

$$y(x) = \frac{d - fx}{1 + \exp(b(\log(x) - \log(e)))},$$

where x denotes the dose and y(x) the response, the shape of which is determined by the parameters *b*, *d*, *f* and *e*.

Akaike's information criterion (AIC) (Akaike, 1998) was calculated for each fitting of a data set according to both models. This approach adds to the logarithm of the residual sum of squares the number of parameters used, thus penalizing smaller residual sum of squares with the number of parameters. The model yielding the smaller AIC was chosen as parametric part of the approach. For the endpoints LDH and extrinsic IL-1α a four-parametric log-logistic model was chosen (Ritz, 2005). The non-parametric fitting was performed by local linear regression (Nottingham and Birch, 2000).

2.14. Calculation of laboratory-specific and overall mean and median IC_{50} and log IC_{50}

For WST-1, BCA, and IL-1α assay the IC_{50} value refers to the concentration, at which curve fitting as described above predicts half the value of the corresponding non-exposed control. For LDH assay the IC_{50} was defined as concentration, which predicts the average of the corresponding non-exposed control and the mean value of the TPC. To determine this point a bisection procedure was applied using the R function *bisect* from the package pracma, which finds roots of univariate functions in bounded intervals (Borchers, 2015). The search for IC_{50} was performed in an interval from 0 to 5 times the maximum concentration applied. IC_{50} values were first calculated endpoint-specifically for each substance and laboratory. Majority rules were applied in this process for the three independent runs of a single laboratory. Laboratory-independent overall IC_{50} values were calculated as mean or median of the individual laboratory-specific IC_{50} values.

Median log IC_{50} values of each substance for each endpoint and laboratory were used for regression and classification of the substances. The median was employed, because it is a robust estimate of the central tendency, which minimizes potential biases originating from outliers. Additionally, in contrast to the mean the median can be calculated, even if for some curves only minimal (>) IC_{50} values have been estimated. In case only minimal IC_{50} values could be estimated for the majority of replicates, i.e. the only information obtained was that the IC_{50} was greater than a certain value, this value was used for further analysis.

2.15. Analysis of intra- and inter-laboratory variances

The standard deviation (Stdev) and coefficient of variation (Cv = Stdev / mean * 100%) of the logarithmically (decimal) transformed IC_{50} values, denoted as log-transformed (log IC_{50}), were used as measures of dispersion to describe variability of IC_{50}. This provides a first estimation of the variance per endpoint and substance within (intra-laboratory variance) and across (inter-laboratory variance) laboratories.

2.16. Two-group classification model

In the two-group classification model, substances were divided into two classes based on their in-vivo acute inhalation toxicity (toxic and non-toxic). Substances with GHS categories 1, 2, and 3 were assigned to the toxic class and substances with GHS categories 4 and 5 to the non-toxic class. To classify the results of the present study the median of log IC_{50} values from the WST-1 assay was calculated for each laboratory and each substance. The in-vitro class center of the toxic class was then calculated as of the median for all substances that had been assigned to this class in-vivo. The same was done for the non-toxic class. Finally, the mean of these two medians was defined as the

threshold for binary classification. Each substance whose median was below this threshold was classified as "toxic", whereas substances with median values above the threshold were classified as "non-toxic".

2.17. Prediction of the LC_{50} by IC_{50}

The in-vivo lethality values of the substances used in this project were used to develop a prediction model (see also Supplement Table 1). For some chemicals only inhalation LC_{50} values from studies with exposure durations different from 4 h were found. In these cases, the LC_{50} values were converted to 4-h LC_{50} values.

In order to develop a prediction model, which allows for estimation of LC_{50} by IC_{50} a linear regression analysis was performed with values from laboratory C. The log IC_{50} values of WST-1 assay were regarded as predictors and the log-transformed LC_{50} values were regarded as response:

$$\log_{10}(LC_{50}) = \beta_0 + \beta_1 \log_{10}(IC_{50}).$$

2.18. Sensitivity and specificity

Sensitivity and specificity provide measures for the ability of an assay to correctly identify positive and negative results (in the present study to correctly classify substances as toxic or non-toxic). The sensitivity was calculated as the number of correct positives (correctly identified as toxic) / (number of correct positives + number of false negatives (incorrectly identified as non-toxic)). Specificity was calculated as the number of correct negatives (correctly identified as non-toxic) / (number of correct negatives + number of false positives (incorrectly identified as toxic)).

2.19. Statistics

Statistical analysis was performed by using the one-way ANOVA (software: GraphPad Prism 4, version 4.03). The outcomes of the post-hoc treatment to control comparisons with Bonferroni adjustment of p-values to account for multiple testing are reported. Differences between treated samples and controls were considered statistically significant at the level of $p < 0.05$. Comparison of data was analyzed as indicated in the appropriate section using either parametric correlation (Pearson's product moment correlation and two-tailed P value) or non-parametric correlation (Spearman's rank correlation r).

3. Results

3.1. Acceptance criteria

Absorbance values represented the raw data in the different endpoint assays. Consequently, most of AC referred to the optical density of the samples (Table 2). Deviations from any of the acceptance criteria indicated a problem likely related to assay quality or experimental conduct. In the cytotoxicity assays, quality controls indicated when viability of the tissue was insufficiently low. Responsiveness of the tissue to an effective toxic substance was included as positive control. Based on concentration–response curves, minimum and maximum values of absorption were defined and had to be met for subsequent experiments. Further ACs were used to control the quality of standard curves as generated in the BCA and ELISA assays. These ACs (AC2 and AC3) guaranteed sufficient spreading of absorption values, providing a high dynamic range as well as a close fitting between predicted and observed responses of standard curves. An important issue was the definition of a minimum absorption value for non-treated tissue sections (tissue and vehicle controls). Instead of measuring and standardizing the thickness of slices, measurement of total protein content was used to indirectly monitor thickness of lung slices.

To ensure that the requirements (Table 2) were fulfilled, all samples were entered for evaluation, automatically checked and accepted for further analysis only if the defined criteria were met. Supplement Table 2 gives the number of samples entered for evaluation and the proportion [%] of finally accepted cases for the lastly tested substances (DMF, ethanol, acetone, formaldehyde). AC1, which represents the magnitude of technical errors such as pipetting errors for every assay, was met in ≥95% of all experiments. In the WST-1 and LDH assays AC2 and AC3 were also met in most cases (≥83%). For BCA assay the quality requirement for the standard curves (AC2 and AC3) was fulfilled in all experiments in all laboratories. The minimum optical density of untreated samples (AC4) of the BCA assay was also reached in all experiments. Fulfillment of AC2 and AC3 as quality criteria for the ELISA standard curves ranged from 92% to 100%. Experiments that did not meet all endpoint-specific AC were not used for calculation of the IC_{50} and correlation with in-vivo. These experiments were repeated and the data used for subsequent determinations of IC_{50} values. In general, fulfillment of test acceptance criteria improved over the entire project time.

3.2. Substance-induced cytotoxicity in rPCLS: log IC_{50} results for all endpoints

In each participating laboratory rPCLS were prepared and exposed to five concentrations of 20 preselected substances (Table 1). This paper reports the final results for all substances. Concentration–response curve data from the various qualified assays were used for non-linear sigmoidal curve fitting using mathematical models. Logarithmically transformed IC_{50} [μM] values (log IC_{50}) were calculated for each substance, laboratory, and endpoint (Supplement Tables 3 to 6). More than 900 curves have been fitted. Data meeting acceptance automatically were used to determine overall mean and median log IC_{50} values for each endpoint and most of the substances (Table 3).

For 18 of 20 substances defined median log IC_{50} values could be calculated for WST-1 assay for all laboratories (Supplement Table 3, Fig. 2A, Table 3). For MMA and lactose, only minimal (>) IC_{50} values were examined except for laboratory C. Median log IC_{50} values constantly covered a wide range of approximately four orders of magnitude from 1.91 (corresponding IC_{50} value: 81 μM) for TX-100 to 6.17 (corresponding IC_{50} value: 1.45 M) for ethanol. The majority of log IC_{50} values calculated by each laboratory was within the measured concentration range, with few exceptions where IC_{50} values were found to be above the maximum applied concentrations (IPDI, TMA, TDI, paracetamol, see Supplement Table 3). In three cases outliers were detected with a Grubbs test (Grubbs, 1950) (laboratory C: SDS, OC, acetone). For the calculation of the median the outliers were not excluded. For a single laboratory a defined log IC_{50} for lactose could only be obtained by extrapolation.

The second cytotoxicity assay used was the LDH assay. It finally yielded a less complete data set than the WST-1 assay. Concentration–response curves provided defined log IC_{50} values for only 11 out of 20 substances in all laboratories (Supplement Table 4, Fig. 2B, Table 3). Of those 11 defined IC_{50} values three could only be determined in one laboratory and two additional values in only two laboratories. Median log IC_{50} values covered approximately four orders of magnitude from 2.46 for SDS (corresponding IC_{50} value: 288 μM) to 6.52 for acetone (corresponding IC_{50} value: 3.3 M). After lactose and MMA exposure no effect was observed on LDH release. The reason why many substances failed to be fitted must be a matter of some considerable concern and needs explanation. Please refer to the discussion chapter.

Total protein content was quantified using BCA assay. Log IC_{50} values could also be calculated for 20 substances (Supplement Table 5, Fig. 2C, Table 3). Median log IC_{50} values ranged from 2.25 for TX-100 (corresponding IC_{50} value: 178 μM) to 6.59 for acetone (corresponding IC_{50} value: 3.9 M) (Table 3). MMA treatment in most cases had no effect on the total protein content of lung slices. IPDI and lactose treatment provided defined values in one laboratory.

