Mutation Research - Genetic Toxicology and Environmental Mutagenesis 879-880 (2022) 503503

Contents lists available at ScienceDirect

Mutation Research - Genetic Toxicology and Environmental Mutagenesis

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

A comparison of the lowest effective concentration in culture media for detection of chromosomal damage *in vitro* and in blood or plasma for detection of micronuclei *in vivo*

David Kirkland [a,*], James Whitwell [b], Robert Smith [b], Kiyohiro Hashimoto [c], Zhiying Ji [d], Julia Kenny [e], Naoki Koyama [f], David P. Lovell [g], Hans-Jörg Martus [h], Krista Meurer [i], Daniel Roberts [j], Akira Takeiri [k], Yoshifumi Uno [l], Bas-Jan van der Leede [m], Paul White [n], Andreas Zeller [o]

[a] *Kirkland Consulting, PO Box 79, Tadcaster LS24 0AS, UK*
[b] *Labcorp Drug Development, Otley Road, Harrogate HG3 1PY, UK*
[c] *Takeda Pharmaceutical Company Limited, Research, Drug Safety Research and Evaluation, Fujisawa, Japan*
[d] *Bristol Myers Squibb Company, Nonclinical Safety, 1 Squibb Drive, New Brunswick, NJ 08903, USA*
[e] *GlaxoSmithKline, Park Road, Ware SG12 0DP, UK*
[f] *Eisai Co., Ltd, Global Drug Safety, Tsukuba-shi, Ibaraki, Japan*
[g] *Public Health Research Institute (PHRI), St George's Hospital Medical School, University of London, Cranmer Terrace, London SW17 0RE, UK*
[h] *Novartis Institutes for BioMedical Research, Preclinical Safety, 4002 Basel, Switzerland*
[i] *BASF SE, Li444, Speyererstrasse 2, 67117 Limburgerhof, Germany*
[j] *Charles River Laboratories, Skokie, IL, USA*
[k] *Chugai Pharmaceutical Co., Ltd., Translational Research Div., 1-135 Komakado, Gotemba, Shizuoka, Japan*
[l] *LSI Medience Co., 1-13-4, Tokyo 101-8517, Japan*
[m] *Non-Clinical Safety, Janssen Research & Development, a Division of Janssen Pharmaceutica N.V., Turnhoutseweg 30, 2340 Beerse, Belgium*
[n] *Environmental Health Science and Research Bureau, Health Canada, Ottawa, Ontario, Canada*
[o] *F. Hoffmann-La Roche Ltd., Pharmaceutical Sciences, pRED Innovation Center Basel, 4070 Basel, Switzerland*

ARTICLE INFO

Keywords:
Chromosomal damage
In vitro concentration
Plasma concentrations *in vivo*
Clastogens
Aneugens

ABSTRACT

It is often assumed that genotoxic substances will be detected more easily by using *in vitro* rather than *in vivo* genotoxicity tests since higher concentrations, more cytotoxicity and static exposures can be achieved. However, there is a paucity of data demonstrating whether genotoxic substances are detected at lower concentrations in cell culture *in vitro* than can be reached in the blood of animals treated *in vivo*. To investigate this issue, we compared the lowest concentration required for induction of chromosomal damage *in vitro* (lowest observed effective concentration, or LOEC) with the concentration of the test substance in blood at the lowest dose required for biologically relevant induction of micronuclei *in vivo* (lowest observed effective dose, or LOED). In total, 83 substances were found for which the LOED could be identified or estimated, where concentrations in blood and micronucleus data were available *via* the same route of administration in the same species, and *in vitro* chromosomal damage data were available. 39.8 % of substances were positive *in vivo* at blood concentrations that were lower than the LOEC *in vitro*, 22.9 % were positive at similar concentrations, and 37.3 % of substances were positive *in vivo* at higher concentrations. Distribution analysis showed a very wide scatter of > 6 orders of magnitude across these 3 categories. When mode of action was evaluated, the distribution of clastogens and aneugens across the 3 categories was very similar. Thus, the ability to detect induction of micronuclei in bone

Abbreviations: ADME, absorption, distribution, metabolism and excretion; AED, administered equivalent dose; CA, chromosomal aberrations; CBMN, cytokinesis block micronucleus assay; C_{max}, peak plasma concentration; GLP, Good Laboratory Practice; ICH, International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use;; IVIVE, *in vitro* to *in vivo* extrapolation; LOEC, lowest observed effective concentration *in vitro*; LOED, lowest observed effective dose *in vivo*; MLA, mouse lymphoma assay; MN, micronuclei; MoA, mode of action; MTD, maximum tolerated dose; NCE, normochromatic erythrocytes; NTP, National Toxicology Programme; PCE, polychromatic erythrocytes; PoD, point of departure; retics, reticulocytes; SD, standard deviation.

* Corresponding author.
E-mail address: dkirkland@genetoxconsulting.co.uk (D. Kirkland).

https://doi.org/10.1016/j.mrgentox.2022.503503
Received 20 March 2022; Received in revised form 13 May 2022; Accepted 25 May 2022
Available online 28 May 2022

marrow *in vivo* regardless of the mechanism for micronucleus induction, is clearly not solely determined by the concentration of test substance which induced chromosomal damage *in vitro*.

