Swansea University
Prifysgol Abertawe

## SCHOOL OF MEDICINE

# Validation of the *hprt* Mutagenicity Assay for Nitrosamine Assessment

Submitted to Swansea University in fulfilment of the requirements for the degree of

## MSc In Medical And Health Care Studies By Research

Abbie Williams

[redacted]

Supervised by Dr. George Johnson

## DECLARATION

I, Abbie Williams, declare that: This thesis has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed: Abbie Williams

Date: 22 April 2026

## STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s). Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed: Abbie Williams

Date: 22 April 2026

## STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed: Abbie Williams

Date: 22 April 2026

**ABSTRACT**

Nitrosamines are a chemical class considered to have mutagenic and carcinogenic potential, primarily exerting their mechanism through DNA alkylation. Under new guidance from the Food and Drug Administration (FDA), it is recommended that nitrosamine impurities are tested in both the enhanced Ames test as well as a second *in vitro* mammalian-cell mutation assay in order to support an acceptable intake of 1500 nanograms/day. Here, we report set-up and validation of the hypoxanthine-guanine phosphoribosyl transferase (*hprt*) mutagenicity assay in-house using L5178Y cells to establish its application domain within the pharmaceutical industry to support robust nitrosamine mutagenicity assessments.

Four Ames-positive nitrosamines were evaluated: *N*-nitrosodimethylamine (NDMA), *N*-nitrosodiethylamine (NDEA), *N*-nitrosodiisopropylamine (NDIPA) and 1-cyclopentyl-4-nitrosopiperazine (CPNP). The results demonstrated concordance with parallel mouse lymphoma assay (MLA) data for three of the four compounds, with two showing positive results (NDMA/NDEA) and one yielding a negative outcome (NDIPA). It was noted that the Lowest Observed Genotoxic Effect Levels (LOGELs) for the positive nitrosamines were significantly higher than those observed in the MLA, suggesting differences in the sensitivity of the two assays. Due to significant inter-replicate variability, a statistically significant increase in mutation frequency for CPNP was not detected using the *hprt* assay whereas CPNP was positive by MLA. This project identified *hprt* assay sensitivity, response data variability and experiment duration as potential drawbacks for effective nitrosamine hazard and risk assessment in the industry setting. However, the project also identified opportunities to address these problems through proof-of-concept development of an automated, image-based, machine learning scoring approach and employment of combined benchmark dose (BMD) modelling to leverage all generated data during point-of-departure estimation.

## Section 1: Introduction

### 1.1 History of Nitrosamines

*N*-nitrosamines are a class of compound which were first documented in the nineteenth century and defined as being created as a result of a chemical reaction between nitrous acid and primary aromatic amines (Griess, 1984). Their role as environmental carcinogens was investigated in the 1970's (Love et al., 1977), prompting many years of research on exposure and carcinogenic effect in humans.

Most small alkyl nitrosamines are known mutagens as determined by the bacterial reverse mutation (Ames) test, which consequently translates to carcinogenicity in rodents (Thresher et al., 2020). However, little data is available on the mutagenicity of these compounds within *in vitro* mammalian test systems, and even less in human-based cell types (Li et al., 2022). It is important to note that metabolic activation by mammalian enzyme systems is required before nitrosamines can exert a mutagenic effect (Shank, 1975), and it has been demonstrated that a key enzyme for the metabolism of small-alkyl nitrosamines is CYP2E1(Kang et al., 2007). *N*-nitrosodimethylamine (NDMA) is one of the most studied nitrosamines and its mechanism has been widely researched. Under metabolic activation NDMA forms highly reactive diazonium ions which can specifically alkylate the O-6 and N-7 positions of guanine, resulting in pro-mutagenic DNA adducts that induce strand breakages if not repaired (Snodin et al., 2024) (Anderson et al., 1996). Additionally, upon metabolic activation, some nitrosamines are also able to generate reactive oxygen species (ROS), which can further increase DNA damage (de la Monte and Tong, 2009), as well as equimolar formation of formaldehyde which is a known mutagen and inhaled carcinogen (Lynch et al., 2024).

In more recent years, exposure to nitrosamines as impurities in drug products has become a problem for the pharmaceutical and healthcare industries. This was first highlighted by the US Food and Drug Administration (FDA) and the UK Medicines and Healthcare Products Regulatory

Agency (MHRA) in 2018, when regulatory bodies were alerted to the presence of NDMA in an angiotensin II receptor blocker (Valsartan) products ((WHO), 2019). Since this, a number of other marketed products have been found to contain nitrosated impurities, resulting in these medicines being recalled and/or voluntarily withdrawn from the market.

There are a number of risk factors that increase the likelihood of both finished products and active pharmaceutical ingredients (APIs) containing nitrosamine impurities. These include use of nitrosating agents (*e.g.*, solvents such as dimethylformamide) in the synthetic route of API manufacture, use of water that may contain trace amounts of nitrite capable of nitrosation of the API, or degradation of the finished product as a result of incorrect storage (*e.g.*, temperature, humidity). Contamination has even been observed in blister packaging of finished products whereby the foil cover contains nitrocellulose, which is capable of releasing nitrogen oxides during the heat-sealing process (EMA, 2025). As a result of the wide-ranging opportunities for APIs to potentially nitrosate, stringent testing and risk management of pharmaceutical products is now required at both candidate selection and post-market stages of development. In addition, industry can attempt to mitigate risks by implementing proper storage practices and using alternative packaging.

The European Medicines Agency (EMA) is also currently examining the potential for nitrosated APIs to form naturally within the human body (Eisenbrand et al., 2024). This underscores the need to fully understand their genotoxic potential and mutagenic potency, enabling characterisation and contextualisation of any associated safety risks. As a result of this, at the current time, much ongoing research is focussed on understanding the carcinogenic potency of nitrosamine-containing compounds, as well as demonstrating that the standard battery of genotoxicity tests proscribed by the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) are sufficient for the detection and effective hazard and risk management of nitrosated APIs.

### 1.2 The Standard Battery for Genotoxicity Testing

Genotoxicity assessment is a significant part of the pre-clinical lifecycle of pharmaceuticals in development prior to their administration to humans in clinical settings according to the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) regulation ICH M3(R2). The genetic toxicology discipline aims to assess candidate compounds at different stages of their development to determine their ability to damage DNA; this testing encompasses intermediates within the synthetic route, to potential degradation products, all the way through to the assessment of the active pharmaceutical ingredient itself (including any unique human metabolites not observed in the pre-clinical phase).

In the pharmaceutical industry, genetic toxicology sits within a group of non-clinical safety disciplines with the primary aim to identify potentially hazardous agents that adversely interact with DNA in order to prevent them from reaching later stages of clinical development (thereby avoiding potential human exposure to genotoxic compounds). This is not only to ensure that new pharmaceuticals are safe for patients, but also to reduce time and cost losses, through successful 'positive attrition'. This is why genetic toxicology screening methods have been developed to be performed at the pre-candidate drug selection phase. Genetic toxicology testing is highly regulated, and a battery of tests has been determined by the ICH to ensure effective testing of pharmaceuticals that covers all genotoxicity endpoints. The 'standard battery' testing strategy includes an *in vitro* test for mutagenicity in bacteria *e.g.,* the Ames test, an *in vitro* mammalian test for mutation and/or chromosomal damage *e.g.,* mouse lymphoma mutation assay (MLA), micronucleus test (MNT) or chromosome aberration test, and lastly an in vivo test in rodents. The default *in vitro* mammalian cell assays are those that assess micronucleus formation or chromosome aberration (ICHS2(R1), 2011). Other assay endpoints, *e.g., in vivo* Comet assays and transgenic rodent gene mutation assays can be used to follow-up

a specific *in vitro* result, as appropriate. If a pharmaceutical achieves negative outcomes in the test battery it is considered that there is sufficient evidence to discharge any genotoxic liability and the compound can continue in its development towards clinical trial (ICHM3 (R2), 2009).

For nitrosamine drug substance-related impurities (NDSRIs), the testing strategy is constantly evolving. Currently, the FDA recommends an enhanced Ames assay as well as a second *in vitro* mammalian cell mutation assay in order to support an acceptable intake (AI) of 1500 ng/day. To consider increasing the AI above 1500 ng/day, an additional *in vivo* mutagenicity study (*i.e.,* a transgenic rodent assay) can be conducted. Currently, the FDA does not express a preference for *in vitro* mammalian test type (*i.e.,* MLA, *hprt* or MNT). However there is increasing interest from the regulatory bodies to include *hprt* data in supporting documentation when requesting an increase in the AI of the related nitrosamine impurity, due to its recognised ability to detect small mutational events *e.g.*, point mutations and exon deletions (Johnson, 2012), which are specific to the type of genetic damage that nitrosamines are known to cause (de la Monte and Tong, 2009).

### 1.3 Hypoxanthine–Guanine Phosphoribosyl Transferase (*hprt)* Assay

The hypoxanthine–guanine phosphoribosyl transferase (*hprt*) assay is an *in vitro* test system used in the genetic toxicology discipline to detect chemically induced gene-mutations through utilisation of the sex-linked loci of the *hprt* gene. The *hprt* gene is located on the X-chromosome and exists as a single functional copy in mammalian cells, due to either male cells only having one copy of the X-chromosome and female cells undergoing inactivation of one the copies (Aidoo et al., 1997). The *hprt* gene codes for hypoxanthine phosphoribosyltransferase 1, a transferase which is essential to the purine synthesis salvage pathway (Guo et al., 2022). Its specific function is to enable phosphoribosylation of guanine or hypoxanthine which allows them to be salvaged to produce guanosine 5'-monophosphate (GMP) or inosine 5'-monophosphate (IMP), respectively. IMP can then be converted to adenosine 5'-

monophosphate (AMP) or GMP, which can be further phosphorylated to produce nucleotide di- or tri-phosphates which can be used for DNA synthesis (Tran et al., 2024).

Purine synthesis occurs via one of two pathways (*de novo* or salvage), enabling purines to be created either from scratch or by recycling readily available nucleotide components from either the diet or of the results of nucleotide catabolism. As the *de novo* pathway is a metabolically costly process involving 10 catalytic steps to assemble the purine ring, it is thought that cells predominantly use the salvage pathway where possible to synthesise nucleic acids as it requires only 1 ATP molecule per purine created and therefore is a more efficient method of DNA synthesis (Tran et al., 2024).

As mentioned above, a critical enzyme in the salvage pathway is *hprt* which not only helps recycle purines but also acts on activation of purine analogues such as 6-thioguanine (6TG) which, due to its cytotoxicity, is used in the *hprt* assay as a selective agent (Albertini, 2001). In the assay, wild-type cells (*hprt*$^{+}$) with a functional *hprt* enzyme activate and incorporate the toxic 6TG through the salvage pathway which leads to cell death (Christie et al., 1984), whereas mutant cells (*hprt*$^{-}$), lack functional *hprt* and therefore cannot utilize this pathway. Instead, they rely on the *de novo* synthesis route, which avoids the incorporation of 6TG, allowing these mutant cells to survive and form scorable colonies thus enabling creation of the *hprt* test system (see Figure 1).

Traditionally, Chinese hamster ovary (CHO) cells have been used to investigate these mutations (Moore et al., 1991), however use of the assay with suspension cells such as mouse lymphoma L5178Y and human lymphoblastoid AHH-1 have also been documented (Crespi and Thilly, 1984, Parry et al., 2005). In this work, L5178Y cells have been selected to allow for direct comparison against mutant frequency in the thymidine kinase (TK$^{+/-}$) mouse lymphoma assay.

The *hprt* assay is designed to detect small mutations to DNA such as base pair mutations, frameshift mutations and small deletions or insertions. This is due to the fact that large losses

to the X chromosome (*i.e.*, as a result of clastogenicity) are fatal and typically result in cell death. This makes the *hprt* assay highly specific for detecting mutations only, compared with the mouse lymphoma assay (MLA) where mutant colonies can be as a result of either mutagenicity or clastogenicity. The MLA assay is described in more detail in Section 1.4. Due to this ability to detect small mutation events, the *hprt* assay has been discussed within industrial groups such as the European Federation of Pharmaceutical Industries and Associations (EFPIA) as the potential best option to evaluate the mutagenicity of nitrosamines, as it rules out any questions that may be raised over other mechanisms of genotoxicity (*i.e.,* clastogenicity).