Table 3
Mean and median of endpoint specific log-transformed IC_{50} [log IC_{50} (μM)] for all tested chemicals ± inter-laboratory standard deviation (Stdev) and coefficient of variation (Cv). Log IC_{50} values are given for A) WST-1 assay, B) LDH assay, C) BCA assay, and D) intrinsic IL-1α (intIL-1α) as measured by ELISA. Chemicals are ranked by increasing median log IC_{50} values. The number of each substance refers to known order in inhalation toxicity (see also Table 1). Abbreviations of chemicals are given according to Table 1. ">" values are described in the Materials and methods section.

	Chemical	Median log IC_{50}	Mean log IC_{50}	Inter-lab. Stdev	Inter-lab. Cv
A	WST-1 assay				
3	TX-100	1.91	1.89	0.05	2
7	SDS	2.42	2.39	0.06	2
5	AHCP	2.61	2.60	0.03	1
14	ZnO	2.86	2.99	0.12	4
6	GA	3.14	3.10	0.17	5
2	IPDI	3.16	3.25	0.21	7
1	Paraquat	3.53	3.34	0.15	4
10	FA	3.67	3.70	0.12	3
11	TMA	4.00	4.16	0.43	11
8	OC	4.18	4.19	0.04	1
12	Ac2O	4.19	4.19	0.04	1
9	HCF	4.32	4.23	0.26	6
13	Aniline	4.77	4.61	0.31	6
4	TDI	4.83	4.84	0.27	6
20	Paracetamol	4.94	4.92	0.20	4
16	MMA	>5.70	>5.70	NA	NA
19	Lactose	>5.70	5.56	0.07	1
18	Acetone	6.11	6.16	0.19	3
15	DMF	6.14	6.11	0.11	2
17	Ethanol	6.17	6.18	0.04	1
B	LDH assay				
7	SDS	2.46	2.46	–	–
5	AHCP	2.94	2.90	0.10	3
14	ZnO	2.99	3.09	0.32	11
2	IPDI	3.64	3.71	–	–
6	GA	3.71	3.71	–	–
1	Paraquat	4.24	4.20	0.45	11
10	FA	4.38	4.31	0.04	1
13	Aniline	4.94	5.01	0.06	1
20	Paracetamol	5.15	5.12	0.03	1
15	DMF	6.29	6.29	0.48	8
18	Acetone	6.52	6.52	0.08	1
C	BCA assay				
3	TX-100	2.25	2.24	0.04	2
7	SDS	2.63	2.63	0.08	2
5	AHCP	3.13	3.17	0.09	3
14	ZnO	3.46	3.42	0.33	9
6	GA	3.68	3.76	0.28	8
11	TMA	4.06	4.06	0.54	13
2	IPDI	4.10	2.93	0.68	16
8	OC	4.24	4.26	0.01	0
12	Ac2O	4.29	4.33	0.05	1
10	FA	4.50	4.58	0.19	4
9	HCF	4.64	4.65	0.03	1
1	Paraquat	4.84	4.80	0.31	6
13	Aniline	4.93	4.85	0.10	2
4	TDI	4.98	4.88	0.49	10
20	Paracetamol	5.28	5.25	0.25	5
19	Lactose	5.70	5.59	0.06	1
16	MMA	>5.70	>5.70	0.00	0
15	DMF	6.37	6.35	0.07	1
17	Ethanol	6.43	6.52	0.01	0
18	Acetone	6.59	6.45	0.44	7
D	intIL-1α ELISA				
3	TX-100	2.21	2.17	0.13	6
7	SDS	2.40	2.21	0.57	24
5	AHCP	2.83	2.83	0.23	8
14	ZnO	3.08	3.02	0.09	3
6	GA	3.23	3.21	0.18	6
2	IPDI	4.10	3.40	0.20	6
10	FA	3.59	3.67	0.20	5
11	TMA	4.03	4.00	0.58	14
9	HCF	4.17	4.10	0.44	11
8	OC	4.23	4.21	0.10	2
12	Ac2O	4.29	4.23	0.12	3
1	Paraquat	4.58	4.72	0.24	5
D					
13	Aniline	4.75	4.46	0.64	13
4	TDI	4.8	4.49	0.58	12
20	Paracetamol	4.84	4.88	0.29	5
16	MMA	5.42	5.09	0.49	9
19	Lactose	5.70	5.20	0.29	5
15	DMF	6.13	6.14	0.04	1
17	Ethanol	6.27	6.24	0.06	1
18	Acetone	6.30	6.20	0.57	9

3.3. IL-1α as a biomarker of inflammation

IL-1α as a biomarker of inflammation was assessed using ELISA. Intracellular IL-1α was either unchanged or decreased to levels below tissue control (Table 3D). Only paraquat very significantly increased intracellular IL-1α levels of up to 140% vs. controls (laboratory A). As intracellular IL-1α decreased with increasing cytotoxicity, it was used to determine mean and median log IC_{50} values (Supplement Table 6, Fig. 2D, Table 3D). Median values hardly covered 5 orders of magnitude from 2.21 for TX-100 (corresponding IC_{50} value: 162 μM) to 6.30 for acetone (corresponding IC_{50} value: 2.0 M). Extracellular IL-1α was very low in control tissue. Even if tissue was stimulated with LPS, which is an activator of the innate immune system, most of the induced IL-1α could only be detected intracellularly after 24 h. Nevertheless, the results showed extracellular IL-1α to be increased significantly after exposure to Ac2O, AHCP, aniline, ethanol, FA, GA, paracetamol, paraquat, SDS, TMA, and TX-100 in at least one laboratory. No increase or decrease could be observed after exposure with lactose, MMA, HCF, and TDI. Other chemicals induced a non-significant increase of extracellular IL-1α in at least two laboratories (Fig. 3). The increases were observed in the herein reported three independent experiments.

3.4. Intra- and inter-laboratory variability

Experiments were performed under as near identical conditions as possible. Beside all steps that were performed to minimize variability and are already described in the Materials and methods (e.g. material, equipment, SOPs), the staff was intensively trained — on site if necessary. We performed no staff swapping and blind evaluations but all data were submitted to and evaluated by an independent partner.

Intra-laboratory variability of log IC_{50} was evaluated for each of the biological replicates (N = 3) for each substance and laboratory. Results of the three laboratories were used to determine inter-laboratory variances. Mean values, standard deviation (Stdev), and coefficient of variation (CV) of log IC_{50} were calculated for each substance within each laboratory (Supplement Tables 3 to 6) and across the laboratories (Table 3).

Log IC_{50} showed best inter-laboratory consistency and good agreement for both the WST-1 and BCA assay (Table 3, Supplement Tables 3 and 5). For 52 combinations (out of 60 for three laboratories and 20 substances) log IC_{50} values for WST-1 could be estimated numerically, enabling calculation of intra- and inter-laboratory CV. About two thirds of the inter-laboratory CV values were lower than 5%; the maximum value was 16%. For 7 substances an intra-laboratory CV greater than 10% was determined. All inter-laboratory CV values were lower than 10%. For all experiments conducted with lactose and MMA log IC_{50} values were estimated greater than the maximum applied concentration (except one replication of laboratory A). Similar results have been observed for the BCA assay where the intra- and inter-laboratory CV values were only in few cases between 10% and 20%. Thus, these results were (at least) qualitatively reproducible.

Analysis of inter-laboratory correlation of log IC_{50} values reached satisfactory levels of agreement (Spearman r > 0.9, Fig. 4). BCA assay

proved to be as robust as WST-1 assay. Consistency between laboratories was also good (CV < 20%) for LDH assay. Intra- and inter-laboratory variability was high for intracellular IL-1α ELISA with intra- and inter-laboratory CV values reaching up to 50% and 25%, respectively. In general, agreement between the readings of laboratory A versus B was higher (Spearman) for all endpoints than between laboratories A versus C and B versus C.

3.5. Correlation of in-vitro (ex-vivo) IC_{50} with in-vivo inhalation LC_{50} and oral LD_{50} values

The substances used in this project were selected to represent different chemical classes and modes of action for which acute inhalation and oral lethality data are available (i.e. Sigma, RTECS, ChemIDplus Lite, TOXNET, ECHA, NIOSH Pocket Guide to Chemical Hazards, and unpublished BASF studies, Supplement Table 1). For some chemicals we found only inhalation LC_{50} values from studies with exposure durations different from 4 h. In these cases, the LC_{50} values were converted to 4-hour LC_{50} values (see also material and method section). The in-vivo data were then converted to molar concentrations, ranging from 0.01 μM for paraquat to 1.2 mM for acetone. Published oral rat LD_{50} values were also converted to molar-based doses which ranged from 222 μmol/kg for paraquat to 153 mmol/kg for ethanol and (Supplement Table 1). Thus, the most toxic substance via inhalation and oral uptake was paraquat. The least toxic substances were acetone for inhalation and ethanol for oral administration.

The results of the WST-1 assay were used for development two prediction models (prediction of LC_{50} by IC_{50} and two-group classification model). The WST-1 assay was selected because (i) it provides toxicologically relevant mechanistic evidence for cell injury, and ii) delivered a more complete data set as the LDH assay. Secondly we wanted to establish a model which has highest prospects for further validation. All endpoints were quite similar (high correlation) and the combination of two assays provided no additional information.

For correlation analysis, IC_{50} values i.e. the medians of the laboratory means obtained by WST-1 assay in rPCLS and/values of acute inhalation (LC_{50}) and oral lethality (LD_{50}) in rats were used for linear regression. No in-vivo inhalation toxicity data were available for paracetamol (inhalation), lactose (inhalation), and HCF (oral) (Supplement Table 1). Medians [μM] of IC_{50} values of rPCLS correlated highly significantly with rat LC_{50} values [μM] from inhalation studies (Pearson correlation coefficient $r = 0.87$, $p < 0.0001$) and with rat LD_{50} values [μmol/kg] from oral studies (Pearson correlation coefficient $r = 0.75$, $p = 0.0002$) (Fig. 5).