1. Introduction

It is often assumed that genotoxic substances will be detected more easily by using *in vitro* rather than *in vivo* genotoxicity tests because testing *in vitro* is usually conducted at higher concentrations, and the target cells are usually subjected to greater cytotoxicity than is likely to occur in an organism such as a rat or a mouse. Moreover, exposure of target cells *in vitro* is typically static in a closed vessel, test chemical concentrations are maintained over the entire exposure period in the absence of metabolic activation, or metabolites are produced in direct contact with the target cells. On the other hand, plasma concentrations leading to exposure of target cells *in vivo*, after a usually quick absorption phase, generally follows first-order elimination kinetics and concentrations rapidly decline. Therefore, *in vivo* systemic exposure of target cells to xenobiotic agents is generally expected to be lower when compared to *in vitro* exposures. In terms of detecting the genotoxic potential of a test substance, it is therefore reasonable to assume that genotoxic substances will be detected more easily by using *in vitro* rather than *in vivo* testing, which supports the hierarchical approach in use for genotoxicity testing adopted by most regulatory agencies. However, is this difference in ability to detect which substances are genotoxic driven by similar differences in the concentrations at which they are first detected as genotoxic? In other words, for a substance to be detected as genotoxic *in vivo*, is it necessary to achieve similar or higher concentrations in blood than are required to produce genotoxic effects in culture *in vitro*? Indeed, if this is the case it could lead to the rejection of negative *in vivo* studies for substances that are genotoxic *in vitro* when the blood plasma concentration does not reach or exceed the concentrations that are genotoxic *in vitro*. Such questions are also relevant for the developing IVIVE (*in vitro*-to-*in vivo* extrapolation) models, which use *in vitro* concentration-response data for hazard evaluation and employ toxicokinetic data to determine the oral dose required to elicit a steady state blood concentration equivalent to the *in vitro* point of departure (PoD) value (*e.g.*, the concentration at 50 % of maximum activity, or AC_{50}; [1,2]). Since it is now accepted that "misleading positive" results can occur *in vitro* as a result of physiological disturbances or damage to non-DNA targets, with substances that are not genotoxic or carcinogenic *in vivo* [3], and that choice of cell type and cytotoxicity measures are relevant in avoiding misleading positive results [4,5], it is important to have confidence that a negative *in vivo* result with a substance that was genotoxic *in vitro* is robust. It is therefore essential to investigate if there is a correspondence between culture medium concentrations required to elicit a response *in vitro* and blood or plasma concentrations *in vivo* required to elicit a tissue-specific genotoxic response.

In order to investigate this issue, we compared the lowest concentration at which there was biologically relevant induction of chromosomal damage *in vitro* (lowest observed effective concentration, or LOEC) with the blood or plasma concentration at the lowest dose at which there was biologically relevant induction of micronuclei (MN) *in vivo* (lowest observed effective dose, or LOED). The paper describes how relevant data were collected, and whether the blood plasma concentration at the LOED was lower than, higher than or similar to the LOEC.

2. Methods

2.1. Collection of relevant data

Initially, companies and contract laboratories that participate in the Genetic Toxicology Technical Committee (GTTC) of the Health and Environmental Sciences Institute (HESI) were asked if they had in-house data on substances that were positive for induction of MN in bone marrow *in vivo*, for which there were relevant (*i.e.*, by route of administration and species) concentration data in blood or plasma, and for which there were *in vitro* results for chromosomal damage, namely induction of MN, chromosomal aberrations (CA) or small colony mouse lymphoma (MLA) mutants. If a substance was positive *in vivo* but negative *in vitro* in a robust test that reached an acceptable maximum concentration (*i.e.*, in terms of cytotoxicity, solubility limit or the maximum required concentration for a freely soluble non-toxic substance), it also was included. This call for data resulted in an initial database of 36 compounds with relevant data. These substances were predominantly from pharmaceutical companies, and due to commercial confidentiality, they were identified merely by therapeutic class. Most (if not all) of these studies were carried out according to Good Laboratory Practice (GLP) and followed OECD guideline recommendations. They are therefore considered reliable.

We then conducted a search of the US Food and Drug Administration (FDA) database (Drugs@FDA) for oncology drugs, where 7 substances were found with positive *in vivo* MN data, chromosomal damage data *in vitro*, and relevant blood or plasma concentration data. These were added to the database. In order to support regulatory submissions, it is likely most of these studies would have been conducted to GLP and followed OECD guidelines in place at the time and are therefore considered reliable.