**1.4 Mouse Lymphoma Assay (MLA)**

The mouse lymphoma assay (MLA) forms part of the genotoxicity test battery recommended by the ICH for testing of pharmaceuticals and is outlined in The Organisation for Economic Co-operation and Development (OECD) Test Guideline 490. Similar to the *hprt* assay, the premise of the MLA relies on induced genotoxic damage to the thymidine kinase (*tk*) gene of L5178Y cells resulting in phenotypic resistance to the selective agent 5-triflurothymidine (TFT), a cytotoxic purine nucleotide analogue. Unlike the *hprt* assay, the MLA utilises a heterozygous *tk* gene ($tk^{+/-}$) which is not essential for survival and therefore mutation to its inactive state does not necessarily result in cell death. *Tk* catalyses the phosphorylation of thymidine deoxyriboside to form deoxythymidylate, at which point two further phosphate groups are added and the complex is incorporated into the DNA to create a phosphodiester bridge and a functional piece of DNA (Lloyd and Kidd, 2012). The functional *tk* enzyme (product of $tk^{+/-}$ genotype) allows incorporation of the toxic analogue TFT into the DNA, thus killing the cell. Cells that are homozygous recessive and possess 2 copies of the mutated form of the *tk* gene (*i.e.,* $tk^{-/-}$) are able to survive due to the inactivation of the salvage pathway (like in the *hprt* assay) and form mutant cell colonies when cultured in 96-well plates containing TFT (see Figure 1).

However, unlike the *hprt* assay, mutant clones may form small or large colonies. It is thought that small colonies originate from intergenic mutations (*i.e.,* mutations that extend outside of the *tk*-locus due to, for example, clastogenic events) whereas large clones originate from intragenic mutations (*i.e.,* mutations that occur within the *tk*-locus, for example single base substitutions, and small indels). The small colony phenotype is due to the autosomal location of the *tk* gene, whereby larger chromosomal alterations such as large deletions, chromosomal rearrangements and mitotic recombination can result in a loss of heterozygosity leading to a non-functional enzyme and also the loss of genes that regulate growth, which are located close to the *tk* gene. The colonies are smaller since loss of growth regulation genes extends cell doubling times causing colonies to grow at a much slower rate (Amundson and Liber, 1992). In contrast, the *hprt* gene is X-linked, which means loss of large fragments of DNA typically result in cellular death.

### 1.5 Summary: *Hprt* versus MLA test systems

The MLA and *hprt* assays are both important components of the genotoxicity testing battery, used to assess the potential mutagenic (and clastogenic) properties of chemicals. However, while they share some similarities in their reliance on nucleotide salvage pathways and the use of toxic purine analogues, the differences in target gene locations allow them to be used for different testing objectives.

One similarity between the two assays is their utilisation of cytotoxic purine analogues such as TFT and 6-TG to select for mutations in *tk* or *hprt*, respectively, as mentioned in the sections above. In both cases, the functional mechanism depends on a disruption in the associated salvage pathway, thereby permitting mutant cells to survive and form colonies under selective pressure using toxic nucleotide analogues. They can also both be performed in the L5178Y cell line which means, for the purposes of this work, that results from both assays can be compared to look at the effect of different chemicals on different reporter genes.

Despite this commonality, a significant difference between these assays lies in the reporter genes they use. As mentioned in Section 1.4, the MLA utilises the heterozygous *tk* gene and is sensitive to a broader range of DNA damaging agents; it identifies mutagens, clastogens, and some aneugens, owing to the potential formation of both small and large colonies reflecting different types of genetic disruptions. This flexibility renders the MLA a versatile tool for assessing a wider range of genotoxic influences. Meanwhile, the *hprt* assay targets the *hprt* gene situated on the X-chromosome, exploiting its sex-linked nature, and is sensitive to intragenic mutations *i.e.,* base pair mutations, small frameshift mutations, and small deletions or insertions (indels) (Keohavong et al., 2005).

An additional difference between the assays is that the *tk* gene typically has a relatively short expression period for mutations of around 2 days (OECD Test Guideline 490). As described, *tk* is a highly expressed enzyme essential for the salvage pathway of **pyrimidine** nucleotide synthesis, making it quick to demonstrate loss of function when mutations occur. In contrast, the *hprt* gene has a longer expression period for mutations which is approximately 7-9 days. This extended timeframe is primarily due to the cellular and metabolic processes related to **purine** salvage pathways, where mutations in the *hprt* gene gradually result in loss of function (Guo et al., 2022). In essence, the *hprt* mutation assay requires a longer manifestation period than the MLA. More specifically, the time to reach detectable colonies is longer as it takes several cell divisions for the effects of the mutation to manifest in terms of loss of function, and there is lower cellular turnover to generate visible colonies. This means that performing a *hprt* assay takes double the time of an MLA.

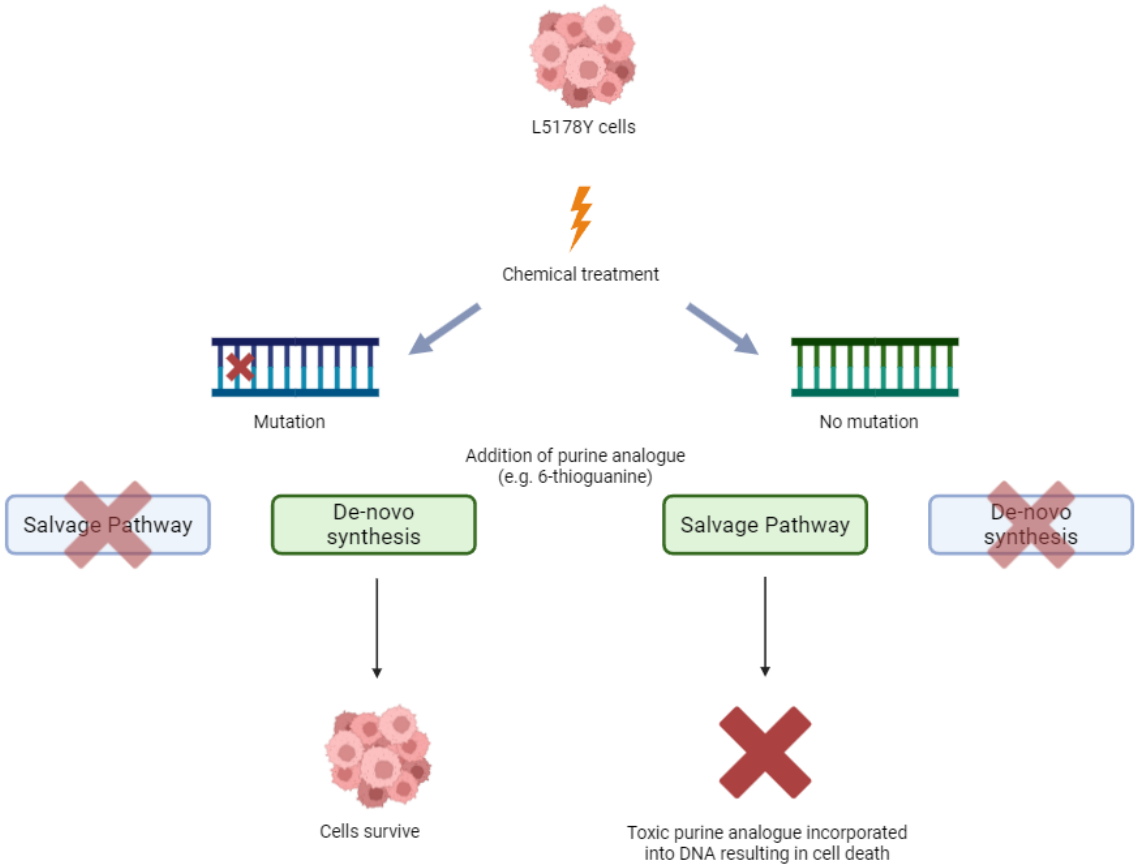



**Figure 1 – Schematic overview of the hprt test system.** Mutational inactivation of the hprt gene required for the salvage pathway forces cells to synthesise DNA via the *de-novo* pathway, allowing cells to survive 6TG selection to form scorable colonies.

### 1.6 Objectives

The objective of this project was to set up an OECD-compliant *hprt* assay in-house using tool mutagens and small-alkyl nitrosamines to establish a reliable test-system that is ready to use to support drug discovery impurity qualifications and aid robust safety assessment at GSK. To achieve this, the first main aim was to establish the assay with known genotoxic compounds benzo[a]pyrene and methyl methanesulfonate, as instructed by OECD guidelines. This allowed progression onto screening of four tool nitrosamine compounds (N-nitrosodimethylamine (NDMA), N-nitrosodiethylamine (NDEA), N-nitrosodiisopropylamine (NDIPA) and 1-cyclopentyl-4-nitrosopiperazine (CPNP)) at increasing concentrations to assess their behaviour in the assay

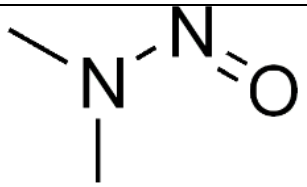
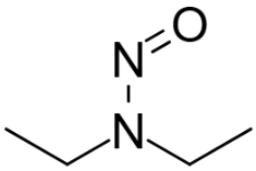
and to provide data to support regulatory decision making and robust intake limits. In this way, development using model nitrosamine compounds established the application domain of the *hprt* assay within the pharmaceutical setting in terms of its ability to support robust nitrosamine mutagenicity assessments. The generated *hprt* data was compared to that obtained from the MLA assay helping to assess assay sensitivity and elucidate the potential of the *hprt* assay in the industrial setting in terms of discharging human safety concerns surrounding nitrosamine impurities. This work also highlighted the practical benefits and drawbacks of each assay system, providing not only a scientific justification for the use of one *in vitro* mammalian assay over the other, but one shaped on practical experience. Additionally, to maximise the utility of the collected data, a further aim of this work was to explore how other newly-evolving methods such as artificial intelligence can be harnessed to automate *hprt* colony scoring with the goal of reducing human resource and enabling more replicates to be analysed to increase assay sensitivity. Combined benchmark dose modelling is also explored as a means to improve the utility of *hprt* dose-response data to improve potency ranking and point-of-departure estimations.

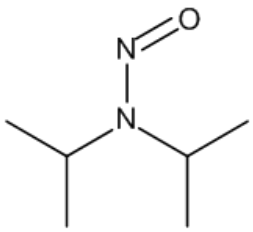
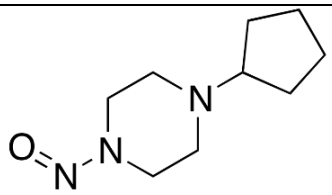
### 1.7 Tool Nitrosamines and Mechanisms

Although all nitrosamines require the same initiating step of metabolic activation via cytochrome P450 enzymes, they differ in their mechanism of genotoxicity based on their structure. This can affect their downstream potency due to the type of damage caused *i.e.*, what kind of adducts form, how stable those adducts are and whether they are efficiently repaired by DNA repair mechanisms. Table 1 below outlines the four tool nitrosamines used in this work and their mode of action and potency, where data has been generated.

**Table 1 – Structural and mechanistic information for the four tool nitrosamines investigated in this work.**

| Name | Structure | Mode of Action | Potency |
|---|---|---|---|

| | | | |
|---|---|---|---|
| NDMA |  | Undergoes α-hydroxylation to produce a methyldiazonium ion and formaldehyde which are both DNA reactive. Rodent experimental data shows that this causes increases in N7 and O6-methylguanine adducts (Li and Hecht, 2022, Carlsson et al., 2025). These adducts result in G → A transitions which can be repaired by O6-methylguanine-DNA methyltransferase (MGMT) however, when not repaired, manifest primarily as liver tumours in rats and mice (Zhou et al., 2025). | • IARC Group 2A.<br>• Acceptable Intake (AI) of 96 ng/day.<br>• NDMA is among the highest relative potency simple dialkylnitrosamines based on mutagenicity and carcinogenicity datasets (Bartsch et al., 1976). |
| NDEA |  | Undergoes α and β-hydroxylation to produce ethyldiazonium and 2-hydroxyethyldiazonium ions, respectively. This results in N7 and O6-ethylguanine adducts which also result in G → A transitions which can be repaired by MGMT. O2 and O4-ethylthymine adducts are also formed however are poorly repaired, and these are strongly mutagenic. Rodent experimental data links NDEA adducts to potent liver and oesophageal carcinogenesis (Li and Hecht, 2022). | • IARC Group 2A.<br>• Acceptable Intake (AI) of 26.5 ng/day.<br>• NDEA is a very high potency nitrosamine, with data suggesting that it is quantitatively more potent that NDMA in-vivo based on differences in methyl vs. ethyl repair efficiency (Li and Hecht, 2022, Bartsch et al., 1976). |

| | | | |
|---|---|---|---|
| NDIPA |  | Undergoes α-hydroxylation however steric hindrance from the isopropyl groups slows metabolism. The production of alkyldiazonium ions result in branched alkyl adducts, such as N7-isopropylguanine (Cross and Ponting, 2021).<br><br>Although there is no literature to support NDIPA-specific adduct mapping, it is predicted that NDIPA results in O6-isopropylguanine adducts which are much harder to form due to steric hindrance at the O6 position (Li and Hecht, 2022). | • IARC Group 2A.<br>• Acceptable Intake (AI) of 26.5 ng/day.<br>• NDIPA is harder to metabolise due to its steric hindrance, therefore requiring more complex activation and potentially lowering its potency. Additionally, due to steric hindrance limiting O6 adducts, this again potentially reduces potency. |
| CPNP |  | Has multiple α-hydroxylation sites which may require several CYP enzymes to metabolise.<br><br>Again, there is currently no literature to support CPNP-specific adduct mapping. The nitroso-piperazine class generally results in O6-alkylguanine or hydroxyalkyl guanine adducts, however extrapolation to CPNP remains hypothesis level (Dos Santos et al., 2022). | • Considered a possible human carcinogen based on its structural similarity to the known rodent carcinogen 1-nitrosopiperazine.<br>• Little data is available on the mechanism or potency of this nitrosamine. |

## Section 2: Materials

### 2.1 Test article

Test Article Formulation:

Solubility assessments were performed prior to testing in order to select the most appropriate solvent that allowed maximum exposure up to the solubility limit or the maximum test concentration (10 mM or 2 mg/mL) in accordance with OECD guidelines. Additionally, solvent choice was prioritised based on guidance from the Health and Environmental Sciences Institute Genetic Toxicology Technical Committee (HESI GTTC) and the Food and Drug Administration (FDA) on testing of nitrosamine impurities in the bacterial reverse mutation assay (Ames Test). In this project, test articles were solubilised in either water, methanol or dimethyl sulphoxide.