3.6. Two-group classification model

In the two-group classification model, substances were divided into two classes (toxic and non-toxic) based on their in-vivo inhalation toxicity. Substances with GHS categories 1, 2, and 3 were assigned to the toxic class. Substances with GHS categories 4 and 5 were classified as "non-toxic". Subsequently, the results of the WST-1 assay were used to determine thresholds. Each substance whose median was below this threshold was classified as "toxic", whereas substances with median values above the threshold were classified as "non-toxic. The goodness of prediction of the two-group classification model was assessed in terms of specific measures of the Cooper statistics, namely: sensitivity and specificity.

Median log IC_{50} values of each substance were used for classification. Since for lactose and MMA no effect could be determined up to a logarithmic dose of 5.7 on a μM scale, the median of log IC_{50} values of these two substances could not be expressed precisely. Therefore, we assumed for lactose and MMA the highest applied concentration as minimal possible IC_{50} value. To test robustness of the model even lower values were assumed for both substances which finally resulted only in small differences in classification. For the two-group classification model, IC_{50} values were transformed to mg/L units, considering the molecular weights of each substance. Thresholds of log IC_{50} for two-group classification were calculated to be 3.83 for laboratory A, 3.88 for laboratory B, and 3.91 for laboratory C. Classification of the substances using the two-group classification model is presented in Table 4 and Fig. 6. Sixteen out of 20 substances were classified correctly in all laboratories, thereof six substances of the non-toxic class and ten of the toxic class. Four substances, namely three non-toxic and one toxic

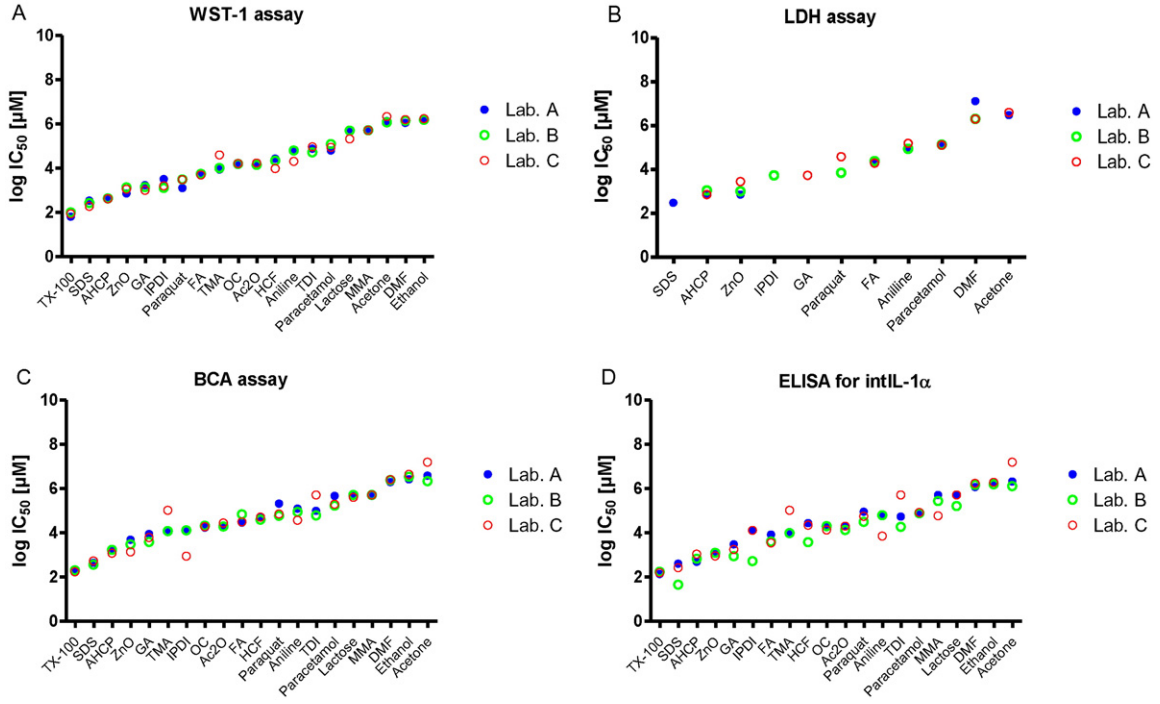


Fig. 2. Mean of log transformed IC_{50} [μM] values (log IC_{50}) for all chemicals ordered by magnitude is shown for (A) WST-1 assay, (B) LDH assay, (C) BCA assay, and (D) intracellular IL-1α (intIL-1α). Lab.: laboratory. For ">" values, minimum values have been entered.

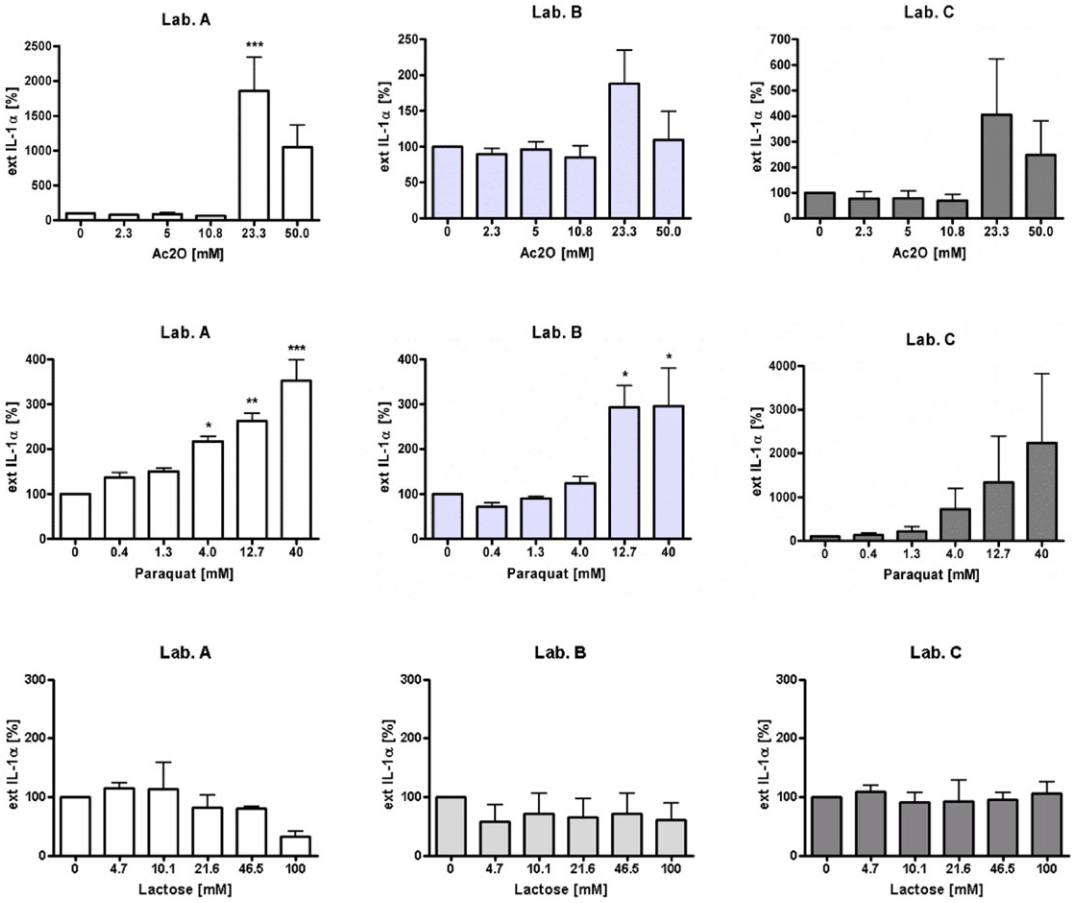


Fig. 3. Extracellular IL-1α (extIL-1α) of rPCLS after 1 h exposure to increasing concentrations of indicated substances. Cytokine levels were determined by ELISA. Data are presented as mean ± S.E.M. n = 3. *p < 0.05, **p < 0.01, and ***p < 0.001. Lab.: laboratory.

in-vivo, were classified falsely in all laboratories. TDI was falsely classified as non-toxic, whereas TMA, AHCP, and ZnO were classified falsely as positive in all laboratories (Table 4, Fig. 6). Thus, only one substance was false negative. Consequently, sensitivity was 91% and specificity was 66% in all participating laboratories. Assuming lower values for lactose and MMA changed the sensitivity of laboratory A to 82% and the specificity of laboratory C to 78%.

All results were obtained in leave-one-out cross-validation which is a model technique for evaluating how the results of a statistical analysis will generalize to an independent data set. This gives the possibility to estimate how accurately a predictive model will perform in practice (Geisser and Seymour, 1993; Kohavi, 1995). Thus, the toxicity of each substance per endpoint and laboratory combination was predicted by a model trained on subsets of data of all other substances for the same endpoint and laboratory combination without the subset data of the same substance.

3.7. Prediction of the LC_{50} by IC_{50}

The two-group classification model used in this study is a simplification of the actual toxicity. When using only toxic versus non-toxic class, the steepness of toxicity is not included, i.e. highly toxic (GHS category 1) and less toxic (GHS category 3) are all combined in the toxic class. Therefore, we have also viewed on how highly toxic substances (e.g. members of GHS category 1 such as TDI, IPDI, and paraquat) came out.