We then identified substances with published positive *in vivo* MN data (mainly from [6] and [7]). Searches were conducted on Google Scholar and PubMed to see if more recent or previously unidentified *in vivo* MN data were available, and to fill in any gaps for *in vitro* chromosomal damage data. Searches were then conducted through PharmaPendium, and again through Google Scholar and PubMed for blood or plasma concentration data. For pharmaceuticals we also searched the Drugs@FDA website for any relevant data included as part of the drug approval process. In several cases, it was not possible to identify a LOED *in vivo*, for example, where a substance was included in a publication as a positive control or reference substance at a single dose inducing large MN responses. However, if a substance was tested at a single dose and induced a weak response that was considered biologically relevant (*e.g.*, around 2-fold above concurrent negative control) it was included. Also, many of the substances with relevant *in vivo* MN and *in vitro* chromosomal damage data did not have blood plasma concentration data in the same species and using the same route of administration as used for the study giving the LOED *in vivo*. However, we were able to add another 40 substances to the database, increasing the number of relevant substances to 83. Most of the data obtained from the published literature would not have been from studies conducted to GLP, and may not have complied with OECD guideline recommendations, but no other formal reliability or quality criteria were applied in selecting these studies for inclusion in the database. We therefore acknowledge that not all published studies will be of the same quality as each other, or as the GLP studies referred to above.

2.2. Identification of LOEC, LOED, and blood plasma exposures at the LOED

The LOEC *in vitro* and LOED *in vivo* were identified, where possible, as the lowest concentration or daily mg/kg dose (respectively) giving rise to a statistically significant response or, in the absence of statistical analysis, giving a $>$ 2-fold increase in response relative to concurrent control, which would usually be considered the minimum biologically meaningful response. For some substances, only a single *in vitro* or *in vivo* study was identified and therefore a LOEC and LOED could be identified

only from these single studies. However, where several studies were performed, the one giving the lowest LOEC, irrespective of cell type used, and the lowest LOED, irrespective of species tested, were used for the comparisons described herein (*i.e.*, comparing the LOEC with the blood or plasma concentration at the LOED).

It is important to note that *in vivo* exposure data from different laboratories and in different publications were presented in different ways. Sometimes data were given as blood concentration and sometimes as plasma concentration, and therefore both terms are used in this paper. Data were rarely given separately for parent compound and metabolite (s), so we assume that, in most cases, the reported concentrations apply to total substance (*i.e.*, parent plus metabolites). Also, in most cases, no data were provided on free *versus* protein-bound test substance. It should also be noted that the *in vitro* LOEC values used herein are based on responses observed at the nominal concentration. Although actual concentrations *in vitro* (*i.e.*, in culture medium), as well as *in vivo* (*i.e.*, in blood or plasma), will be affected by losses due to protein binding, the actual *in vitro* concentration will also be affected by factors such as adsorption to plastic and substance volatility. For example, although dependant on experimental setup, Stadnicka-Michalak et al. [8] recently noted that the actual *in vitro* concentration of a volatile substance can be 1–2 orders of magnitude below the nominal concentration. Importantly, that observation was based on results for cells cultured and exposed in multi-well plates. Hence, the degree to which substance volatility affects the LOEC values used for our analysis, is not known. Twelve substances were positive even at the lowest concentration tested *in vitro* and, in those cases, we noted that the LOEC would either be "$<$ ", or, if the response was weak (*e.g.*, around 2-fold concurrent control) then as "$\leq$ ", the lowest concentration tested. However, as mentioned above, if the increase in CA, MN or small colony MLA mutant frequencies at the lowest concentration tested *in vitro* was of the order of 2–3-fold concurrent control frequency, this was accepted as the LOEC. Seven substances that induced MN *in vivo* were negative for chromosomal damage *in vitro*, and in those cases the highest concentration tested *in vitro* was used for comparison with blood or plasma concentrations at the LOED *in vivo*.

Eighteen substances were positive *in vivo* at all doses tested, and therefore the LOED was likely to be lower than the lowest dose tested. In these cases, the plasma concentration at the LOED was also most likely lower than the plasma concentration at the lowest dose tested, and these were therefore presented as "$\leq$ ".

For some substances, particularly those provided by industry and oncology drugs found on the FDA website, plasma concentration data were obtained in the same study in which MN frequencies were determined, or in a separate in-house study using the same batch of test material, and same strain of rodents. Often, the plasma concentration data were obtained at the same doses used for the MN evaluation, and therefore an accurate determination of the blood or plasma concentration at the LOED was possible. In many of these cases full toxicokinetic profiles were performed and it was therefore possible to identify the C_{max}, which was included in the database wherever possible. However, for many of the published studies, blood or plasma concentration data usually came from a different laboratory to that conducting the MN evaluation, and it was not always clear whether C_{max} was being reported. It was therefore possible (perhaps even likely) that different batches of test material, from different sources, and different vehicles were used. Also, although the same species and route of administration were used, the rodent strain may have been different. Moreover, blood or plasma concentration data, including C_{max} values, were often only available at doses different from those used for MN evaluation, and therefore simple linear extrapolation was used to estimate the blood or plasma concentration at the LOED for MN induction. Since the T_{max} may be different for different dose levels, again we could not be certain whether or not the extrapolated blood/plasma concentration was the C_{max}. Clearly, these uncertainties can lead to inaccuracies when compared to the data from in-house industry studies. However, it was considered useful to include these substances, and thus enlarge the database, despite these uncertainties.