Concentration Selection:

Preliminary dose range-finder (DRF) experiments were performed in single cultures to establish an appropriate limiting test concentrations based on precipitation, cytotoxicity or solubility. In this initial DRF, a range of at least 10 concentrations ranging down from the OECD Test Guideline limit (10 mM, 2 mg/mL or 2 µL/mL, whichever is lowest) were tested. Following this, in cases of cytotoxicity, assessment of relative survival (RS) measured post-treatment influenced concentration selection for the main test, in which at least 6 concentrations ranging from toxic to non-toxic were chosen to be tested in triplicate. In this case, the maximum concentration chosen for analysis was where the relative survival (RS) post-treatment was reduced to 10-20%. In cases where test article precipitation was observed by eye at the end of treatment, the maximum concentration chosen for analysis in the main test was the lowest concentration at which precipitation was observed. Finally, in cases where no precipitation of test article or cytotoxicity were observed, the maximum test concentration for the main test was 10 mM, 2 mg/mL or 2 µL/mL, whichever was lowest, in accordance with OECD Test Guideline 476.

**2.2 Cell Line**

Vials of Mouse Lymphoma L5178Y $TK^{+/-}$ cells were obtained from the European Collection of Authenticated Cell Cultures (ECACC) and subsequently cultured in-house to provide a working stock of cells which were frozen down and stored in liquid nitrogen at GSK, Stevenage (termed the 'GSK Master Stock (MS) 12'). Following some preliminary testing using GSK MS 12, vials of L5178Y $TK^{+/-}$ cells were obtained from Labcorp (Harrogate, UK) and cultured in-house to provide an additional working stock of cells which were frozen down and stored in liquid nitrogen at GSK, Stevenage (cell stock hereafter termed 'HPRT WS1'). Cells were cultured with specific culture medium (CM10) comprised of RPMI 1640 GLUTAMAX-1 medium containing 3.0 mM L-glutamine and 25 mM HEPES and supplemented with 1.8 mM sodium pyruvate, 50 µg/mL streptomycin sulphate, 50 IU/mL penicillin, 0.1% pluronic F-68 and 10% (v/v) heat inactivated donor horse serum (HIDHS).

Both cell stocks (GSK MS 11 and HPRT WS1) had been cleansed when the new working stock was created, using hypoxanthine-aminopterin-thymidine (HAT) medium to ensure spontaneous mutant frequency (MF) was within the acceptable range (2-50 x $10^{-6}$) determined by OECD guidelines.

**2.3 Metabolising System**

An exogenous metabolising system of co-factor containing nicotinamide adenine dinucleotide phosphate (NADP) and glucose-6-phosphate (G6P) supplemented with post-mitochondrial liver fraction (S9) was used in 3-hour treatments that specify '+S9'. Phenobarbital, 5,6- benzoflavone-induced Sprague-Dawley rat S9 was used (supplied by Trinova Biochem, Germany). The S9-mix was prepared at 10% to give a final concentration of 1% (v/v) in the test-system.

**2.4 Controls**

Six vehicle negative controls were included in each test (*i.e.,* if the test compound was solubilised in methanol, six methanol vehicle controls (1% v/v) were included concurrently). Positive controls were not included in DRF experiments. In the main test, benzo[a]pyrene at concentrations of 0.5-1.5 µg/mL were used as a positive control in the presence of S9-mix. Positive controls were formulated as stock solutions and stored as frozen aliquots, ready for use prior to testing.

**Section 3: Methods**

Approximately 6 days prior to day 0, a 1 mL vial of frozen cell stock was thawed and added to 49 mL CM10 in a 75 $cm^3$ tissue culture flask. These were subsequently sub-cultured at 2 and 5 days using 175 $cm^3$ tissue culture flasks to ensure density did not exceed $1x10^6$ cells/mL in order to reach and sustain exponential growth phase. Cells were incubated at 37 ˚ C in a humidified $CO_2$ incubator (5% $CO_2$).

**3.1 Treatment**

Stock solutions of the various concentrations of test article were prepared and stock aliquots of positive controls were thawed ready for use.

An exponentially-growing stock culture of cells in CM10 media were shaken to produce a single cell suspension. From this, 0.5 mL was added to 9.5 mL isotonic solution and the cell titre was determined using a Coulter counter (Beckman Coulter, UK). Aliquots of the cell stock were transferred into 50 mL centrifuge tubes and centrifuged at 200 x g (approximately 1000 RPM) for 5 minutes. The supernatant was removed and the cell pellets were thoroughly resuspended and combined in CM5 media. This suspension was mixed gently and a 0.5 mL sample was taken and counted to confirm cell count. The target cell concentration was $1x10^7$ cells per culture (approximately $5x10^5$ cells/mL).

If a +S9 treatment was conducted, S9-mix was prepared using co-factor and S9 fraction at 10% (v/v).

Each culture was set up at the following volumes in a 50 mL falcon tube:

**Table 2 - Final volumes of constituents in each culture.**

| Treatment | Cell volume (mL) | S9-mix volume (mL) | Vehicle/Test substance volume (mL) | Final culture volume (mL) |
|---|---|---|---|---|
| **Presence of S9** | 17.8 | 2 | 0.2 | 20 |

Each culture contained 17.8 mL of cells, 2 mL of S9-mix and 0.2 mL of vehicle or test substance to produce a final treated culture of 20 mL.

Following treatment with either vehicle or test article, the time was recorded and tubes were incubated at 37˚ C in a shaking incubator for the duration of treatment (with sufficient agitation only to prevent the cells from settling).

**3.2 Post Treatment (Day 0)**

At 3 hours post-treatment, cultures were examined by eye for the presence of precipitation. Treatment groups showing precipitation were discarded, leaving the remaining maximum concentration as the lowest concentration at which precipitation was visible by eye. Following this, the remaining cultures were centrifuged at 200 x g (1000 RPM) for 5 minutes to pellet the cells. The supernatant was gently poured off and discarded, ensuring the pellet was not disturbed. The pellet was then resuspended in 5 mL pre-warmed phosphate buffered saline (PBS). The cells were centrifuged again at 200 x g (1000 RPM) for 5 minutes and the supernatant removed and discarded as before. The pellet was then resuspended in 20 mL of CM10. At this point, 0.5 mL of each culture was taken and added to 9.5 mL isotonic solution and each culture was counted using the Coulter counter. Cell counts were adjusted with CM10 medium for each

sample in order to achieve $2x10^5$ cells/mL culture. This was then transferred to 25 $cm^3$ vented tissue culture flasks and incubated at 37˚C. All cell titres and dilutions were recorded.

*Survival Plating*

Following adjustment of cultures to $2x10^5$ cells/mL as described above, samples were diluted to ~8 cells/mL using the dilution scheme in the table below.

**Table 3 – Scheme for dilution of viability plate samples to achieve a final concentration of 8 cells/mL.**

| **Initial cell concentration** | **Dilution** | | **Intermediate cell concentration** | **Dilution** | | **Final cell concentration** |
|---|---|---|---|---|---|---|
| (A) | mL A | mL CM20 | (B) | mL B | mL CM20 | |
| $2x10^5$/mL | 0.1 | 9.9 | $2x10^3$/mL | 0.2 | 50 | 8 cells/mL |

0.1 mL of the original $2x10^5$ cells/mL culture (A) was added to 9.9 mL of CM20 culture medium to produce an intermediate cell concentration of $2x10^3$ cells/mL (B). 0.2 mL of B was added to 50 mL of CM20 culture medium to produce a final sample containing 8 cells/mL.

Using a multichannel pipette, 0.2 mL of ~8 cells/mL culture was placed into each well of 2 x 96 well microtitre plates (total of 192 wells, averaging 1.6 cells/well). These plates were then incubated at 37 ˚ C for 7 days. In the case of a DRF study design, the experiment would conclude at this point and the data from the survival plating would inform concentration selection for the main test.

**3.3 Days 1, 3 and 6 (Expression Period)**

On days 1, 3 and 6, flasks were removed from the incubator and vigorously shaken to ensure that cultures were single-cell suspensions. 0.5 mL of cells were added to 9.5 mL isotonic

solution and each culture was counted using the Coulter counter. Cell concentrations were adjusted to $0.25x10^5$, $0.08x10^5$ and $2x10^5$ cells/mL on days 1, 3 and 6, respectively.

**3.4 Day 7**

*Survival Plate Scoring*

Following 7 days of incubation, the survival plates were scored using a mirror box for the absence of colonies in each well. The number of empty wells were recorded in the *hprt* Excel workbook to calculate mean percentage survival relative to the vehicle control. This survival data was then used to inform concentration selection for viability and mutation plating.

*Plating for Viability*

Following 7 days of sub-culturing enabling mutation fixation, flasks were removed from the incubator and vigorously shaken. Similar to the survival plating process, cells were counted and adjusted to $1x10^5$ cells/mL in CM20 medium. From this $1x10^5$ cells/mL stock, 0.2 mL was removed and added to 9.8 mL of CM20 to give an intermediate cell concentration of $2x10^3$ cells/mL. From this, 0.2 mL was removed and added to 50 mL of CM20 to give a final concentration of 8 cells/mL.

Using a multichannel pipette, 0.2 mL of 8 cells/mL culture was placed into each well of 2 x 96 well microtitre plates (total of 192 wells, averaging 1.6 cells/well). These plates were then incubated at 37˚C for 7 days. Following the incubation period, these plates were scored in the same way as survival plates (Section 3.4, above).

*Mutant Frequency Plating*

In order for a culture to be accepted for mutation plating, the cell concentration must have been high enough to yield at least 80 mL of cells at the desired concentration ($1x10^5$ cells/mL). Following removal of 0.2 mL of the $1x10^5$ cells/mL stock for the viability dilution, a pre-calculated volume of stock 1.5 mg/mL 6-thioguanine (6TG) was added to each culture to give a

final concentration of 15 µg/mL. Using a multichannel pipette, 0.2 mL of $1x10^5$ cells/mL culture containing 6TG was placed into each well of 4 x 96 well microtitre plates (total of 384 wells, averaging $2x10^4$ cells/well). These plates were then incubated at 37˚ C for 10-14 days. Following the incubation period, individual wells were assessed for the presence of colonies >20 cells in diameter (Johnson, 2011) and recorded in the *hprt* Excel workbook. As there was regularly debris in each well as the result of $hprt^+$ cells that had been killed by the 6TG selection, it was important not to score these as mutant colonies.

**Section 4: Data Analysis**

All calculations were performed using an automated Excel workbook.

The following calculations were performed:

**4.1 Survival Cultures (plated on Day 0)**

Cloning Efficiency (CE)

The zero term of Poisson distribution (P) is initially calculated, based on the ratio of empty wells to total wells scored:

P = -ln (total empty wells/total wells counted) (Equation 1)

The CE for each culture is then calculated based on the number of cells plated per well (*i.e.,* 1.6 cells per well):

CE = P/1.6 (Equation 2)

Survival

Initially, a day-zero adjustment factor is applied to the cell count of each culture on day 0, post wash-off. This adjusts for any loss of cells during the 3 hour treatment period:

(Equation 1)

$$\text{Day zero factor} = \frac{\text{cell count on Day 0 (post wash off)}}{\text{mean vehicle control cell count on Day 0 (post wash off)}}$$

%Survival is then calculated:

$$\%\text{survival} = CE * \text{day zero factor} \qquad \text{(Equation 2)}$$

Percentage relative survival (%RS) for each concentration is determined by comparing the cloning efficiency of the test article treatment (mean) vs the concurrent vehicle control (mean):

$$\%RS = [\ \text{Mean } CE_{(test)} / \text{Mean } CE_{(vehicle)}\ ] * 100 \qquad \text{(Equation 3)}$$

**4.2 Mutant and Viability Cultures (plated on Day 7)**

In order to calculate mutant frequency (expressed as mutants/$10^6$ viable cells), the cloning efficiency of both mutant and viable cells from the same culture are calculated.