We used linear regression analysis to develop a prediction model which allows predicting in-vivo toxicity based on observed in-vitro cytotoxicity. In order to obtain a prediction model allowing estimation of LC_{50} by $IC_{50,}$ linear regression analysis was performed with values from every laboratory. The log IC_{50} values of the WST-1 assay were regarded here as predictors and the log-transformed LC_{50} values as response. The obtained formula is given in the material and methods section. The coefficients were estimated by the least-square method as $\beta_0 = -3.11$ and $\beta_1 = 0.9252$ for laboratory A, $\beta_0 = -3.12$ and $\beta_1 = 0.9287$ for laboratory B, and $\beta_0 = -2.53$ and $\beta_1 = 0.7993$ for laboratory C. Variances explained (i.e. the square of correlation coefficient) by this model were 63% for laboratory A, 63% for laboratory B, and 0.55% for laboratory C.

4. Discussion

4.1. Application of ACs to the different endpoint assays

At the beginning of this project, a standardized rPCLS protocol was established and variances in assay performance were determined. The basic protocol and technique for preparation of rat lung tissue sections closely followed those published in previous reports (Held et al., 1999; Ressmeyer et al., 2006). Initial sets of experiments using a first selection of six substances were focused on the definition of AC and subsequently applied to the remaining substances. When applying the ACs to each endpoint, WST-1, LDH, and BCA assays showed good validity, as the ACs were met by more than 95%. More difficulties in meeting the ACs were seen with the ELISA. The endpoint was highly sensitive to experimental conditions (e.g. differences in technical equipment, see also Materials and methods) beyond description in SOPs. Standardization was therefore not entirely possible. Experiments that did not meet all AC were not used for calculation of IC_{50} values but, instead, they were repeated and the data used for subsequent determinations of IC_{50} values.

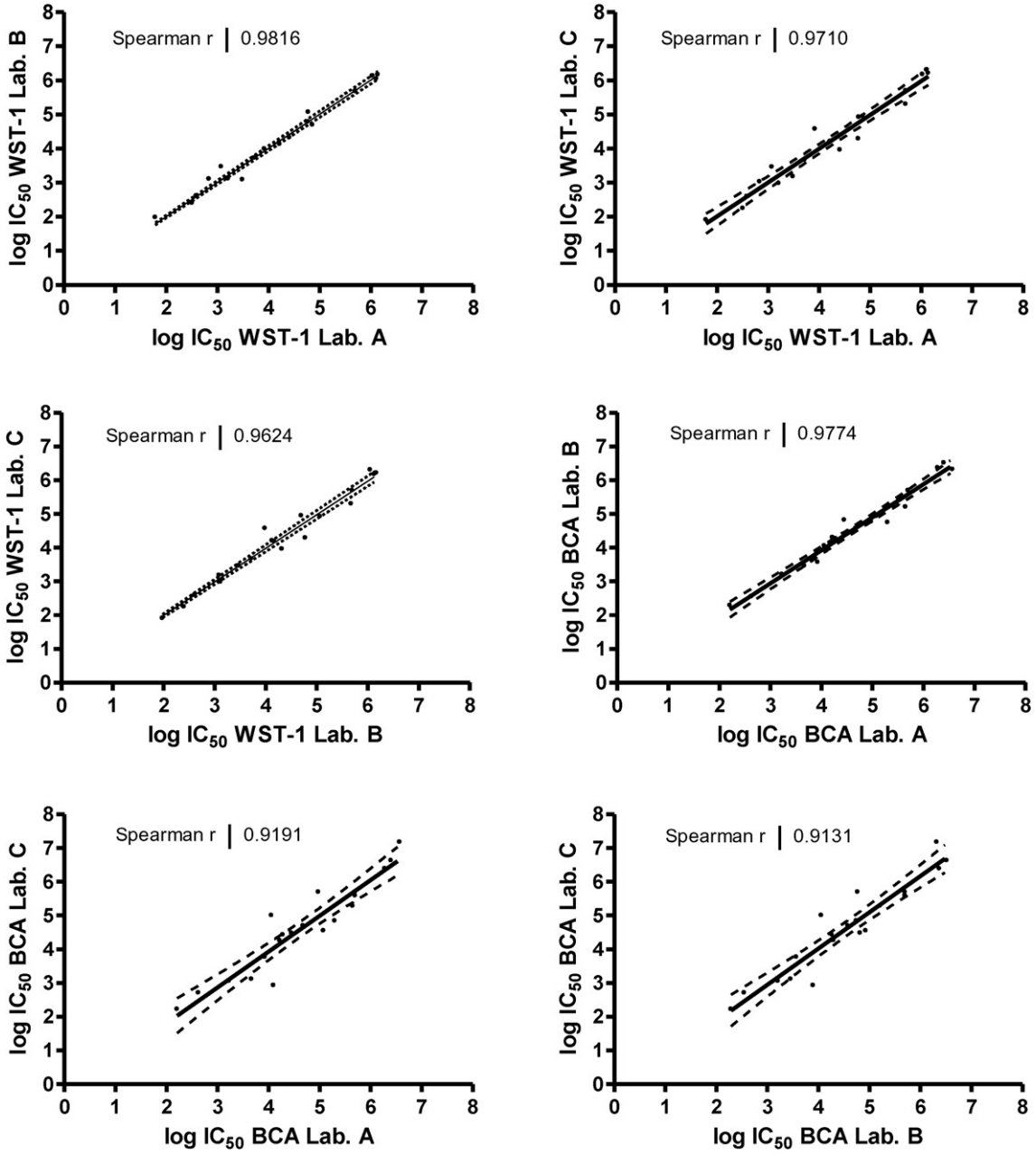


Fig. 4. Inter-laboratory correlation of mean log-transformed IC_{50} values [μM] obtained by WST-1 assay and BCA assay. The data of every laboratory was plotted against the data of every other laboratory. Comparison was performed by non-parametric Spearman correlation. For illustration only result of linear regression is shown as line and 95% confidence intervals are shown as dashed lines. Lab.: laboratory.

4.2. WST-1 assay revealed to be the best endpoint assay

Two endpoint assays were used for cytotoxicity: (i) determination of cell death as measured by leakage of the cytosolic enzyme LDH into culture medium (Legrand et al., 1992), and (ii) loss of metabolic enzyme activity assessed with the water-soluble dye WST-1 (Vistica et al., 1991).

Initially we thought that comparison of several viability endpoints could provide useful complementary data for (mechanistic) cytotoxicity, avoiding misinterpretation of chemicals interfering with the assay system. WST-1 assay provided the most complete and promising data set. LDH assay also proved to be responsive, ranking the chemicals in nearly the same order as the WST-1 assay but high substance concentrations induced wash-out effects. Thus, changing the medium after 1 h of incubation influences the outcome of the LDH assay. The wash-out effects were observed at high concentrations of AcO2, ethanol, HCF, OC, SDS, TDI, TMA, and TX-100. It was the main reason why these substances failed to be fitted. The use of later time points for LDH is often hampered by interferences of substances (e.g. AHCP, TMA) with the LDH activity (Lauenstein et al., 2014). Although the LDH assay was performed exactly the same as the WST-1 assay, it was actually not less predictive but less "practicable". Manual analysis and fitting of the big data set was not feasible. Altogether, the LDH assay was as responsive – also for 1 hour exposure – as the WST-1 assay but its further use was hindered by the limited efficiency during automatic data analysis.

It is worth mentioning here that BCA assay was standardized for determination of total protein content, but also reflected substance-induced cytotoxicity. The BCA assay was less responsive, as shown by the fact that WST-1 assay indicated cytotoxicity for some substances at lower concentrations already.