The ratio of the blood or plasma concentration at the LOED to the LOEC *in vitro* was calculated. We then applied arbitrary criteria as follows:

- When the ratio was between 0.5 and 2.0 (*i.e.*, the blood or plasma concentration at the LOED for MN induction was within a factor of 2 of the LOEC), we accepted that these concentrations were similar. If the LOEC and/or LOED could not be clearly identified (and so the concentrations were marked as $<$ or $\leq$) and the blood or plasma concentration at the LOED for MN induction was within a factor of 4 of the LOEC (*i.e.*, the ratio was between 0.25 and 4.0), then we also accepted these as being similar.
- When the ratio was $>$ 2.0, or $>$ 4.0 in the case of concentrations marked as $<$ or $\leq$, the blood or plasma concentration at the LOED was classified as being clearly higher than the LOEC.
- When the ratio was $<$ 0.5, or $<$ 0.25 in the case of concentrations marked as $<$ or $\leq$, the blood or plasma concentration at the LOED was classified as being clearly lower than the LOEC.

These arbitrary approaches were considered reasonable given that dose-spacing was often $>$ 2-fold, experiments and toxicokinetic measurements were often not repeated, and the uncertainties of comparing data from different studies in different laboratories.

2.3. Mode of action (MoA) and treatment conditions

In addition to absorption, distribution, metabolism and excretion (ADME) impacting tissue exposure levels, substances that perturb mitotic machinery (*i.e.*, aneugens) often induce MN only after a critical exposure level has been reached. Since MN can be induced by both clastogenic or aneugenic MoAs, it is reasonable then to question if MoA influences the LOED to LOEC comparisons. Substances provided by industry that were positive for MN *in vitro* were often evaluated for MoA by probing the induced MN with pan-centromeric probes or anti-kinetochore antibodies. In some cases, MoA was determined *in vivo*. Thus, it was possible to identify MoA as being predominantly clastogenic or aneugenic, or mixed (both clastogenicity and aneugenicity). For some of the published substances the MoA is well known, or data on MoA have been published separately, and this was noted together with supporting references. However, where there was uncertainty over the reliability of MoA data, no clear conclusion could be reached. For example, substances that were positive for chromosomal aberrations *in vitro* were initially considered to be clastogenic, unless it was specifically noted that polyploidy and/or hyperdiploidy were induced (*i.e.*, indicators of chromosome loss and a potential aneuploidy mechanism) or there were data from MN studies on the same substance where an aneugenic MoA was also identified. However, if there were only structural chromosomal aberration data or induction of small colony mouse lymphoma mutants, and a mixed MoA could not be excluded, no category was allocated.

The treatment conditions for the *in vitro* studies (*i.e.*, with or without metabolic activation, short or long treatment times) were also identified. This allowed for an assessment of whether any particular treatment condition was more likely to lead to genotoxic effects at the lowest concentration (*i.e.*, where the LOEC was found) than other treatment conditions.

2.4. Data analysis

The degree of agreement between the log transformed LOEC and blood or plasma concentration values at LOED was analysed using a Bland/Altman (B-A) plot, the method of choice for the investigation of agreement between two quantitative methods of measurements [9]. This plots the differences against the means of the pairs of log values of plasma concentration at the LOED and LOEC values for the 83

substances. Analysis was carried out using Minitab (Minitab 16 Statistical Software, Minitab, In State College, PA).

In addition, JMP v15 software's Contingency Platform was used to examine the relationship between the categorical variables Blood or Plasma Concentration at LOED Relative to LOEC (lower than, similar to, higher than) and MoA (clastogenic, aneugenic, and mixed). The hypothesis tested was that the response rates for the different MoAs are the same in each sample category (higher than, similar to or lower than blood/plasma concentration relative to LOEC). The Likelihood Ratio Chi-square test statistic was computed as twice the negative log-likelihood for Model in the Tests table. The Pearson Chi-square test statistic was calculated by summing the squares of the differences between the observed and expected cell counts. The Pearson Chi-square test exploits the property that frequency counts tend to a normal distribution in very large samples.

3. Results

In total, 83 substances (64 with MN, 14 with CA and 5 with MLA data *in vitro*) were identified that met the criteria described above. The relevant results are given in Supplementary Tables 1–3. It can be seen from these tables that:

- 33/83 (39.8 %) substances were positive *in vivo* at blood or plasma concentrations that were **lower than** the LOEC *in vitro*
- 19/83 (22.9 %) substances were positive *in vivo* at blood or plasma concentrations that were **similar to** the LOEC *in vitro*
- 31/83 (37.3 %) substances were positive *in vivo* at blood or plasma concentrations that were **higher than** the LOEC *in vitro*.

It should be noted that the ratios of LOEC to blood/plasma concentration at LOED ranged over several orders of magnitude from the blood/plasma concentration at the LOED being > 10,000x lower than the LOEC (#3 Drug for cancer, Supplementary Table 1) to being > 2000x higher (Antifolate for Oncology, Supplementary Table 3).