Viability Cloning Efficiency ($CE_{viable}$)

The zero-term of the Poisson distribution ($P_{viability}$) is initially calculated, based on the ratio of empty wells to total wells scored:

$$P_{viability} = -\ln(\text{total empty wells/total wells counted}) \qquad \text{(Equation 1)}$$

The CE for each culture is then calculated based on the number of cells plated per well (*i.e.,* 1.6 cells per well):

$$CE_{viability} = P/1.6 \qquad \text{(Equation 2)}$$

Mean Relative Viability (%RVc) is then calculated for each test article concentration:

$$\%RVc = [\ \text{Mean } CE_{(test)} / \text{Mean } CE(_{vehicle})\ ] * 100 \qquad \text{(Equation 3)}$$

Mutant Frequency

The zero term of the Poisson distribution ($P_{mutant}$) is initially calculated, based on the ratio of empty wells to total wells scored:

$$P_{mutant} = -\ln (\text{total empty wells/total wells counted}) \quad \text{(Equation 1)}$$

The CE for each culture is then calculated, based on the number of cells per well plated for mutation (*i.e.*, 20,000)

$$CE_{mutant} = P_{mutant} / 20000 \quad \text{(Equation 2)}$$

The mutant frequency can then be calculated:

$$MF = CE_{mutant} / CE_{viability} \times 10^6 \quad \text{(Equation 3)}$$

Mean MF for each treatment and vehicle, as well as the relative fold increase over the control were also calculated.

**Section 5: Statistical Analysis**

**5.1 Pair-wise testing**

A tiered approach was applied to statistically analyse the *hprt* dose-response data, using the framework detailed by Johnson et. al.(2014). Raw, log10 and square-root transformations of the response data were prepared before assessment for best transformation using Bartlett and Shapiro-Wilk tests to consider homogeneity of variance and distribution normality respectively. Alongside these preliminary distribution tests, an assessment for the presence of a monotonic trend (Jonckheere-Terpstra test) was also run. Datasets that passed these tests ($p > 0.05$) were then taken forward for pair-wise testing response significance relative to untreated vehicle controls by one-sided Dunnett's tests with alpha set at 0.05. Statistical analyses were conducted in the R programming environment using the DRSMOOTH R-package (Johnson et al., 2014).

**5.2 Benchmark dose modelling**

A pilot sub-set of compounds exhibiting a positive mutant frequency dose-response were then taken forward for non-linear regression analysis using the Benchmark Dose (BMD) approach to rank compound potencies (Wills et al., 2016a). The data for these compounds were analysed using the PROAST (version 71.1) R-package using the exponential model family recommended by the European Food Safety Authority (EFSA) for the analysis of continuous toxicity data (EFSA, 2022). When combined modelling was used, data were analysed using compound as covariate. PROAST uses the Akaike information criterion to assess whether the use of increasingly complex models with additional parameters achieve a significant improvement of model fit where more complex parameterisations were only accepted if the improvement exceeded $p < 0.05$ to avoid overfitting the dose-response data. In turn, interpolation at a benchmark response size of 50% was used to determine the Critical Effect Dose (CED) as well as its upper and lower 95% confidence limits (termed the CEDU and CEDL, respectively) providing a quantitative definition of compound potency. Confidence interval plots of the results of both individual and combined analyses were used to provide a visually intuitive representation of compound potencies (EPA, 2012).

## Section 6: Assay Acceptance and Evaluation Criteria

### 6.1 Assay Acceptance Criteria

The assay was considered valid if the following criteria were met:

- In the absence of a laboratory historical control range, vehicle negative controls were assessed against published literature. It was expected that for L5178Y cells, vehicle spontaneous MF should fall within the published range of 2-50 x $10^{-6}$ (Moore et al., 1991, DeMarini et al., 1988).
- The positive control must produce a statistically significant increase when compared against the concurrent negative controls. After appropriate data transformation for

distribution normality and variance homogeneity this was assessed using a Dunnett's one-tailed T-test.

- The top concentration must be limited by either 10 mM, solubility, precipitation or cytotoxicity (whichever occurs first), as described above in Section 2.1.

**6.2 Assay Evaluation Criteria**

In order to call a test compound positive for induction of gene mutations within mammalian cells, the following criteria had to be met:

1. At least one of the test concentrations must demonstrate a statistically significant increase in MF when compared against concurrent controls using an appropriate statistical test (*e.g.*, Dunnett's T-test).
2. This increase must be trend-tested in order to show that it is concentration-dependent.
3. OECD test guidelines state that the positive test concentration must be outside the distribution of the historical negative control data. However, in this instance the laboratory does not have enough data to establish a historical control range, and therefore the published range of 2-50 x $10^{-6}$ was used as a guide.

In the case of an equivocal result (*i.e.*, wherein the data does not meet all the criteria to support justification for a positive response) a repeat experiment may be performed using a modified concentration range to investigate the dose-response further.

## Section 7: Results and Discussion

### 7.1 Method Development

To provide a basic start-point for the experimental method, a standard assay protocol was sourced from a laboratory with a well-established, performant *hprt* assay. This provided a baseline experimental method upon which to build and optimise the GSK approach. In this Method Development section, the optimisation experiments conducted, and the learnings from each optimisation run that enabled development of the final protocol are described. These initial experiments were conducted using known genotoxins benzo[a]pyrene (BaP) and methyl methanesulfonate (MMS) which acted as positive controls. Using known genotoxins to establish the assay protocol provided critical data to demonstrate that the assay system was functioning correctly and was capable of reliably detecting genotoxic effects as intended. If the assay failed to respond to these established positive controls, it signalled an issue with the protocol itself or the optimisations that had been made. Ultimately, the goal was to arrive at an optimised protocol that was ready for assessment and demonstration of laboratory proficiency.

One of the first adaptations to the protocol was made before any testing began. The initial protocol for cell subculturing required cells to be passaged every two days to maintain exponential growth. This schedule inherently necessitated subculturing on Sundays, posing a logistical challenge. To address this issue, an adaptation was made with the aim of preventing overgrowth and ensuring cell cycle integrity over weekends. On Fridays, cells were seeded at a lower concentration of $0.08 \times 10^5$ cells/mL. This reduction minimised the risk of cell overgrowth during the two-day period, preventing cells from surpassing a density of $1 \times 10^6$ cells/mL, which would potentially disrupt their cell cycle and exponential growth-phase due to contact inhibition and culture medium nutrient depletion.

The first optimisation experiment (HPRT 1) was conducted using the baseline protocol with positive and negative controls as an initial trial to establish laboratory performance and

familiarisation with the methodology (Table 4, Figure 2). Dimethyl sulphoxide (DMSO) was selected as the negative control and benzo(a)pyrene (BaP) was selected as a known mutagenic positive control. In this experiment, five vehicle controls and two positive controls (at two concentrations 0.4 and 0.6 µg/mL) were selected, based on the information provided, to assess which concentration induced the highest mutant frequency without showing excess toxicity. These initial studies yielded five negative controls with a mean mutant frequency (MF) of $19.6x10^{-6}$ which was within the expected range (DeMarini et al., 1988), and positive controls which clearly showed an increase in MF to $73.8x10^{-6}$ at 0.4 µg/mL and $118.5x10^{-6}$ at 0.6 µg/mL (Figure 2). This established the basic assay at GSK and confirmed the ability to detect mutation at the *hprt* locus using the known genotoxicant, BaP.

**Table 4 – Survival and Mutant Frequency Data for DMSO and Benzo[a]pyrene (HPRT-1)**

| Compound | Concentration (µg/mL) | % Survival | Mutant Freq. ($x10^{-6}$) | Mean Mutant Freq. (per million cells) | Fold Increase |
|---|---|---|---|---|---|
| DMSO | 0 | 69.86 | 27.4 | 19.6 | N/A |
| DMSO | 0 | 90.23 | 6.18 | | |
| DMSO | 0 | 89.16 | 38.1 | | |
| DMSO | 0 | 98.92 | 24.86 | | |
| DMSO | 0 | 82.69 | 1.49 | | |
| BaP | 0.4 | 34.92 | 73.76 | 73.76 | 3.8 |
| BaP | 0.6 | 38.82 | 128.39 | 118.5 | 6.1 |
| BaP | 0.6 | 44.34 | 108.52 | | |

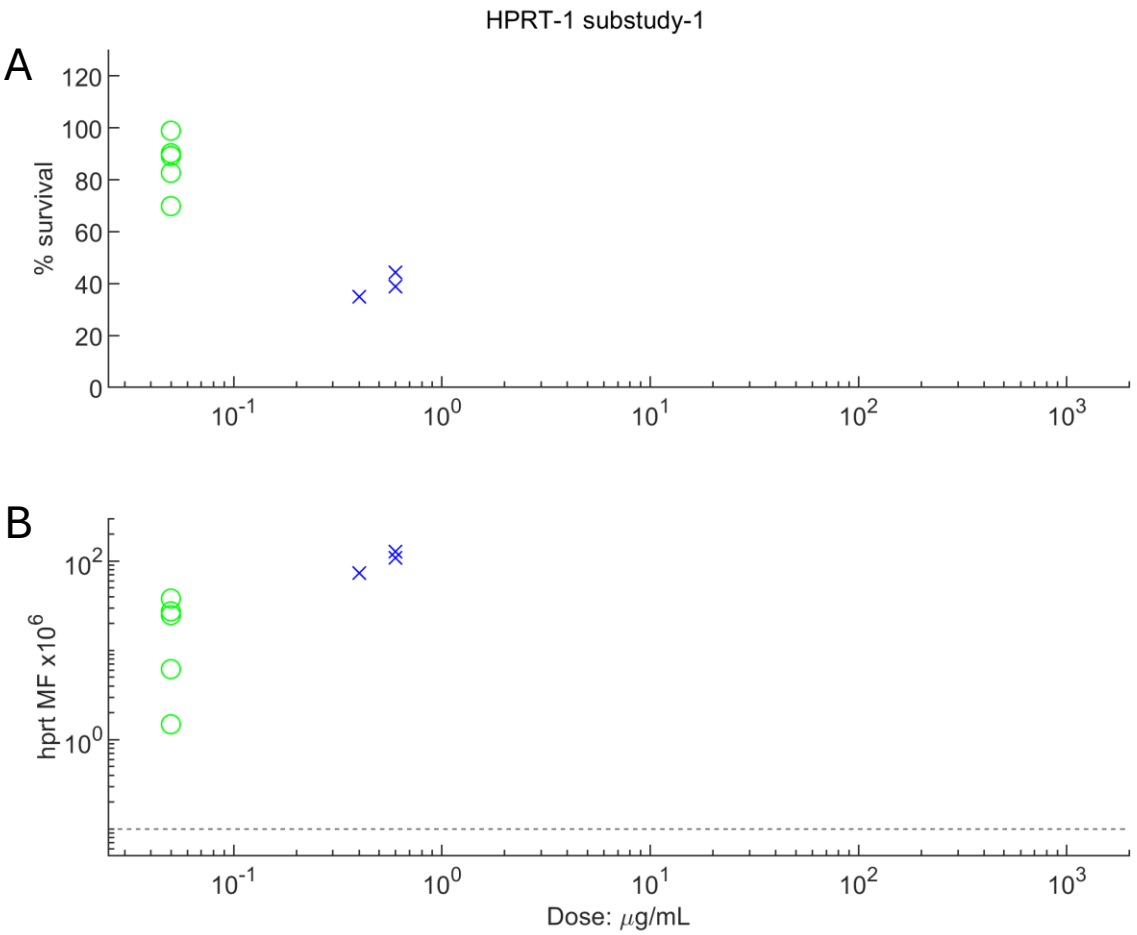



**Figure 2 – Survival and Mutant Frequency Dose-Response Relationships for DMSO and Benzo[a]pyrene (substudy "HPRT-1").** (A/B) Scatter plots on double-log scales showing (A) cell survival percentages (relative to vehicle control) and (B) *Hprt* mutant-frequency responses (MF $x10^{-6}$) for each test concentration. (A/B) Green circles depict replicate vehicle control data whilst blue crosses denote responses to the test compound.