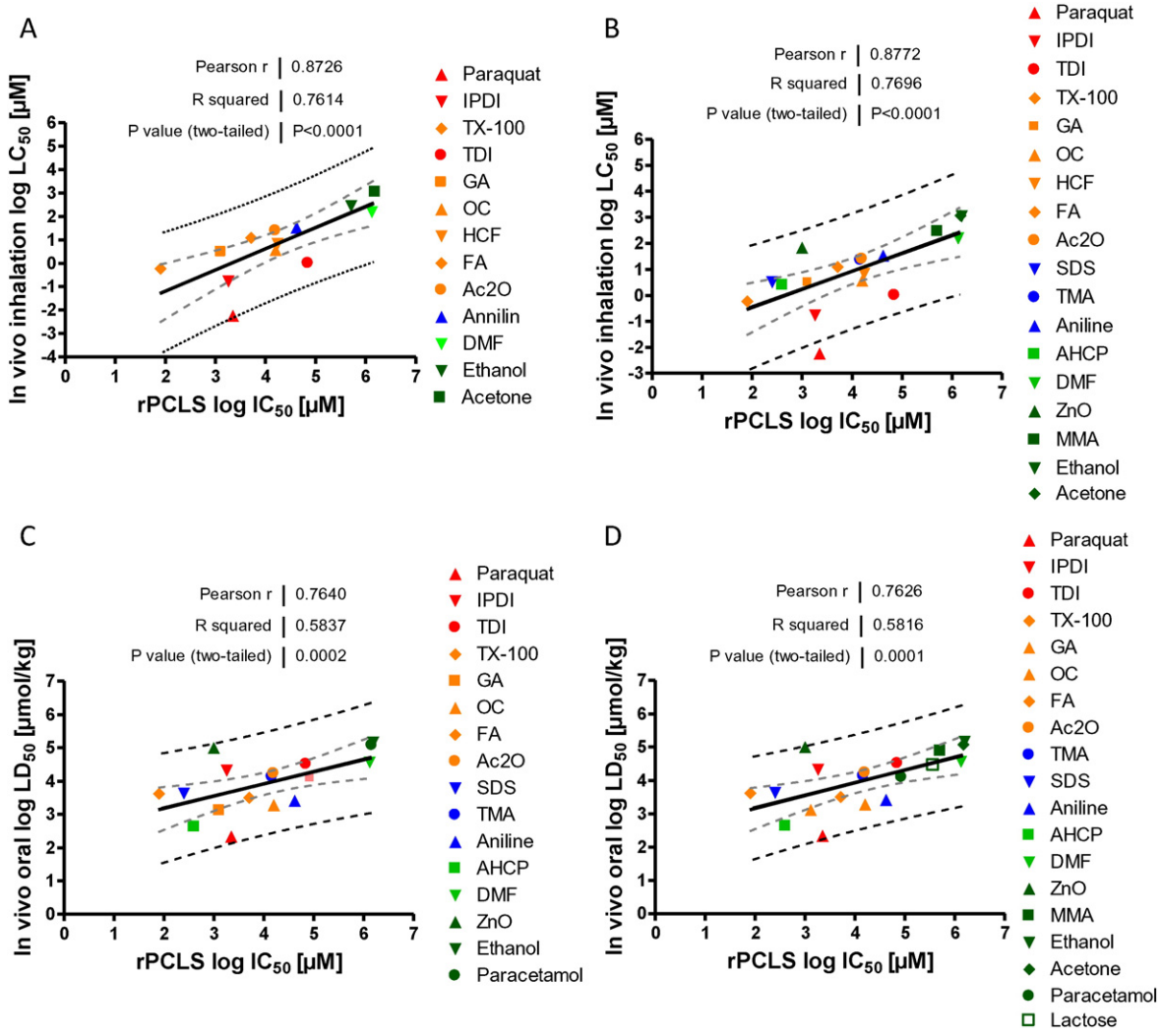


Fig. 5. Correlation of ex-vivo IC_{50} (WST-1 assay) with in-vivo inhalation LC_{50} (A, B) and oral LD_{50} values (C, D). In-vivo data were taken from online databases (Sigma, RTECS, ChemIDplus Lite, TOXNET, ECHA, and NIOSH Pocket Guide to Chemical Hazards) and converted to molar amounts (details are given in Supplement Table 1). IC_{50} values obtained by WST-1 assay in rPCLS and LC_{50}/LD_{50} values of acute inhalation and oral lethality in rats were used for linear regression analysis. A) Correlation shown for all substances except lactose, paracetamol (unknown LC_{50} values), MMA (no defined IC_{50} value), ZnO, TMA, AHCP, SDS (no defined LC_{50} values). B) Correlation shown for all substances except lactose, paracetamol (unknown LC_{50} values). For MMA, ZnO, TMA, AHCP, and SDS minimum values were used. C) Correlation shown for all substances except lactose, MMA (no defined IC_{50} values), and HCF (unknown LD_{50} value). D) Correlation shown for all substances except HCF (unknown LD_{50} value). For MMA and lactose minimum values were used. 95% confidence interval is shown as gray dashed line. 95% prediction interval is shown as black dashed line. Symbol colors refer to GHS classes as follows: 1 = red, 2 = orange, 3 = blue, 4 = light green, 5 = dark green.

Finally, we conclude that the combination of two different enzyme assays (LDH and WST-1) for cytotoxicity (with or without BCA assay) was not more sensitive and reliable than the WST-1 assay alone. The combination of the two assays did not provide additional complementary data. All substances that reduced the metabolic activity of the tissue also induced an increase in LDH activity in supernatant. Using the twenty described substances we found no case where it was contrary, e.g. decrease of tissue activity without detectable LDH leakage (e.g. due to substance interference).

4.3. Can IL-1α be used as biomarker for inflammation?

For inflammation we used only one endpoint. Substance-induced pro-inflammatory responses were assessed by quantification of IL-1α. This is a protein of the interleukin-1 family that has strong pro-inflammatory effects in the body, promoting the development of sepsis and fever (Dinarello, 2009; Frank et al., 2008; Sims and Smith, 2010). In lung tissue, IL-1α is produced mainly by monocytic cells such as activated macrophages, but also by epithelial and endothelial cells (Suwara et al., 2014). IL-1α has been reported to be an indicator of inflammation in lung tissue, e.g. due to exposure to LPS, adjuvants, or chemicals (Cao et al., 2011; Henjakovic et al., 2008; Xing et al., 1994). In our study, IL-1α responsiveness of rPCLS to LPS and substance treatments was quantitatively extremely variable within all laboratories. In rPCLS IL-1α peaks intracellularly at 24 h after permanent exposure to LPS (Henjakovic et al., 2008). Interestingly, some substances such as Ac2O and AHCP, known to induce respiratory inflammation accompanied by cellular infiltration and edema, significantly increased extracellular IL-1α levels in rPCLS (Fig. 3). On the other hand, substances such as lactose and MMA showed no effect on IL-1α levels. Although IL-1α could not be used to derive IC_{50}, these results suggest that IL-1α could be a marker for the prediction of substance induced inflammatory effects. Further efforts are needed to assess the usefulness of this information, and to establish it as valid and reliable endpoint as the inter-laboratory concordance was insufficient.

4.4. Intra- and inter-laboratory variabilities

The results from each laboratory were evaluated concerning intra- and inter-laboratory variability. Variability within and between

Table 4
Classification of substances using the binary prediction model. The number of each substance refers to known order in inhalation toxicity (see also Table 1). Lab.: laboratory; tox: toxic; non: non-toxic.

#	Substance	In vivo	Ex vivo		
			Lab. A	Lab. B	Lab. C
1	Paraquat	Tox	Tox	Tox	Tox
2	IPDI	Tox	Tox	Tox	Tox
3	TX-100	Tox	Tox	Tox	Tox
4	TDI	Tox	Non	Non	Non
5	AHCP	Non	Tox	Tox	Tox
6	GA	Tox	Tox	Tox	Tox
7	SDS	Tox	Tox	Tox	Tox
8	OC	Tox	Tox	Tox	Tox
9	HCF	Tox	Tox	Tox	Tox
10	FA	Tox	Tox	Tox	Tox
11	TMA	Non	Tox	Tox	Tox
12	Ac2O	Tox	Tox	Tox	Tox
13	Aniline	Tox	Tox	Tox	Tox
14	ZnO	Non	Tox	Tox	Tox
15	DMF	Non	Non	Non	Non
16	MMA	Non	Non	Non	Non
17	Ethanol	Non	Non	Non	Non
18	Acetone	Non	Non	Non	Non
19	Lactose	Non	Non	Non	Non
20	Paracetamol	Non	Non	Non	Non

laboratories were evaluated for each constellation. Best consistency, low variability, and strong agreement were found for WST-1 and BCA assays. Both assays were comparably robust. Results for LDH assay were found to be moderately matching, whereas intra- and inter-laboratory variability was high for IL-1α determination by ELISA. In summary, WST-1 assay seems to be best suited for analysis of cytotoxicity in rPCLS. The assay was transferable between the laboratories and provided relevant and reproducible results. The same was true for BCA assay used for determination of total protein content. The transferability of the ELISA based quantification of IL-1α was only limited, and results were less reproducible and very sensitive to laboratory specific conditions. In general, with a sample size of three biological replicates, the results have to be interpreted very carefully and need further validation.

Ranking of substances in the ex-vivo assay correlated well with underlying toxicological mechanisms. TX-100 is one of the most toxic compounds in-vivo and was ranked as most toxic in the ex-vivo model. Among the most toxic substances was also SDS, which is used as detergent, too. Both chemicals are known to disrupt cell membranes. SDS denatures proteins at higher concentrations (Chaturvedi and Kumar, 2011). Other chemicals such as isocyanates are known to attack proteins which lead to disturbance of cellular processes in e.g. metabolism, proteolysis, proliferation and protein synthesis. Substances such as Ac2O change their chemical structure. In aqueous media, Ac2O hydrolyses to acetic acid within few minutes. In our study the effects of Ac2O and acetic acid cannot be distinguished. In-vivo toxic threshold concentrations for Ac2O were found to be lower (about half) than for acetic acid. This is suggesting initial toxic effect by the anhydride before it hydrolyzed to acetic acid (OECD SIDS, 1997). Moreover, some chemicals were able to change the pH of the medium. They were not neutralized before they were applied to the test system in order to mimic real-life exposure. As test substances were incubated with rPCLS under submerged conditions, chemicals with limited water solubility such as TMA were initially dissolved in appropriate solvents. These solvents showed no effects at the final concentrations used.

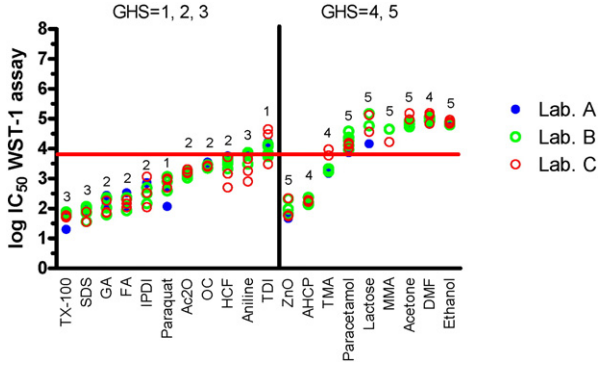


Fig. 6. Classification of substances using the binary prediction model. Means of log IC_{50} as obtained by WST-1 assay were converted to mg/L and ranked according to GHS category. GHS categories 1, 2, and 3 were sorted in toxic class. GHS categories 4 and 5 were sorted as non-toxic class. Red line indicates the ex-vivo threshold value. Small numbers above the symbols refer to the GHS class of each substance. Lab.: laboratory.