The B-A plot is shown in Fig. 1. The B-A plot contrasts the divergence between the *in vivo* and *in vitro* measures compared with a measure of the 'potency' of the chemical (the combined log blood/plasma concentration at the LOED and log LOEC values). The Y-axis is the ratio of the log values for LOEC and blood/plasma concentration at LOED.

The central dotted line shows the overall mean value of log differences between the two values. As it is based on log values, if there is no overall difference between the two logged measures the overall mean will be 0. This line gives an indication of bias, in the context of one of the measures (LOEC or blood/plasma concentration at LOED) giving higher values than the other. In this case, the mean value is close to zero (−0.078 for the 83 compounds; 95 % CI: −0.40 to 0.24). This allows an assessment of any directional bias in the differences. If the blood/plasma concentration at the LOED is larger (in numerical terms) than the LOEC the substance will be plotted above the central horizontal (mean) dotted line; conversely, if the LOEC is larger than the blood/plasma concentration at the LOED it will be below the central horizontal (mean) dotted line. The B-A plot clearly shows the variability in the differences between the *in vivo* and *in vitro* results. The log data points are approximately normally distributed (Figure not shown).

The dashed lines, the Limits of Agreement, approximately 2 standard deviations above and below the mean line show the upper (2.81) and lower (−2.97) Limits of Agreement (ULoA and LLoA, respectively). Approximately 95 % of values would be expected to fall within the limits. Outliers, where there are marked differential values between the two measures, may be of interest. Compounds above the mean line, in general, have smaller LOEC than blood/plasma concentration at LOED values, while those below the mean line have smaller blood/plasma concentration at LOED than LOEC values. Specific outliers might be investigated further. It is clear there is appreciable variability in the difference as can be seen by the width of the 'limits of agreement/2 SD' lines (>3 orders of magnitude in each direction) and six substances are outside these limits, although not appreciably different from the 5 % 'expected' from these limits. Three compounds were above the ULoA (Antifolate for Oncology, Potassium chromate, Sodium arsenite) and three were below the LLoA (#3 Drug for cancer [same class as #2 Drug for cancer], #2 Drug for cancer and #8 Drug for asthma). All were results from *in vitro* micronucleus studies. Differential effects that large are somewhat unexpected. However, whether the width of the 'limits of agreement' are acceptable is not a statistical issue but depends upon the particular field of investigation. In this case the plot shows no obvious systematic difference between the methods as most of the differences are around the line of equality (zero).

The X-axis in Fig. 1 is the mean of the two log transformed measures and the spread of these points along the X-axis shows the variation/divergence in magnitude of the mean LOEC and blood/plasma LOED values. The further to the left a compound is the greater the biological activity, in this case the greater the ability to detect a genotoxic effect. Obviously, if the blood/plasma concentration at the LOED is much

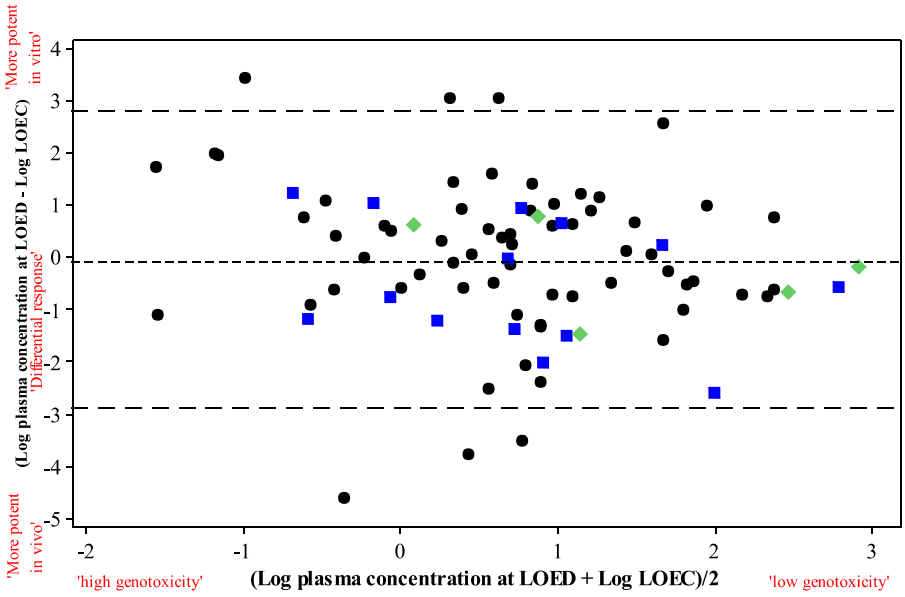

Fig. 1. Bland/Altman plot of the agreement between the log plasma concentration at the LOED and log LOEC values for 83 substances. Data points represent results obtained from *in vitro* AND *in vivo* studies, and all *in vivo* data are from MN studies, The centre line (dotted) is the mean difference between the log plasma concentration at the LOED and log LOEC values for the 83 substances The two other dashed lines, the Upper and Lower limits of Agreement (ULoA and LLoA) are approximately 2 SD above and below the mean difference. Approximately 95 % of values would be expected to fall within this ± 2 SD boundary. For simplicity the plot has been labelled to note that higher (log plasma conc. at LOED+log LOEC) values are 'less genotoxic' while smaller (log plasma conc. at LOED+log LOEC) values are labelled as 'more genotoxic'. (*i.e.*, negative log values are more 'genotoxic'). Therefore, small LOECs or plasma concentrations at LOED indicate easier detection of genotoxic activity *in vitro* or *in vivo* respectively, Legend to colours: black, *in vitro* micronucleus (MN); blue, *in vitro* chromosomal aberrations (CA); green, mouse lymphoma assay (MLA).