In the HPRT 1 study, the lower seeding density used to carry the cells over the weekend was successful as the cells did not exceed the density limit of $10^7$ cells/mL, however a crucial observation was made regarding cell density at the start of the assay. This was initially set at $1x10^5$ cells/mL on Day 1 but this density was found to be too high as it resulted in cell counts on Day 3 that exceeded the maximum concentration of $1x10^6$ cells/mL in all cultures which meant that the cells had overgrown. Given this observation, the HPRT-2 study included an adjusted Day 1 cell concentration to prevent overgrowth which may have influenced mutation fixation, cell health and viability. The cell concentration was halved, yet Day 3 cell counts exceeded maximum recommended suggesting further decreases were necessary. In HPRT-2, the five negative controls had a mean MF of $12.0x10^{-6}$, and BaP (0.4 and 0.6 µg/mL) induced increases in MF to 126.4 and $130.7x10^{-6}$, respectively (Table 5, Figure 3). Although increases in MF were observed for both concentrations of BaP, excess toxicity was also observed (cell survivals were below 20%). In this study, one of the negative controls had a MF of zero (*i.e.,* indicating that no cell colonies formed in any of the wells of the mutation plates). Further examination using light microscopy showed a high volume of cell debris (Figure 4), confirming the plates were seeded with cells before test-article incubation and ruling out human error in the final dilution step.

**Table 5 – Survival and Mutant Frequency Data for DMSO and Benzo[a]pyrene (HPRT-2)**

| Compound | Concentration (µg/mL) | % Survival | Mutant Freq. ($x10^{-6}$) | Mean Mutant Freq. (per million cells) | Fold Increase |
|---|---|---|---|---|---|
| DMSO | 0 | 94.97 | 9.92 | 12.01 | N/A |
| DMSO | 0 | 123.89 | 0 | | |
| DMSO | 0 | 103.53 | 25.8 | | |
| DMSO | 0 | 106.17 | 7.05 | | |
| DMSO | 0 | 108.07 | 17.31 | | |
| BaP | 0.4 | 10.4 | 126.38 | 126.38 | 10.5 |
| BaP | 0.6 | 12.69 | 130.74 | 130.74 | 10.9 |

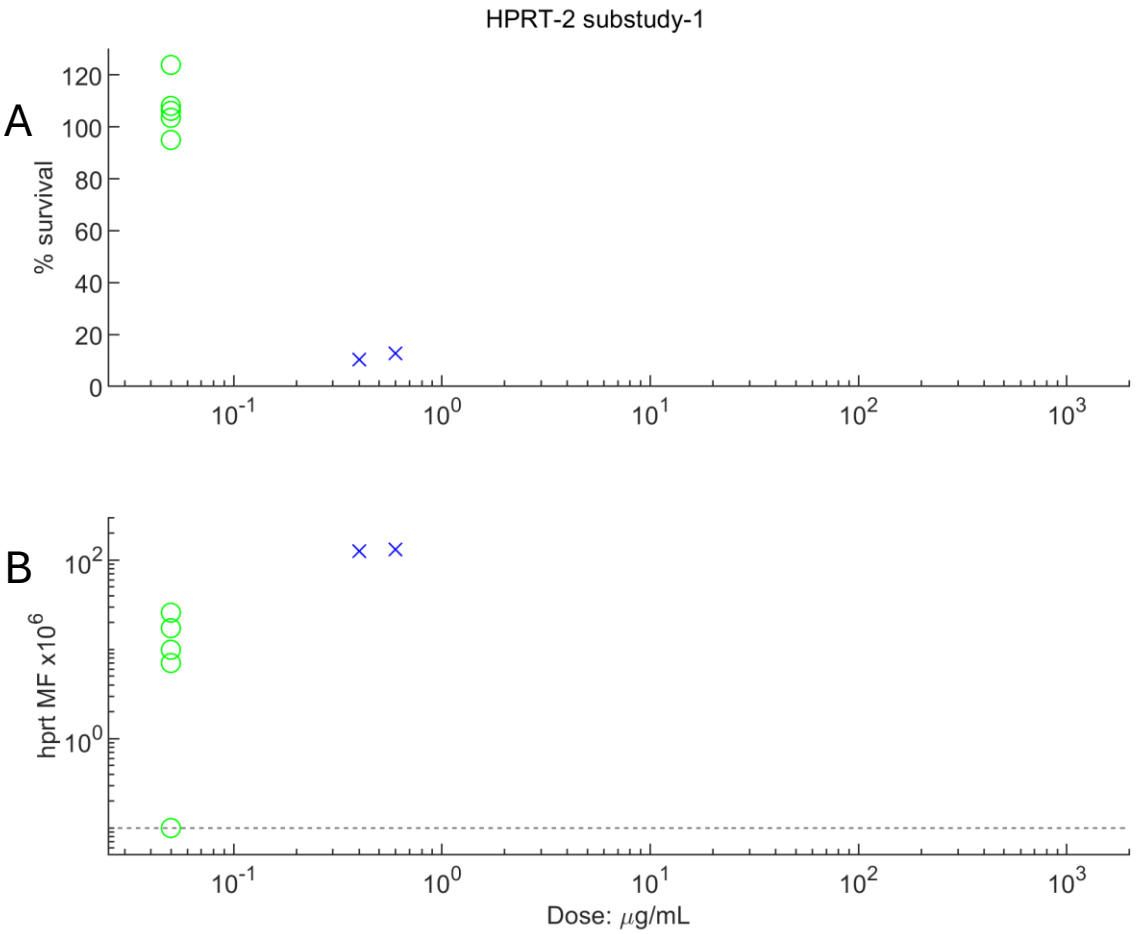



**Figure 3 – Survival and Mutant Frequency Dose-Response Relationships for DMSO and Benzo[a]pyrene (substudy "HPRT-2").** (A/B) Scatter plots on double-log scales showing (A) cell survival percentages (relative to vehicle control) and (B) *Hprt* mutant-frequency responses (MF $x10^{-6}$) for each test concentration. (A/B) Green circles depict replicate vehicle control data whilst blue crosses denote responses to the test compound. (B) Data-points placeholdered on the dashed horizontal line indicate zero mutant-frequency responses that fell below the limit-of-detection of the assay.

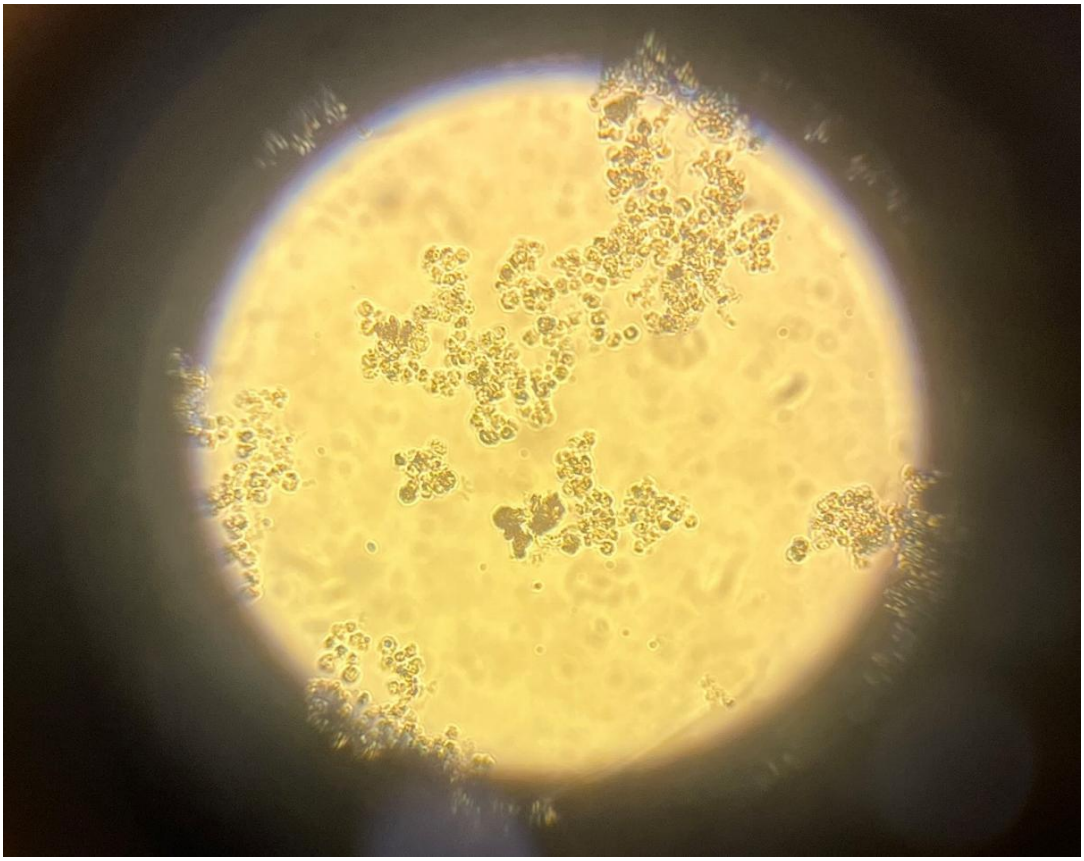


**Figure 4 – Light Microscopy Image of a Well in a 96-Well Plate Initially Seeded with Viable Cells.** The image shows intact and fragmented cellular material that confirms the prior presence of a viable cell population. The distribution and morphology of the debris suggest a process of cell degeneration, possibly due to apoptosis.

Due to the excessive BaP toxicity observed in HPRT-2, a wider concentration range of positive controls were assessed in HPRT-3 in an attempt to identify an optimal concentration in the presence of S9-mix. The concentration range assessed was 0.1-1 µg/mL (Table 6, Figure 5). The Day 1 cell seeding concentration was again halved to $0.25x10^5$ cells/mL yielding cell counts below $1x10^6$ on Day 3 demonstrating that, this time, the cells had successfully been maintained at densities compatible with exponential growth. For BaP cultures, cell survivals ranged from 90% at 0.1 µg/mL to 33% at 1 µg/mL, and increases in MF were observed at 0.2 µg/mL and above, with a 56.9-fold increase in MF over vehicle controls observed at 1 µg/mL (Table 6). The survival in this experiment had shifted slightly as in HPRT-2 cell survival at 0.6 µg/mL BaP was 13% compared to 42% in HPRT-3. This volatility was further investigated in following experiments. Again, two of the vehicle control replicates and also one of the 0.1 µg/mL BaP test concentrations yielded MFs of zero, which skewed the overall vehicle mean MF to $1.0x10^{-6}$, (notably lower than the published range ($2 – 50x10^{-6}$) (Moore, 1988)). Similar to HPRT-2, the plates with no growth were investigated using light microscopy and showed high levels of cell debris, indicating the presence of cell death.

**Table 6– Survival and Mutant Frequency Data for DMSO and Benzo[a]pyrene (HPRT-3)**

| Compound | Concentration (µg/mL) | % Survival | Mutant Freq. ($x10^{-6}$) | Mean Mutant Freq. (per million cells) | Fold Increase |
|---|---|---|---|---|---|
| DMSO | 0 | 80 | 0 | 1.03 | N/A |
| DMSO | 0 | 67.94 | 0 | | |
| DMSO | 0 | 84.22 | 3.08 | | |
| BaP | 0.1 | 89.64 | 0 | 0 | 0 |
| BaP | 0.2 | 75.44 | 35.75 | 35.75 | 34.8 |
| BaP | 0.3 | 65.05 | 3.34 | 3.34 | 3.3 |
| BaP | 0.4 | 59.65 | 6.64 | 6.64 | 6.5 |
| BaP | 0.5 | 55.42 | 19.63 | 19.63 | 19.1 |
| BaP | 0.6 | 42.34 | 78.09 | 78.09 | 76.1 |
| BaP | 1 | 32.9 | 58.38 | 58.38 | 56.9 |