4.5. The two prediction models

For comparison with in-vivo data we developed two prediction models. The first prediction model was a two-group classification model where substances were divided into two classes based on their in-vivo acute inhalation toxicity. Substances with GHS categories 1, 2, and 3 were assigned to the toxic class and substances with GHS categories 4 and 5 to the non-toxic class. With this approach, 16 out of 20 test substances were classified correctly in all laboratories (Fig. 6). The reason why TDI and ZnO were falsely classified is most probably the same as that explaining the outliers in the second prediction model proposed in the next section below. A sensitivity of 91% was found for all laboratories, whereas the specificity was calculated at 66% for all laboratories.

The second prediction model was based on linear regression analysis of ex-vivo and in-vivo data. IC_{50} values obtained by WST-1 assay in rPCLS showed significant correlation with LC_{50} of in-vivo inhalation rats. The results can potentially be used to predict starting concentrations of acute inhalation toxicity studies. Significant correlation was also observed with LD_{50} values of oral studies in rats. Nevertheless, some outliers were obvious: (Akaike, 1998) Although paraquat is the substance with the lowest LC_{50} among the selected substances, it is ranked only on seventh position in the ex-vivo model (WST-1 assay) and was about 30 times less toxic than the most toxic substance ex-vivo TX-100. The lung is the primary target organ of paraquat. However, the primary underlying mechanism of paraquat is oxidative damage, which – and this is remarkable – leads to delayed toxic damage of lung tissue via pulmonary fibrosis. In the end, this condition is the usual cause of death after exposure to paraquat and which most commonly occurs 7–14 days after acute exposure (EPA Handbook, 2006). It shows that substances with long term effects such as paraquat can be tested only to a limited extent in acute models. (Archer et al., 1997) ZnO is less toxic in in-vivo studies (GHS category 5). In the rPCLS model, it ranked among the five most toxic substances. The difference is that ZnO was dissolved in acetic acid before application to rPCLS, whereas in-vivo it is inhaled in the undissolved form as dust (Klimisch et al., 1997). Toxicity of dissolved ZnO has been reported to be due to zinc ion concentrations leading to mitochondrial dysfunction, caspase activation, and apoptosis (Kao et al., 2012). Therefore, it can also not be compared to Sauer et al., 2014 where ZnO was applied as nanomaterial for 24 h. (Balls and Fentem, 1999) The diisocyanates IPDI and TDI are highly toxic in-vivo (GHS category 1) and ranked among the five most potent substances used in this study, but showed only minor cytotoxic effects ex-vivo. TDI in particular reacts quickly with water and its half-life time in water is only a few minutes (Collins, 2002). IPDI is more stable and remains in aqueous solutions for some hours (OECD SIDS, 2006). This might be the reason why the ex-vivo IC_{50} value of IPDI is about 30 times lower than that of TDI, whereas in-vivo both compounds differ only by a factor of six. It shows that substances being unstable in solution such as

isocyanates have only limited applicability for testing at submerse exposure conditions.

Both prediction models were based on the use of the median IC_{50} values. In general, models using concentrations different from IC_{50} may be of interest in future. Benchmark Dose (BMD) for example can also rank potencies of substances based on in-vitro dose–response curves even when no IC_{50} can be calculated (Louisse et al., 2010). With this approach doses for low effect levels such as a 5–10% change from baseline can be estimated and subsequently converted into predicted in-vivo doses. Although it was not applied in this study it could be an interesting approach in future.

4.6. How does the results compare to other in-vitro models?

In the following discussion our results are compared with other published in-vitro models. A recently published study (Sauer et al., 2013) investigated responses of the two commercially available 3D models, MucilAir™, and EpiAirway™, and of two monolayer cell lines, human A549 epithelial cells and 3T3 mouse embryonic fibroblasts, to the same chemicals as assessed in the present study. IC_{50} values using the MTT and WST-1 assays have been reported. We used these published IC_{50} values for direct correlation to in-vivo lethality data by linear regression analysis. The IC_{50} values of each model correlated significantly with oral LD_{50} and inhalation LC_{50} values (Supplement Fig. 1). Remarkable is the dynamic range spreading the data, which for the same chemicals was less pronounced for single-cell cultures (4 orders of magnitude for A549and 3T3) than for in-vivo toxicity data (6 orders of magnitude). Among these models, rPCLS was the first model used for prevalidation. In comparison with the other models, rPCLS correlated best with in-vivo data (Pearson $r = 0.87$). But also A549, MucilAir™, and EpiAirway™ showed good correlation with in-vivo (Pearson r 0.85, 0.84 and 0.83, respectively) (Supplement Fig. 1). Only 3T3 cells did not correlate well (Pearson $r = 0.71$). Interestingly, all models showed similar outlier pattern, as discussed already above for the results of the present study.

The question needs to be discussed whether for endpoints such as cytotoxicity single cell lines can be taken. The results of the study show that among the tested cell lines only alveolar epithelial cells such as A549 correlated well with in-vivo whereas the fibroblast cell line 3T3 did not. Lauenstein et al. (2014) showed that cells in human lung tissue behave more like monocyte-derived cells regarding their resistance to chemicals than skin-derived epithelial cells. This supports the theory that different cell populations have different responsiveness to chemicals. In the lung more than 45 different cell types can be found. Epithelial cells, nerve fibers, monocytes/macrophages, dendritic cells are in the first line of defense. All of them may respond differently to the same substance, some cell types even possess receptors that provide a chemical sensation (Guilliams et al., 2013). In the context of likely exposed cells the advantage of an ex-vivo tissue culture model is its multicellular microanatomy which is very attractive for a first assessment of acute local respiratory toxicity.

Moreover, we want to develop a model in future that is not merely based on cell injury but also provide early biomarkers for respiratory inflammation and chronic diseases, such as sensitization, proliferation, and fibrosis. The use of cell cultures is limited as published by Lauenstein et al. (2014). Here it was shown that biomarkers developed in cell culture were not proven to be predictive and reliable in human lung tissue. Switalla et al. (2010) showed that biomarkers in human lung tissue ex-vivo correlated very well with in-vivo response of probands. Nevertheless, in face of the different adverse effects after inhalation of toxicants it seems to be unlikely to develop a single approach – be it cell lines or tissue models – predicting the entire diversity of biological responses of the respiratory tract in acute toxicity studies.

4.7. Translation from in-vitro to in-vivo

The comparison of animal toxicity data with in-vitro/ex-vivo data remains critical due to the following issues: i) Differences can occur, for example, in the site of action, which can be different from the site of exposure. ii) Only a small number of substances was used in our study. Retrospective analysis of more than 300 chemicals tested according to the guidance document for oral toxicity has shown that a larger number of test chemicals provides different outcomes than the initial studies using less substances (Schrage et al., 2011). iii) Biotransformation in rPCLS remains questionable. Although, the lungs have been reported to have considerable capacity for metabolism (De Kanter et al., 2004), the rPCLS model has to be characterized using substances requiring biotransformation. Among the so far twenty tested substances only paracetamol requires biotransformation. Its metabolite is known to cause liver toxicity. Paracetamol showed weak toxicity in rPCLS suggesting somewhat biotransformation in rPCLS. iv) Exposure was done under submersed conditions. Whether or not rPCLS can be used to test, for example, nanomaterials, pesticides, metal compounds, and other industrial substances could not be addressed in this study. In particular, the use of rPCLS for the testing of chemicals quickly reacting with aqueous solutions might be limited (Landsiedel et al., 2010; Monopoli et al., 2011; Sauer et al., 2014). v) The model was set up as an acute exposure model. Long-term effects leading to development of fibrosis cannot be studied (Bonfield et al., 1995; Gwyer Findlay and Hussell, 2012; Yankaskas et al., 2004). Nevertheless, within this pre-validation study the respiratory toxicity data obtained for the selected substances reflected the in-vivo situation very closely. Issues that need to be reflected and one has to be aware of have been discussed above.

4.8. Conclusions and perspective

In conclusion, rPCLS as a test system for cytotoxicity can be part of a test battery of different in-vitro or in-vivo models to reflect the various processes associated with inhalation toxicity (Sauer et al., 2013). Whether or not this alternative can eventually be used to reduce or even replace in-vivo studies can only be answered by analyzing a larger set of new reference chemicals different from the set used in our study. However, according to the internationally recognized principle of a modular approach to validation (Hartung et al., 2004; OECD GD 34, 2005), the current study may be accepted already as a sufficient proof of intra- and interlaboratory reproducibility, and reproducibility over time. Thus, a new, independent data set might not necessarily have to be tested again in three laboratories. The main goal of such a special study would be a robust verification of the prediction models developed here post hoc with the data obtained from twenty reference chemicals in three laboratories. Since results of statistical cross-validation applied here to the training data set of twenty chemicals suggest robustness of the prediction models also when applied to new data, the current study offers a particularly good basis for a special follow-up study focused only on verification of the prediction models as one important part of the validation module “relevance”. A pre-submission of this approach to EURL/ECVAM and a check by the independent Boards of EURL/ECVAM (PARERE, ESTAC and ESAC) could aid deciding about the design of follow-up Validation Study (module).

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.tiv.2016.01.006.

Transparency document

The Transparency document associated with this article can be found, in the online version.

Acknowledgments

This research was supported by a grant FKZ 0315720 funding research and development projects on methods to replace animal testing ("Alternative Methods to Animal Testing") from the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung BMBF).