larger that the LOEC (or *vice versa*) the point on the plot will be further from the horizontal line but moved towards the centre as one measure will be different from the other. The plot also showed that there is no clear pattern between the log LOEC and log blood/plasma concentration at the LOED values, and no overall relationship with potency. There was, however, a relatively small but highly significant correlation between the log LOEC and log blood/plasma concentration at the LOED values for the 83 chemicals ($r = 0.28$ $P = 0.009$). However, if the objective is to use this relationship to predict the log plasma concentration at the LOED of a substance from its log LOEC value, this would not give accurate predictions.

Thus, there was no clear evidence that *in vivo* plasma exposures needed to reach or exceed the lowest concentrations at which chromosomal damage was induced *in vitro* in order to detect a biologically relevant induction of MN *in vivo*.

The MoA (clastogenic, aneugenic or mixed) could be identified for 52 of the 83 substances. The distribution of substances that were identified as clastogenic, aneugenic or mixed MoAs across the different concentration categories (*i.e.*, positive at lower, similar or higher blood/plasma concentrations than the LOECs *in vitro*) is shown in Table 1. Again, it appears that there was no clear pattern to the responses, with both clastogens and aneugens occurring with approximately equal frequencies amongst those positive at lower, similar or higher blood/plasma concentrations than the LOECs *in vitro*. The results for the two Chi-square statistical tests are shown in Table 1. The p-values for both statistics were > 0.6. These provided no evidence for an association between the ratios and the MoA of the compounds.

For approximately half of the chemicals in the database, it was possible to identify the *in vitro* LOEC from both short (in the absence and presence of S9) and long treatments for the same chemical. However, since different treatment periods occurred in different publications, sometimes with different cell types, such comparisons of the effect of treatment time should be viewed with caution. Nonetheless, from careful analysis of the data in Supplementary Tables 1–3, it can be seen that the LOEC was most often found following long (*e.g.*, at least 20–24 hrs) continuous treatments in the absence of S9, being approximately 7x more frequent than identifying the LOEC following a short treatment (*e.g.*, 3–4 hrs) in the absence or presence of S9.

Table 1
Distribution of clastogenic, aneugenic or mixed modes of action.

Mode of action (MoA)	Blood/plasma concentration at LOED lower than LOEC	Blood/plasma concentration at LOED similar to LOEC	Blood/plasma concentration at LOED higher than LOEC	Total
No. (% of total) clastogenic	5 (29.4 %)	4 (23.5 %)	8 (47.1 %)	17
No. (% of total) aneugenic	7 (26.9 %)	6 (23.1 %)	13 (50.0 %)	26
No. (% of total) mixed clastogenic & aneugenic	5 (55.6 %)	1 (11.1 %)	3 (33.3 %)	9
Total No. (% of total) of substances where MoA identified in each blood/plasma category	17 (32.7 %)	11 (21.15 %)	24 (46.15 %)	52
Test of 3×3 contingency table				
	Chi-square	P		
Likelihood Ratio	2.25	0.69		
Pearson	2.33	0.68		

4. Discussion

It is widely accepted that if a substance being tested for induction of MN in bone marrow is present in plasma, then the bone marrow, which is a well-perfused tissue, will have been exposed. In Kirkland et al. [7] some data were presented showing that the bone marrow is often exposed to higher concentrations of test substance than found in plasma. However, the question of what is "sufficient" exposure of the bone marrow in order to detect a genotoxic effect has not previously been addressed. In this paper, by comparing blood/plasma concentrations at the lowest micronucleus-inducing dose *in vivo* (*i.e.*, the LOED) with the lowest effective concentration *in vitro* (*i.e.*, the LOEC) for induction of chromosomal damage, across 83 chemicals, we have attempted to shed some light on whether blood/plasma concentration alone can be considered "sufficient" to determine whether a negative *in vivo* result can be considered robust.

Given the expected effectiveness of *in vitro* tests at detecting genotoxic substances, for the reasons given earlier, it was quite surprising to find that > 40 % of substances that induce MN *in vivo* do so at blood or plasma concentrations that are lower than cell culture concentrations at which positive responses for chromosomal damage are seen. Indeed, 7 out of these 33 chemicals were negative *in vitro* yet induced MN *in vivo*. These negative *in vitro* outcomes were generally from single studies (although for Ciclesonide both *in vitro* CA and MN studies were negative), but were GLP studies conducted by industry or contract research laboratories, reached much higher concentrations than could be achieved *in vivo*, and are considered reliable. Overall, these findings suggest that MN induction *in vivo* is not driven by concentration alone, and that other factors such as absorption, distribution including tissue accumulation or retention, metabolism and excretion (ADME) will influence the outcome.