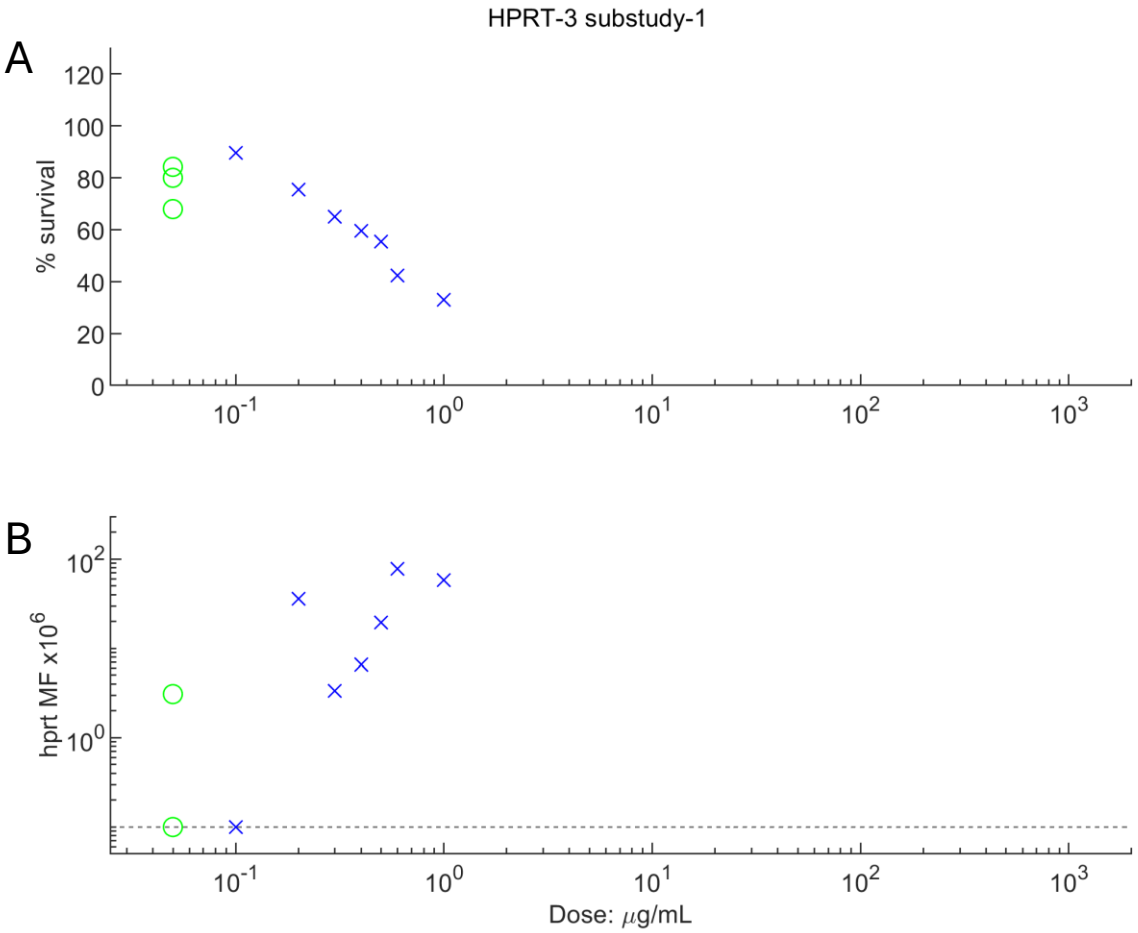



**Figure 5 – Survival and Mutant Frequency Dose-Response Relationships for DMSO and Benzo[a]pyrene (substudy "HPRT-3").** (A/B) Scatter plots on double-log scales showing (A) cell survival percentages (relative to vehicle control) and (B) *Hprt* mutant-frequency responses (MF $x10^{-6}$) for each test concentration. (A/B) Green circles depict replicate vehicle control data whilst blue crosses denote responses to the test compound. (B) Data-points placeholdered on the dashed horizontal line indicate zero mutant-frequency responses that fell below the limit-of-detection of the assay.

To further elucidate optimal dose concentrations for BaP, HPRT-4 was conducted using a similar BaP concentration range from 0.3 - 2 µg/mL (Table 7, Figure 6). At 2 µg/mL survival was reduced to 45% and resulted in a MF of $74.3x10^{-6}$, a 56.5-fold increase over concurrent controls. As a comparison, survival 0.6 µg/mL was 55%, indicating that the low survival seen in HPRT-2 may have been an outlier (possibly due to the high seeding density). Again, one vehicle control yielded a zero mutant frequency response. As this was a frequent occurrence across experiments, this issue was investigated further in the following optimisation experiments.

**Table 7 – Survival and Mutant Frequency Data for DMSO and Benzo[a]pyrene (HPRT-4)**

| Compound | Concentration (µg/mL) | % Survival | Mutant Freq. ($x10^{-6}$) | Mean Mutant Freq. (per million cells) | Fold Increase |
|---|---|---|---|---|---|
| DMSO | 0 | 82.18 | 1.67 | 1.32 | N/A |
| DMSO | 0 | 83.44 | 2.28 | | |
| DMSO | 0 | 76.8 | 0 | | |
| BaP | 0.3 | 54.21 | 0.79 | 0.79 | 0.6 |
| BaP | 0.4 | 69 | 5.17 | 5.17 | 3.9 |
| BaP | 0.5 | 58.96 | 11.98 | 11.98 | 9.1 |
| BaP | 0.6 | 54.89 | 14.6 | 14.6 | 11.1 |
| BaP | 1 | 36.59 | 75.44 | 75.44 | 57.3 |
| BaP | 1.5 | 41.41 | 27.93 | 27.93 | 21.2 |
| BaP | 2 | 44.56 | 74.31 | 74.31 | 56.5 |

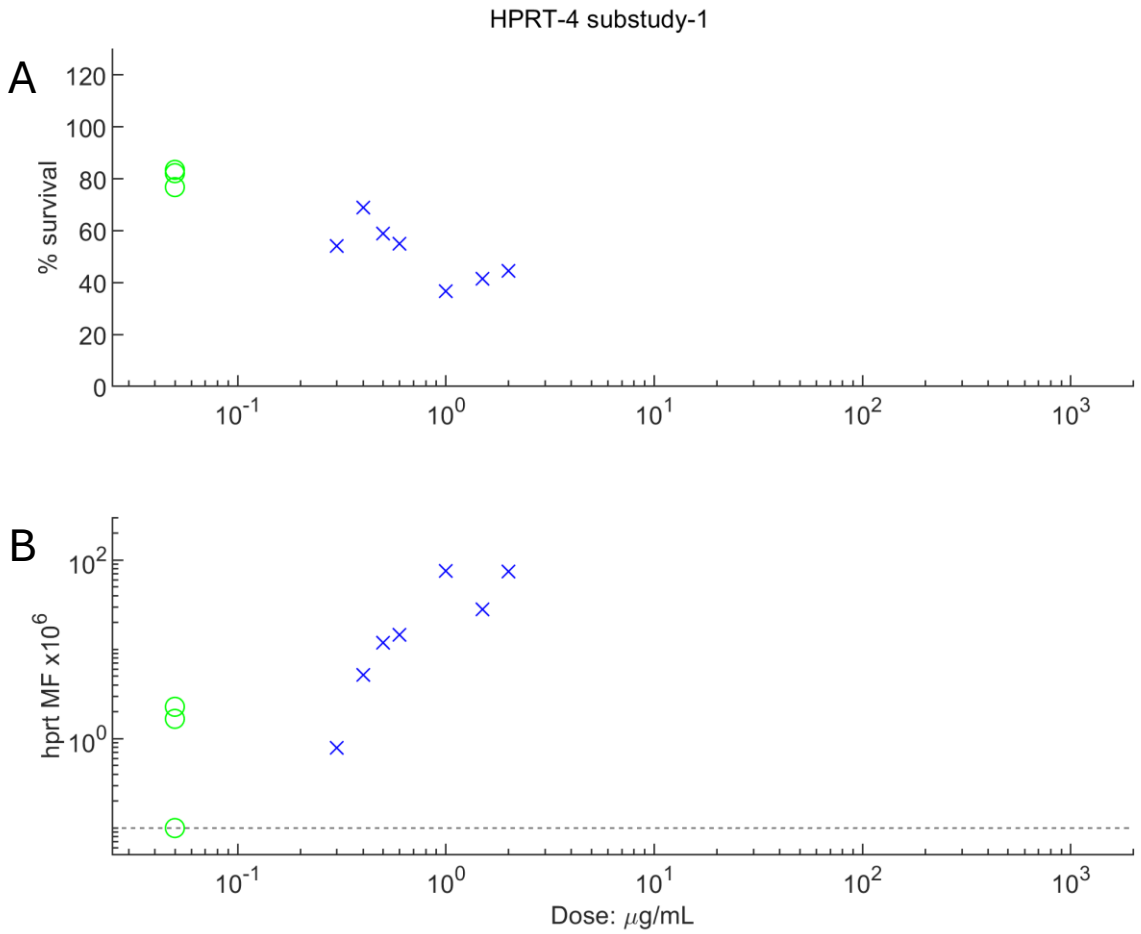



**Figure 6 – Survival and Mutant Frequency Dose-Response Relationships for DMSO and Benzo[a]pyrene (substudy "HPRT-4").** (A/B) Scatter plots on double-log scales showing (A) cell survival percentages (relative to vehicle control) and (B) *Hprt* mutant-frequency responses (MF $x10^{-6}$) for each test concentration. (A/B) Green circles depict replicate vehicle control data whilst blue crosses denote responses to the test compound. (B) Data-points placeholdered on the dashed horizontal line indicate zero mutant-frequency responses that fell below the limit-of-detection of the assay.

For HPRT-5, the primary focus was to address the issue of little/no cell growth in the mutation plates replicating the study design of HPRT-4 (using the same BaP concentrations *i.e.,* 0.3-2 µg/mL) but increasing the cell seeding density by 50% to $1.5x10^5$ cells/mL. The aim was to test the hypothesis that enhancing the initial density of cells plated for mutation would provide a larger starting population and mitigate the zero mutant frequency responses, thereby improving the detection limit for mutations, and overall assay sensitivity. Two of the cultures (one vehicle replicate and 0.4 µg/mL) were contaminated and consequently not scored for mutation. The results from this experiment (Table 8, Figure 7) again showed one vehicle control replicate with a MF of zero, indicating that the increased seeding density was not entirely effective suggesting the need for further optimisation experiments using even higher cell seeding densities.

**Table 8 – Survival and Mutant Frequency Data for DMSO and Benzo[a]pyrene (HPRT-5)**

| Compound | Concentration (µg/mL) | % Survival | Mutant Freq. ($x10^{-6}$) | Mean Mutant Freq. (per million cells) | Fold Increase |
|---|---|---|---|---|---|
| DMSO | 0 | 78.53 | 0 | 0.32 | N/A |
| DMSO | 0 | 93.3 | 0.64 | | |
| BaP | 0.3 | 68.65 | 11.53 | 11.53 | 35.8 |
| BaP | 0.5 | 59.48 | 52.92 | 52.92 | 164.4 |
| BaP | 0.6 | 57.29 | 21.16 | 21.16 | 65.7 |
| BaP | 1 | 32.27 | 54.13 | 54.13 | 168.2 |
| BaP | 1.5 | 26.49 | 9 | 9.00 | 28.0 |
| BaP | 2 | 26.76 | 58.25 | 58.25 | 180.9 |

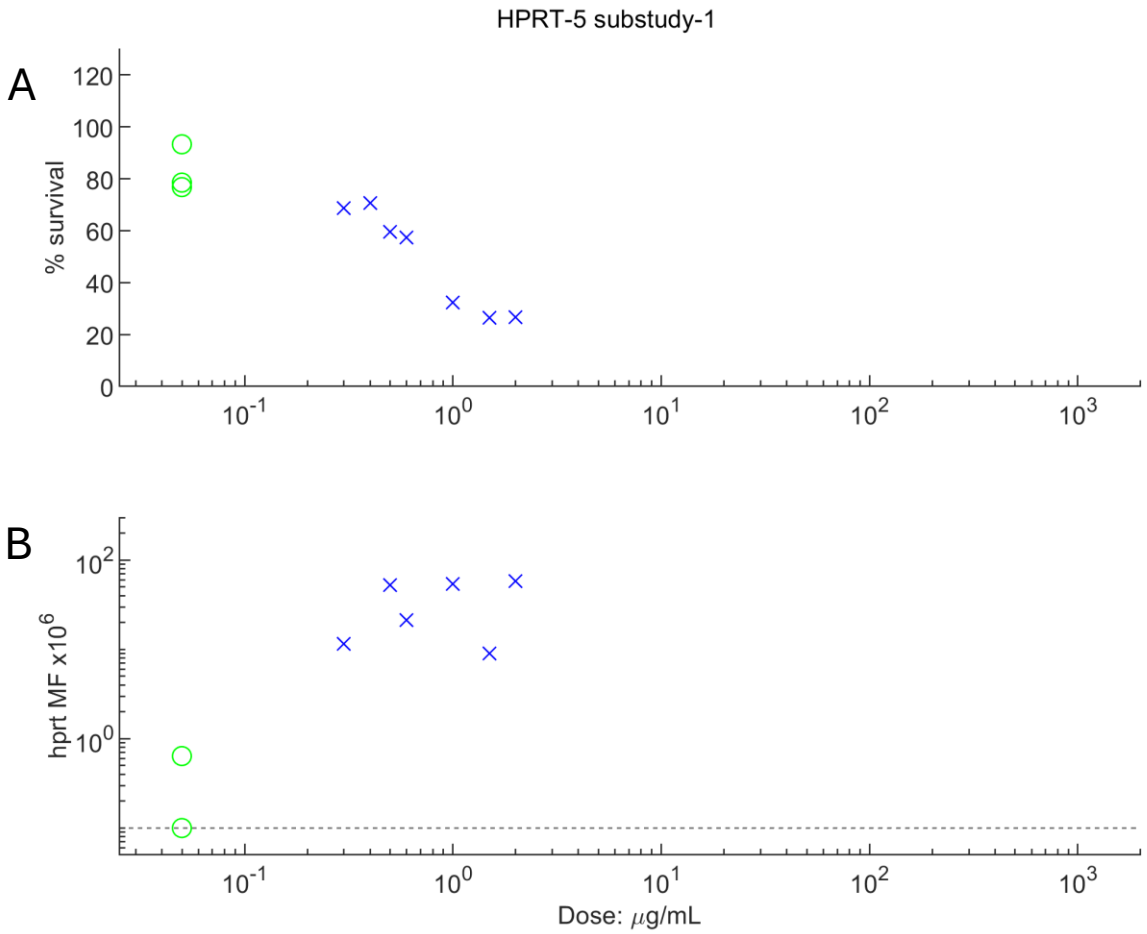



**Figure 7 – Survival and Mutant Frequency Dose-Response Relationships for DMSO and Benzo[a]pyrene (substudy "HPRT-5").** (A/B) Scatter plots on double-log scales showing (A) cell survival percentages (relative to vehicle control) and (B) *Hprt* mutant-frequency responses (MF $x10^{-6}$) for each test concentration. (A/B) Green circles depict replicate vehicle control data whilst blue crosses denote responses to the test compound. (B) Data-points placeholdered on the dashed horizontal line indicate zero mutant-frequency responses that fell below the limit-of-detection of the assay.