References

Akaike, H., 1998. Information theory and an extension of the maximum likelihood principle. Springer, New York, pp. 199–213 (ISBN 978–1-4612-7248-9).
Archer, G., Balls, M., Bruner, L.H., Curren, R.D., Fentem, J.H., Holzhütter, H.G., Liebsch, M., Lovell, D.P., Southee, J.A., 1997. The validation of toxicological prediction models. ATLA-NOTTINGHAM 25, 505–516.
Balls, M., Fentem, J.H., 1999. The validation and acceptance of alternatives to animal testing. Toxicol. in Vitro 13, 837–846.
BeruBe, K., Aufderheide, M., Breheny, D., Clothier, R., Combes, R., Duffin, R., Forbes, B., Gaca, M., Gray, A., Hall, I., Kelly, M., Lethem, M., Liebsch, M., Merolla, L., Morin, J.P., Seagrave, J., Swartz, M.A., Tetley, T.D., Umachandran, M., 2009. In vitro models of inhalation toxicity and disease. The report of a FRAME workshop. Altern. Lab. Anim. 37, 89–141.
Bonfield, T.L., Panuska, J.R., Konstan, M.W., Hilliard, K.A., Hilliard, J.B., Ghnaim, H., Berger, M., 1995. Inflammatory cytokines in cystic fibrosis lungs. Am. J. Respir. Crit. Care Med. 152, 2111–2118.
Borchers, H.W., 2015. Practical Numerical Math Functions. http://cran.r-project.org/web/packages/pracma/pracma.pdf.
Calzetta, L., Spina, D., Cazzola, M., Page, C.P., Facciolo, F., Rendina, E.A., Matera, M.G., 2011. Pharmacological characterization of adenosine receptors on isolated human bronchi. Am. J. Respir. Cell Mol. Biol. 45, 1222–1231.
Cao, Q., Jing, C., Tang, X., Yin, Y., Han, X., Wu, W., 2011. Protective effect of resveratrol on acute lung injury induced by lipopolysaccharide in mice. Anat. Rec (Hoboken) 294, 527–532.
Chaturvedi, V., Kumar, A., 2011. Isolation of a strain of *Pseudomonas putida* capable of metabolizing anionic detergent sodium dodecyl sulfate (SDS). Iran. J. Microbiol. 3, 47–53.
Collins, M.A., 2002. Toxicology of toluene diisocyanate. Appl. Occup. Environ. Hyg. 17, 846–855.
Curren, R.D., Mun, G.C., Gibson, D.P., Aardema, M.J., 2006. Development of a method for assessing micronucleus induction in a 3D human skin model (EpiDerm). Mutat. Res. 607, 192–204.
De Kanter, R., Monshouwer, M., Draaisma, A.L., De Jager, M.H., de Graaf, I.A., Proost, J.H., Meijer, D.K., Groothuis, G.M., 2004. Prediction of whole-body metabolic clearance of drugs through the combined use of slices from rat liver, lung, kidney, small intestine and colon. Xenobiotica 34, 229–241.
Dinarello, C.A., 2009. Immunological and inflammatory functions of the interleukin-1 family. Annu. Rev. Immunol. 27, 519–550.
Escher, S.E., Tluczkiewicz, I., Batke, M., Bitsch, A., Melber, C., Kroese, E.D., Buist, H.E., Mangelsdorf, I., 2010. Evaluation of inhalation TTC values with the database RepDose. Regul. Toxicol. Pharmacol. 58, 259–274.
Frank, J.A., Pittet, J.F., Wray, C., Matthay, M.A., 2008. Protection from experimental ventilator-induced acute lung injury by IL-1 receptor blockade. Thorax 63, 147–153.
Geisser, Seymour, 1993. Predictive Inference. Chapman and Hall, New York, NY (ISBN 0–412-03471–9).
GHS Guidance, 2005. Technical guidance document on the GHS classification. http://www.safe.nite.go.jp/english/ghs/pdf/ghs_guidance_e.pdf.
Grubbs, F.E., 1950. Sample criteria for testing outlying observations. Ann. Math. Stat. 21, 27–58.
Guidance Document on Using In Vitro Data to Estimate In Vivo Starting Doses for Acute Toxicity, 2001. Based on Recommendations from an International Workshop Organised by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) National Institute of Environmental Health Sciences National Institutes of Health US Public Health Service Department of Health and Human Services. NIEHS, Triangle Park, NC, USA (Available at: http://www.epa.gov/hpv/pubs/general/nih, 102 pp.).
Guilliams, M., Lambrecht, B.N., Hammad, H., 2013. Division of labor between lung dendritic cells and macrophages in the defense against pulmonary infections. Mucosal Immunol. 6, 464–473.
Gwyer Findlay, E., Hussell, T., 2012. Macrophage-mediated inflammation and disease: a focus on the lung. Mediat. Inflamm. 2012, 140937.
EPA Handbook, 2006. Paraquat and Diquat. http://www.epa.gov/opp00001/safety/healthcare/handbook/Chap12.pdf.
Hansen, M.B., Skov, L., Menne, T., Olsen, J., 2005. Gene transcripts as potential diagnostic markers for allergic contact dermatitis. Contact Dermatitis 53, 100–106.
Hartung, T., Bremer, S., Casati, S., Coecke, S., Corvi, R., Fortaner, S., Gribaldo, L., Halder, M., Hoffmann, S., Janusch Roi, A., Prieto, P., Sabbioni, E., Scott, L., Worth, A., Zuang, V., 2004. A modular approach to the ECVAM principles on test validity. ATLA 32, 467–472.
Hayes, A., Bakand, S., 2010. Inhalation toxicology. EXS 100, 461–488.
Held, H.D., Martin, C., Uhlig, S., 1999. Characterization of airway and vascular responses in murine lungs. Br. J. Pharmacol. 126, 1191–1199.
Henjakovic, M., Sewald, K., Switalla, S., Kaiser, D., Muller, M., Veres, T.Z., Martin, C., Uhlig, S., Krug, N., Braun, A., 2008. Ex vivo testing of immune responses in precision-cut lung slices. Toxicol. Appl. Pharmacol. 231, 68–76.
Hirakata, Y., Yano, H., Arai, K., Endo, S., Kanamori, H., Aoyagi, T., Hirotani, A., Kitagawa, M., Hatta, M., Yamamoto, N., Kunishima, H., Kawakami, K., Kaku, M., 2010. Monolayer culture systems with respiratory epithelial cells for evaluation of bacterial invasiveness. Tohoku J. Exp. Med. 220, 15–19.
Hope, J.H., Hope, B.E., 2012. A review of the diagnosis and treatment of Ochratoxin A inhalational exposure associated with human illness and kidney disease including focal segmental glomerulosclerosis. J. Environ. Public Health 2012, 835059.
Huang, S., Wiszniewski, L., Derouette, J.P., Constant, S., 2009. In vitro organ culture models of asthma. Drug Discov. Today 6, 137–144.
Huang, S., Wiszniewski, L., Constant, S., 2011. The Use of In Vitro 3D Cell Models in Drug Development for Respiratory Diseases. http://cdn.intechweb.org/pdfs/25227.pdf.
Huang, S., Wiszniewski, L., Constant, S., Roggen, E., 2013. Potential of in vitro reconstituted 3D human airway epithelia (MucilAir) to assess respiratory sensitizers. Toxicol. in Vitro 27, 1151–1156.
Hulette, B.A., Ryan, C.A., Gerberick, G.F., 2002. Elucidating changes in surface marker expression of dendritic cells following chemical allergen treatment. Toxicol. Appl. Pharmacol. 182, 226–233.
Kao, Y.Y., Chen, Y.C., Cheng, T.J., Chiung, Y.M., Liu, P.S., 2012. Zinc oxide nanoparticles interfere with zinc ion homeostasis to cause cytotoxicity. Toxicol. Sci. 125, 462–472.
Kennedy, G.L., 2012. Toxicology of dimethyl and monomethyl derivatives of acetamide and formamide: a second update. Crit. Rev. Toxicol. 42, 793–826.
Klimisch, H.J., Andreae, M., Tillmann, U., 1997. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regul. Toxicol. Pharmacol. 25, 1–5.
Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. Morgan Kaufmann Publishers Inc., Montreal, Quebec, Canada, pp. 1137–1143 (ISBN 1-55860-363-8).
Kroigaard, C., Dalsgaard, T., Nielsen, G., Laursen, B.E., Pilegaard, H., Kohler, R., Simonsen, U., 2012. Activation of endothelial and epithelial K(Ca) 2.3 calcium-activated potassium channels by NS309 relaxes human small pulmonary arteries and bronchioles. Br. J. Pharmacol. 167, 37–47.
Lalko, J.F., Kimber, I., Dearman, R.J., Api, A.M., Gerberick, G.F., 2012. The selective peptide reactivity of chemical respiratory allergens under competitive and non-competitive conditions. J. Immunotoxicol. 10, 292–301.
Landsiedel, R., Ma-Hock, L., Kroll, A., Hahn, D., Schnekenburger, J., Wiench, K., Wohlleben, W., 2010. Testing metal-oxide nanomaterials for human safety. Adv. Mater. 22, 2601–2627.
Larsson, K., Lindstedt, M., Lundberg, K., Dexlin, L., Wingren, C., Korsgren, M., Greiff, L., Borrebaeck, C.A., 2009. CD4+ T cells have a key instructive role in educating dendritic cells in allergy. Int. Arch. Allergy Immunol. 149, 1–15.
Lauenstein, L., Switalla, S., Prenzler, F., Seehase, S., Pfennig, O., Förster, C., Fieguth, F., Braun, A., Sewald, K., 2014. Assessment of immunotoxicity induced by chemicals in human precision-cut lung slices (PCLS). Toxicol. in Vitro 28, 588–599.
Legrand, C., Bour, J.M., Jacob, C., Capiaumont, J., Martial, A., Marc, A., Wudtke, M., Kretzmer, G., Demangel, C., Duval, D., 1992. Lactate dehydrogenase (LDH) activity of the cultured eukaryotic cells as marker of the number of dead cells in the medium [corrected]. J. Biotechnol. 25, 231–243.
Louisse, J., de Jong, E., van de Sandt, J.J.M., Blaauboer, B.J., Woutersen, R.W., Piersma, A.H., Rietjens, I.M., Verwei, M., 2010. The use of in vitro toxicity data and physiologically based kinetic modeling to predict dose–response curves for in vivo developmental toxicity of glycol ethers in rat and man. Toxicol. Sci. 118, 470–484.
Megherbi, R., Kiorpelidou, E., Foster, B., Rowe, C., Naisbitt, D.J., Goldring, C.E., Park, B.K., 2012. Role of protein haptenation in triggering maturation events in the dendritic cell surrogate cell line THP-1. Br. J. Pharmacol. 167, 37–47.
Mitjans, M., Viviani, B., Lucchi, L., Galli, C.L., Marinovich, M., Corsini, E., 2008. Role of p38 MAPK in the selective release of IL-8 induced by chemical allergen in naive THp-1 cells. Toxicol. in Vitro 22, 386–395.
Monopoli, M.P., Walczyk, D., Campbell, A., Elia, G., Lynch, I., Bombelli, F.B., Dawson, K.A., 2011. Physical–chemical aspects of protein corona: relevance to in vitro and in vivo biological impacts of nanoparticles. J. Am. Chem. Soc. 133, 2525–2534.
Nelissen, I., Selderslaghs, I., Heuvel, R.V., Witters, H., Verheyen, G.R., Schoeters, G., 2009. MUTZ-3-derived dendritic cells as an in vitro alternative model to CD34+ progenitor-derived dendritic cells for testing of chemical sensitizers. Toxicol. in Vitro 23, 1477–1481.
Nottingham, Q.J., Birch, J.B., 2000. A semiparametric approach to analysing dose–response data. Stat. Med. 19, 389–404.
OECD, 2010. Guidance document on using cytotoxicity tests to estimate starting doses for acute oral systematic toxicity tests. OECD Series on Testing and Assessment No. 34, OECD, Paris (Available at: http://www.oecd.org/env/ehs/testing/seriesontestingandassessmentpublicationsbynumber.htm).
OECD GD34, 2005. Retrospective Performance Assessment Of The Draft Test Guideline 426 On Developmental Neurotoxicity. http://www.oecd.org/chemicalsafety/testing/37622258.pdf.
OECD SIDS, 1997. SIDS Initial Assessment Report for 6th SIAM (Acetic acid). http://www.chem.unep.ch/irptc/sids/OECDSIDS/79118.pdf.
OECD SIDS, 2006. SIDS Initial Assessment Report 23th SIAM (IPDI). http://www.alipa.org/uploads/Modules/Publications/OECD%20SIDS_IPDI.pdf.
OECD TG423, 2001. Acute Oral Toxicity for testing of chemicals. http://www.oecd.org/chemicalsafety/risk-assessment/1948370.pdf.
OECD TG425, 2002. Acute Oral Toxicity: Up-and-Down Procedure. http://www.oecd-ilibrary.org/environment/test-no-425-acute-oral-toxicity-up-and-down-procedure_9789264071049-en.