It would have been interesting to investigate whether certain cell types (p53-competent *versus* p53-deficient, monolayer cells *versus* suspension cells) were more sensitive than others at predicting *in vivo* positive responses. However, in many cases only one cell type was used for testing a particular chemical, so in those cases it was not possible to identify whether a particular cell type was more sensitive at predicting *in vivo* positive responses. Where different cell types were tested with the same chemical *in vitro* (for example, potassium bromate), the treatment and recovery times are different across the various publications, and therefore a comparison of cell type responses was not possible.

It was also interesting that aneugens and clastogens occurred with approximately equal frequencies amongst those positive at lower or higher plasma concentrations than the LOECs *in vitro*, and therefore there is no reason to expect that an *in vivo* plasma exposure needs to reach or exceed the LOEC *in vitro* in order to specifically detect clastogenic or aneugenic activity *in vivo*.

It is important to note that the results presented herein could provide insight on the performance of directly applying IVIVE (*in vitro*-to-*in vivo* extrapolation) models to *in vitro* concentration-response data for hazard evaluation. IVIVE employs toxicokinetic data (*e.g.*, plasma protein binding and hepatic clearance) to determine the oral dose required to elicit a steady state blood concentration equivalent to the *in vitro* point of departure (PoD) value (*e.g.*, AC_{50}; [1,2]. Calculation of the required oral dose, which is referred to as the administered equivalent dose (AED), assumes that *in vivo* effects will be manifested when the steady state blood concentration reaches the *in vitro* PoD. The results of the analyses presented here suggest that, given the uncertainty in chemical ADME, exposure duration, and *in vitro* disposition, there may be a weak correlation between the expected LOEDs and AEDs derived from LOECs. Specifically, the plasma concentration that corresponds to the *in vivo* PoD (*i.e.*, LOED in this work) is only cursorily related to the *in vitro* PoD (*i.e.*, LOEC in this work). Going forward, it will be necessary to carefully consider the uncertainties related to IVIVE implementation and the impact of these results on the foundational assumptions underlying IVIVE. Indeed, follow-up studies may be warranted, *i.e.*, simultaneous *in*

vivo and *in vitro* studies on carefully selected compounds.

Finally, as was mentioned earlier, collection of LOEC and LOED values in this project were subject to a number of uncertainties. For example, the degree to which substance volatility affects the LOEC values used for our analysis, which are in most cases likely to be from studies that used bent-neck culture flasks with caps, is not known. Future work should investigate the influence of a substance's physicochemical properties on the observed LOEC values. These differences and uncertainties contribute to some inaccuracies in the comparisons reported in this paper. It should be noted that since LOEC and LOED values are affected by study design and dose selection, more robust analyses might be achieved using the BMD (Benchmark Dose) approach [10]. The BMD approach is minimally affected by study design and dose selection, thus providing a less biased PoD metric, *i.e.*, the interpolated dose or concentration required to elicit a set fractional increase above the concurrent control. Follow-up work could explore the ability to use the BMD approach to investigate the relationship between *in vitro* responses and the plasma concentrations associated with *in vivo* responses. That said, the data required to conduct BMD analyses may not be available.

A direct comparison of the effect on the LOEC of different treatment times *in vitro* was not easy because many studies did not include both short and long treatments, or the different treatment periods were reported in different publications, perhaps even using different cell types. However, not unexpectedly the lowest concentration producing a biologically relevant response *in vitro* was associated with long treatments in the absence of S9 on many more occasions (*i.e.*, approximately 7x more frequently) than following short treatments either in the absence or presence of S9. This was even the case for the MLA, where long treatments in the absence of S9 are no longer routinely recommended in OECD guideline 490 [11]. This may suggest that more weight should be placed on the results of longer *in vitro* exposures, but a much larger comparison of *in vitro* and *in vivo* results would need to be performed before such recommendations could be made.

In conclusion, from this analysis the concentration of test substance in blood/plasma that is required to detect induction of MN in bone marrow is clearly not empirically related to the concentration of test substance that results in induction of chromosomal damage *in vitro*. The same lack of an empirical relationship also applies irrespective of whether an aneugenic or clastogenic MoA is involved. What constitutes "sufficient" exposure to detect genotoxic effects *in vivo* is clearly not empirically related to concentrations used in *in vitro* tests and warrants further discussion.