Based on the results observed in HPRT-5, the next experiment was designed with increased concentrations of BaP (0.25 – 20 µg/mL) to push toxicity limits, as well as increasing the seeding density to investigate whether this has any effect on MF (Figure 8, Tables 9 and 10). In this experiment, the MFs for cultures seeded at 1, 1.5 and $2x10^5$ were 17.5, 31.8 and $25.5x10^{-6}$, respectively. In practice, achieving a seeding density of 1.5 and $2x10^5$ cells/mL was more challenging because it required twice the volume of cells to reach the desired final concentration. As a result, significantly larger culture volumes needed to be maintained throughout the week to ensure that there were adequate cells available for plating day. Since there was little difference between the MFs of cultures plated at 1 versus $2x10^5$ cells/mL, and considering the practical limitations, it was decided that we would remain at the standard seeding density (*i.e.,* $1x10^5$). In addition to this, a further adaptation in HPRT-6 was the introduction of four vehicle control cultures and duplicate test concentration cultures to allow for appropriate pairwise statistical analysis in order to fulfil OECD guidelines. It was also thought that increasing numbers of both vehicle and test compound replicates would aid reliability of the assay by compensating for the potential of one of the replicates yielding a zero-response.

Additionally, due to the contamination observed in previous experimental plates, the anti-biotic concentration (Penicillin-Streptomycin) in plating culture medium (CM20) was increased. Upon literature review, it was apparent that increasing supplemental anti-biotic concentrations in cell cultures can lead to cytotoxic and cytostatic effects (Hassan SN, 2020) and therefore, this increase had to be carefully considered as a balance between combatting bacterial contamination without inducing cytotoxicity. This experiment showed that this increase (50%) this did not have an impact on the cell survival, with mean survival of vehicle controls yielding 85%, in-line with previous experiments.

**Table 9 – Survival and Mutant Frequency Data for DMSO and Benzo[a]pyrene (HPRT-6)**

| Compound | Concentration (µg/mL) | % Survival | Mutant Freq. ($x10^{-6}$) | Mean Mutant Freq. (per million cells) | Fold Increase |
|---|---|---|---|---|---|
| DMSO* | 0 | 81 | 17.5 | 11.2 | N/A |
| DMSO | 0 | 99.9 | 11.91 | | |
| DMSO | 0 | 77.09 | 15.52 | | |
| DMSO | 0 | 82.94 | 0 | | |
| BaP | 0.5 | 79.89 | 12.51 | 21.83 | 1.9 |
| BaP | 0.5 | 72.57 | 31.15 | | |
| BaP | 1 | 38.63 | 45.36 | 67.27 | 6.0 |
| BaP | 1 | 44.13 | 89.17 | | |
| BaP | 2.5 | 28.25 | 42.69 | 40.10 | 3.6 |
| BaP | 2.5 | 22.51 | 37.52 | | |
| BaP | 5 | 15.31 | 57.13 | 51.01 | 4.5 |
| BaP | 5 | 28.79 | 44.89 | | |
| BaP | 10 | 17.99 | 98.88 | 109.50 | 9.7 |
| BaP | 10 | 29.76 | 120.11 | | |

* Culture taken forward for seeding in mutation plates at additional varying cell densities (in table below)

**Table 10 – Mutant Frequency Data for DMSO Seeded at Differing Cell Densities**

| Compound | Seeding Density ($x10^5$) | Mutant Frequency ($x10^{-6}$) |
|---|---|---|
| DMSO | 1 | 17.5 |
| DMSO | 1.5 | 31.78 |
| DMSO | 2 | 26.48 |

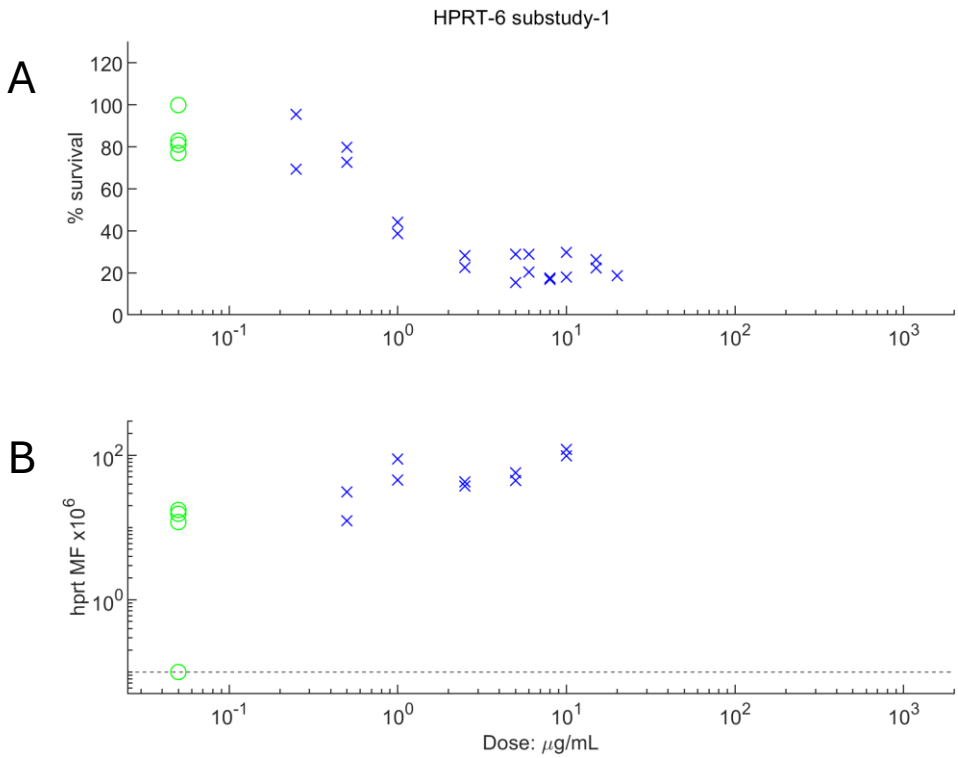



**Figure 8 – Survival and Mutant Frequency Dose-Response Relationships for DMSO and Benzo[a]pyrene (substudy "HPRT-6").** (A/B) Scatter plots on double-log scales showing (A) cell survival percentages (relative to vehicle control) and (B) *Hprt* mutant-frequency responses (MF $x10^{-6}$) for each test concentration. (A/B) Green circles depict replicate vehicle control data whilst blue crosses denote responses to the test compound. (B) Data-points placeholdered on the dashed horizontal line indicate zero mutant-frequency responses that fell below the limit-of-detection of the assay.

In order to further investigate the zero MF-responses, a batch of cells (batch 2) were obtained from a laboratory with a well-established *hprt* assay that did not regularly observe zero responses to test the idea that the cell stock might have been the variable driving these outcomes. This experiment (HPRT-9) was a reproduction of HPRT-8 and is shown in Table 11, Figure 9 below, to assess the differences observed in different cell batch MFs when treated with methyl methanesulphonate (MMS). The concentration range assessed was 10 – 30 µg/mL in triplicate cultures. As can be seen, the mean MF was higher than observed in previous experiments with GSK cells, indicating that the new batch has a higher background spontaneous MF, however a zero response was still observed in one of the vehicle controls.

**Table 11 – Survival and Mutant Frequency Data for DMSO and Methyl Methanesulfonate with Batch 2 L5178Y Cells (HPRT-9)**

| Compound | Concentration (µg/mL) | % Survival | Mutant Frequency ($x10^{-6}$) | Mean Mutant Frequency (per million cells) | Fold Increase |
|---|---|---|---|---|---|
| DMSO | 0 | 126.77 | 2.84 | 20.7 | NA |
| DMSO | 0 | 118.7 | 9.13 | | |
| DMSO | 0 | 108.28 | 13.09 | | |
| DMSO | 0 | 91.87 | 0 | | |
| DMSO | 0 | 109.79 | 78.31 | | |
| MMS | 10 | 119.5 | 57.6 | 33.7 | 1.6 |
| MMS | 10 | 94.87 | 14.22 | | |
| MMS | 10 | 69.95 | 29.34 | | |
| MMS | 20 | 86.07 | 65.86 | 39.5 | 1.9 |
| MMS | 20 | 91.84 | 31.39 | | |
| MMS | 20 | 76.04 | 21.37 | | |
| MMS | 25 | 67.59 | 7.84 | 27.6 | 1.3 |
| MMS | 25 | 51.75 | 35.69 | | |
| MMS | 25 | 72.37 | 39.39 | | |
| MMS | 30 | 47.35 | 97.2 | 72.5 | 3.5 |
| MMS | 30 | 64.84 | 60.07 | | |
| MMS | 30 | 49.1 | 60.19 | | |

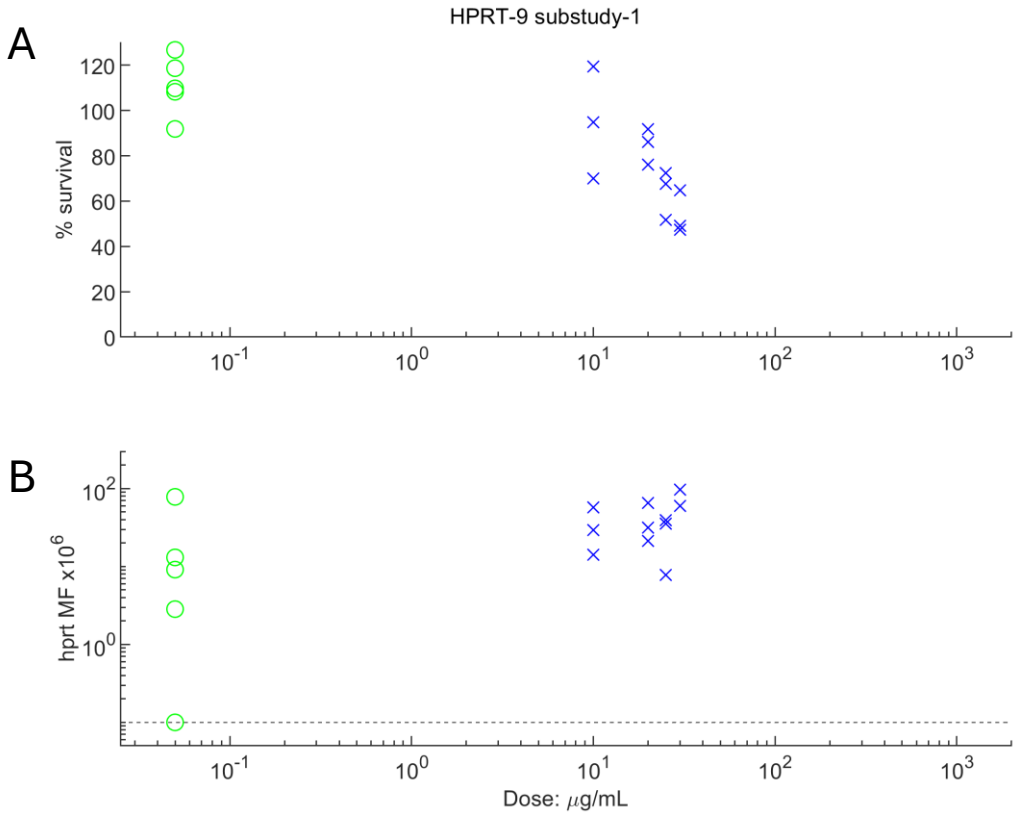



**Figure 9 – Survival and Mutant Frequency Dose-Response Relationships for DMSO and methyl methanesulfonate Using Batch 2 of L5178Y Cells (substudy "HPRT-9").** (A/B) Scatter plots on double-log scales showing (A) cell survival percentages (relative to vehicle control) and (B) *Hprt* mutant-frequency responses (MF $x10^{-6}$) for each test concentration. (A/B) Green circles depict replicate vehicle control data whilst blue crosses denote responses to the test compound. (B) Data-points placeholdered on the dashed horizontal line indicate zero mutant-frequency responses that fell below the limit-of-detection of the assay.