OECD TG433, 2004. Acute Inhalation Toxicity — Fixed Concentration Procedure. http://www.oecd.org/chemicalsafety/testing/32035886.pdf.

OECD TG436, 2009. Acute Inhalation Toxicity — Acute Toxic Class Method. http://www.oecd-ilibrary.org/docserver/download/9743601e.pdf?expires=1371137021&id=id&accname=guest&checksum=90D4541A82AC3C17D2D5A6E8B16833E6.

Patlewicz, G., Dimitrov, S.D., Low, L.K., Kern, P.S., Dimitrova, G.D., Comber, M.I., Aptula, A.O., Phillips, R.D., Niemela, J., Madsen, C., Wedebye, E.B., Roberts, D.W., Bailey, P.T., Mekenyan, O.G., 2007. TIMES-SS—a promising tool for the assessment of skin sensitization hazard. A characterization with respect to the OECD validation principles for (Q)SARs and an external evaluation for predictivity. Regul. Toxicol. Pharmacol. 48, 225–239.

Python, F., Goebel, C.F., Aeby, P., 2009. Comparative DNA microarray analysis of human monocyte derived dendritic cells and MUTZ-3 cells exposed to the moderate skin sensitizer cinnamaldehyde. Toxicol. Appl. Pharmacol. 239, 273–283.

R Development Core Team, 2010. R: A Language and Environment for Statistical Computing. http://www.lsw.uni-heidelberg.de/users/christlieb/teaching/UKStaSS10/R-refman.pdf.

Ren, D., Daines, D.A., 2011. Use of the EpiAirway model for characterizing long-term host-pathogen interactions. J. Vis. Exp., e3261

Ressmeyer, A.R., Larsson, A.K., Vollmer, E., Dahlen, S.E., Uhlig, S., Martin, C., 2006. Characterisation of guinea pig precision-cut lung slices: comparison with human tissues. Eur. Respir. J. 28, 603–611.

Reus, A.A., Maas, W.J., Jansen, H.T., Constant, S., Staal, Y.C., van Triel, J.J., Kuper, F., 2013. Feasibility of a 3D human airway epithelial model to study respiratory absorption. Toxicol. in Vitro 28, 258–264.

Ritz, C., 2005. Bioassay Analysis Using R. http://www.jstatsoft.org/v12/i05/paper.

Sauer, U.G., Vogel, S., Hess, A., Kolle, S.N., Ma-Hock, L., van, R.B., Landsiedel, R., 2013. In vivo-in vitro comparison of acute respiratory tract toxicity using human 3D airway epithelial models and human A549 and murine 3T3 monolayer cell systems. Toxicol. in Vitro 27, 174–190.

Sauer, U.G., Vogel, S., Aumann, A., Hess, A., Kolle, S.N., Ma-Hock, L., Wohlleben, W., Dammann, M., Strauss, V., Treumann, S., Groters, S., Wiench, K., van, R.B., Landsiedel, R., 2014. Applicability of rat precision-cut lung slices in evaluating nanomaterial cytotoxicity, apoptosis, oxidative stress, and inflammation. Toxicol. Appl. Pharmacol. 276, 1–20.

Schrage, A., Hempel, K., Schulz, M., Kolle, S.N., van, R.B., Landsiedel, R., 2011. Refinement and reduction of acute oral toxicity testing: a critical review of the use of cytotoxicity data. Altern. Lab. Anim. 39, 273–295.

Sims, J.E., Smith, D.E., 2010. The IL-1 family: regulators of immunity. Nat. Rev. Immunol. 10, 89–102.

Sullivan, K., Cochrane, S., Enoch, S., Kimber, I., Patlewicz, G., Roggen, E., Sewald, K., Ezendam, J., 2014. Promises and challenges in constructing an adverse outcome pathway for chemical sensitization of the respiratory tract. http://publica.fraunhofer.de/documents/N-315128.html.

Suwara, M.I., Green, N.J., Borthwick, L.A., Mann, J., Mayer-Barber, K.D., Barron, L., Corris, P.A., Farrow, S.N., Wynn, T.A., Fisher, A.J., Mann, D.A., 2014. IL-1alpha released from damaged epithelial cells is sufficient and essential to trigger inflammatory responses in human lung fibroblasts. Mucosal Immunol. 7, 684–693.

Switalla, S., Lauenstein, L., Prenzler, F., Knothe, S., Förster, C., Fieguth, H.G., Pfennig, O., Schaummann, F., Martin, C., Guzman, C.A., Ebensen, T., Müller, M., Hohlfeld, J.M., Krug, N., Braun, A., Sewald, K., 2010. Natural innate cytokine response to immunomodulators and adjuvants in human precision-cut lung slices. Toxicol. Appl. Pharmacol. 246, 107–115.

Trifilieff, A., Baur, F., Fozard, J.R., 2009. Role of sphingosine-1-phosphate (S1P) and the S1P(2) receptor in allergen-induced, mast cell-dependent contraction of rat lung parenchymal strips. Naunyn Schmiedeberg's Arch. Pharmacol. 380, 303–309.

Vandebriel, R., De Jong, W., 2012. A review of mammalian toxicity of ZnO nanoparticles. Nanotechnol. Sci. Appl. 5, 61–71.

Vistica, D.T., Skehan, P., Scudiero, D., Monks, A., Pittman, A., Boyd, M.R., 1991. Tetrazolium-based assays for cellular viability: a critical examination of selected parameters affecting formazan production. Cancer Res. 51, 2515–2520.

Wanner, R., Sonnenburg, A., Quatchadze, M., Schreiner, M., Peiser, M., Zuberbier, T., Stahlmann, R., 2010. Classification of sensitizing and irritative potential in a combined in-vitro assay. Toxicol. Appl. Pharmacol. 245, 211–218.

Worth, A.P., Balls, M., 2001. The Importance of the Prediction Model in the Validation of Alternative Tests. http://www.frame.org.uk/dynamic_files/worth_vol_29_2_.pdf.

Xing, Z., Jordana, M., Kirpalani, H., Driscoll, K.E., Schall, T.J., Gauldie, J., 1994. Cytokine expression by neutrophils and macrophages in vivo: endotoxin induces tumor necrosis factor-alpha, macrophage inflammatory protein-2, interleukin-1 beta, and interleukin-6 but not RANTES or transforming growth factor-beta 1 mRNA expression in acute lung inflammation. Am. J. Respir. Cell Mol. Biol. 10, 148–153.

Yankaskas, J.R., Marshall, B.C., Sufian, B., Simon, R.H., Rodman, D., 2004. Cystic fibrosis adult care: consensus conference report. Chest 125, 1S–39S.