CRediT authorship contribution statement

David Kirkland: Conceptualization, Methodology, Investigation, Resources, Writing – original draft, Writing – review & editing. **James Whitwell:** Investigation, Resources, Writing – review & editing. **Robert Smith:** Investigation, Resources, Writing – review & editing. **Kiyohiro Hashimoto:** Investigation, Resources, Writing – review & editing. **Zhiying Ji:** Investigation, Resources, Writing – review & editing. **Julia Kenny:** Investigation, Resources, Writing – review & editing. **Naoki Koyama:** Investigation, Resources, Writing – review & editing. **Hans-Jörg Martus:** Investigation, Resources, Writing – review & editing. **Krista Meurer:** Investigation, Resources, Writing – review & editing. **Daniel Roberts:** Investigation, Resources, Writing – review & editing. **Akira Takeiri:** Investigation, Resources, Writing – review & editing. **Yoshifumi Uno:** Investigation, Resources, Writing – review & editing. **Bas-Jan van der Leede:** Investigation, Resources, Writing – review & editing. **Andreas Zeller:** Investigation, Resources, Writing – review & editing. **David P Lovell:** Writing – original draft, Formal analysis**. Paul White:** Writing – original draft, Writing – review & editing, Formal analysis.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Dr Stephen Dertinger for the chi-square calculations of the distributions of clastogens and aneugens.

Funding

No funding was provided for this project.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.mrgentox.2022.503503.

References

[1] B.A. Wetmore, J.F. Wambaugh, S.S. Ferguson, M.A. Sochaski, D.M. Rotroff, K. Freeman, H.J. Clewell 3rd, D.J. Dix, M.E. Andersen, K.A. Houck, R.S. Judson, R. Singh, R.J. Kavlock, A.M. Richard, R.S. Thomas, Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment, Toxicol. Sci. 125 (2012) 157–174.

[2] K. Paul-Freidman, M. Gagne, L.-H. Loo, P. Karamertzanis, T. Netzeva, T. Sobanski, J. Franzosa, A. Richard, R. Lougee, A. Gissi, J.-Y.J. Lee, M. Angrish, J.-L. Dorne, S. Foster, K. Raffaele, T. Bahadori, M. Gwinn, J. Lambert, Rasenberg Whelan, T. Barton-MacLaren, R.S. Thomas, Utility of in vitro bioactivity as a lower bound estimate of in vivo adverse effect levels and in risk-based prioritization, Toxicol. Sci. 173 (2020) 202–225.

[3] D. Kirkland, S. Pfuhler, D. Tweats, M. Aardema, R. Corvi, F. Darroudi, A. Elhajouji, H. Glatt, P. Hastwell, M. Hayashi, P. Kasper, S. Kirchner, A. Lynch, D. Marzin, D. Maurici, J.-R. Meunier, L. Müller, G. Nohynek, J. Parry, E. Parry, V. Thybaud, R. Tice, J. van Benthem, P. Vanparys, P. White, How to reduce false positive results when undertaking *in vitro* genotoxicity testing and thus avoid unnecessary follow-up animal tests: report of an ECVAM workshop, Mutat. Res. 628 (2007) 31–55.

[4] P. Fowler, K. Smith, J. Young, L. Jeffrey, D. Kirkland, S. Pfuhler, P. Carmichael, Reduction of misleading ("false") positive results in mammalian cell genotoxicity assays. I. Choice of cell type, Mutat. Res. 742 (2012) 11–25.

[5] P. Fowler, R. Smith, K. Smith, J. Young, L. Jeffrey, D. Kirkland, S. Pfuhler, P. Carmichael, Reduction of misleading ("false") positive results in mammalian cell genotoxicity assays. II. Importance of accurate toxicity measurement, Mutat. Res. 747 (2012) 104–117.

[6] D. Kirkland, L. Reeve, D. Gatehouse, P. Vanparys, A core *in vitro* genotoxicity battery comprising the Ames test plus the *in vitro* micronucleus test is sufficient to detect rodent carcinogens and *in vivo* genotoxins, Mutat. Res. 721 (2011) (2011) 27–73.

[7] D. Kirkland, Y. Uno, M. Luijten, C. Beevers, J. van Benthem, B. Burlinson, S. Dertinger, G.R. Douglas, S. Hamada, K. Horibata, D.P. Lovell, M. Manjanatha, H.-J. Martus, N. Mei, T. Morita, W. Ohyama, A. Williams, In vivo genotoxicity testing strategies: Report from the 7th International workshop on genotoxicity testing (IWGT), Mutat. Res. 847 (2019), 403035.

[8] J. Stadnicka-Michalak, N. Bramaz, R. Schönenberger, K. Schirmer, Predicting exposure concentrations of chemicals with a wide range of volatility and hydrophobicity in different multi-well plate set-ups, Nat. Sci. Rep. 11 (2021) 4680.

[9] J.M. Bland, D.G. Altman, Statistical methods for assessing agreement between two methods of clinical measurement, Lancet 1 (1988) 307–310.

[10] P.A. White, A.S. Long, G.E. Johnson, Quantitative interpretation of genetic toxicity dose-response data for risk assessment and regulatory decision-making: Current status and emerging priorities, Environ. Mol. Mutagen. 61 (2020) 66–83.

[11] OECD, 2016. In Vitro Mammalian Cell Gene Mutation Tests Using the Thymidine Kinase Gene. OECD Guidelines for the Testing of Chemicals. Organisation for Economic Co-operation and Development, Paris. Test guideline 490, adopted 29th July 2016.