An analysis of both sets of MMS data using different cell batches was conducted to evaluate which was most suitable for use moving forward (Figure 10). Plotting the data shows that whereas the responses were similar, batch 2 yielded a slightly higher mean MF in the vehicle control and two lowest dose groups. However, these differences were marginal, and both batches of cells produced zero responses in one of the vehicle control replicates. To understand the similarities and differences across the two datasets, benchmark dose (BMD) modelling was applied. Consistently across all four benchmark dose models it was found that the data from both batches of cells were near-identical, leading to rejection of cell batch as an influential covariate and description of all data from both cell batches using a single curve. (Figure 11). Based on this, the decision was made to continue using the original GSK batch of cells for future tests due to the mouse lymphoma tests being conducted in the same batch and therefore allowing for a more direct comparison between assays. The modelling also highlighted that high variability in the vehicle controls (indicated on left-side of each graph in Figure 11) created considerable uncertainty in determining the true shape of the dose response curve representing a barrier to quantitation from the available dose-response information.

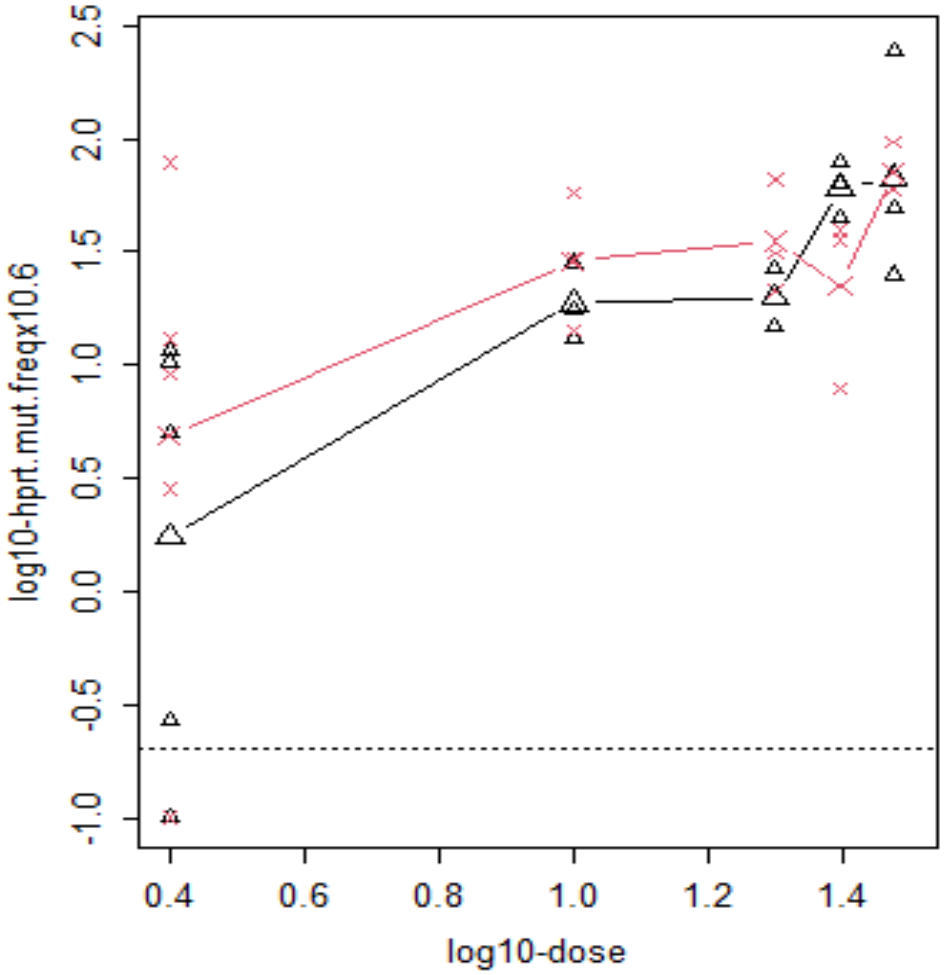



**Figure 10 – Comparison of *hprt* MF responses to MMS across two different batches of cells ("GSK L5178Y" or "Batch 2 L5178Y").** The dose-response from substudy "HPRT-8" using GSK's original batch of L5178Y cells are shown in black whereas red shows the results from "HPRT-9" using Batch 2 L5178Y cells. Dose and response values are log10 transformed. Individual-replicate responses are indicated by the small symbols whereas the mean response at each test concentration is represented by the large symbols. Data-points placeholdered below the dashed horizontal line indicate zero mutant-frequency responses that fell below the limit-of-detection of the assay.

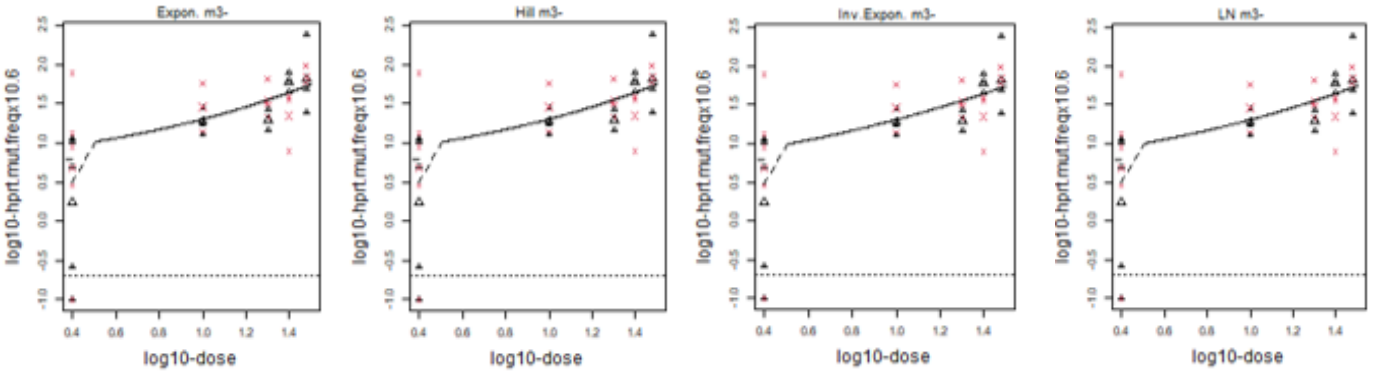



**Figure 11 – BMD model fits to the *hprt* MF data for MMS across two different batches of cells ("GSK L5178Y" or "Batch 2 L5178Y").** The dose-response from substudy "HPRT-8" using GSK's original batch of L5178Y cells are shown in black whereas red shows the results from "HPRT-9" using Batch 2 L5178Y cells. Dose and response values are log10 transformed. Individual replicate responses are indicated with symbols according to test. Data-points placeholdered below the dashed horizontal line indicate zero mutant-frequency responses that fell below the limit-of-detection of the assay. All four BMD models (from left-to-right, exponential, Hill, inverse exponential and log normal) reject cell batch as an influential covariate and show that the data collected across both cell batches are adequately described using a single model fit.

In summary, the optimisation of the *hprt* assay protocol led to improvements in practical elements such as cell subculturing as well as assay reliability. The introduction of triplicate test cultures provided necessary data for performing pairwise statistical analysis, fulfilling OECD guideline requirements and improving assay reliability, as well as increasing sensitivity beyond the standard single-culture protocol. Based on the results from multiple studies with BaP, it was shown that concentrations around 1 – 2 µg/mL were optimal for use as a positive control, inducing slight toxicity whilst also providing a clear and robust increase in MF. The optimisation phase also highlighted key challenges within the assay, such as persistent zero MF responses across cultures, which prompted investigation into cell batches, resulting in rejection of batch variability as a covariate influencing assay outcome. Overall, this led to an improved assay protocol which was ready to be used to demonstrate laboratory proficiency, albeit still with some areas where further development was desirable.

### 7.2 Demonstrating Lab Proficiency with Known Genotoxicants

In-line with OECD recommendations, demonstrated laboratory proficiency is required to confirm sufficient assay experience prior to use for testing. This can be achieved using reference genotoxicants acting via different mechanisms, preferably at least one with and one without metabolic activation (S9-mix) (OECD, 2016). In previous optimisation experiments, the laboratory had repeatedly demonstrated the ability to detect mutations when dosing with BaP in the presence of S9-mix. Therefore, in this section, MMS was used to demonstrate proficiency in detecting mutagens in the absence of S9-mix.

Going into the laboratory proficiency phase, a new dose range-finder design was adopted as a trial for test compounds. Based on the results seen in the optimisation phase, it was clear that increased replicates were required in order to avoid too many zero-responses thereby increasing the sensitivity of the assay. Therefore, it was decided that 6 vehicle control cultures and 3 replicates per dose group would be used. Because this design was more resource

intensive, an initial dose-range finder (DRF) experiment was conducted prior to the main test in single cultures to investigate the toxicity profile of the compound, which in turn tailored the design of the main test and influenced which concentrations were selected for testing (*e.g.,* four dose-plus-control designs). This meant that the main test consisted of a narrower targeted concentration range with more replicates per dose-group – which importantly was still achievable with the same level of laboratory resourcing (*e.g.,* time and support personnel).

MMS is a known genotoxicant that exerts its mutagenic effect primarily through DNA methylation. The majority of methylation occurs at the N7 position of guanine (approximately 80-85%), although this can also occur at other positions such as N1-adenine, N3-cytosine or O6-guanine (Wyatt and Pittman, 2006). This alters the hydrogen bonding of these bases and therefore results in base-pairing mismatches whereby wrong nucleotides are substituted into the sequence, leading to point mutations if not repaired. Additional DNA damage can also be caused as a result of an attempt to correct these mutations by base excision repair, which usually results in double-strand DNA breaks and error-prone repair which can also lead to cell death.

The DRF investigated survival of a dose range from 1 – 40 µg/mL. At 40 µg/mL, survival was reduced to 9%, showing that the MMS had caused extensive DNA damage leading to apoptosis. Using this information, a dose-range of 10 – 40 µg/mL was taken forward for the main test.

In the main test the toxicity profile shifted slightly which resulted in both 35 and 40 µg/mL dose groups exhibiting excessive toxicity at <10% survival. Due to this, these dose groups were not taken forward for mutagenicity assessment. As shown in Table 12, Figure 12, at 30 µg/mL the mutation frequency was increased to 106 $x10^{-6}$, however because mean survival was reduced to 10% this was classified as a cytotoxic positive response. At lower concentrations mutant frequency was also increased to a maximum mean of 62.6 $x10^{-6}$ at 25 µg/mL. Because these

were above the threshold of toxicity, they were concluded as biologically relevant positive responses.

This experiment therefore concluded the assay was capable of detecting mutations in the absence of S9-mix when treated with MMS, as well as informing on potential concentrations for an MMS positive control if required when deployed on future assets in the absence of S9-mix.

**Table 12 – Survival and Mutant Frequency Data for DMSO and Methyl Methanesulfonate (HPRT-8)**

| Compound | Concentration (µg/mL) | % Survival | Mutant Frequency ($x10^{-6}$) | Mean Mutant Frequency (per million cells) | Fold Increase |
|---|---|---|---|---|---|
| DMSO | 0 | 114.66 | 0 | 5.4 | NA |
| DMSO | 0 | 79.48 | 10.38 | | |
| DMSO | 0 | 78.44 | 0.27 | | |
| DMSO | 0 | 81.59 | 4.97 | | |
| DMSO | 0 | 71.75 | 11.51 | | |
| MMS | 10 | 71.8 | 17.59 | 19.7 | 3.6 |
| MMS | 10 | 69.71 | 13.14 | | |
| MMS | 10 | 62.53 | 28.47 | | |
| MMS | 20 | 34.31 | 26.84 | 20.8 | 3.8 |
| MMS | 20 | 38.76 | 14.73 | | |
| MMS | 25 | 15.99 | 45.05 | 62.6 | 11.5 |
| MMS | 25 | 23.09 | 79.35 | | |
| MMS | 25 | 18.93 | 63.29 | | |
| MMS | 30 | 10.08 | 243.52 | 106.0 | 19.5 |
| MMS | 30 | 9.03 | 25.08 | | |
| MMS | 30 | 9.33 | 49.26 | | |

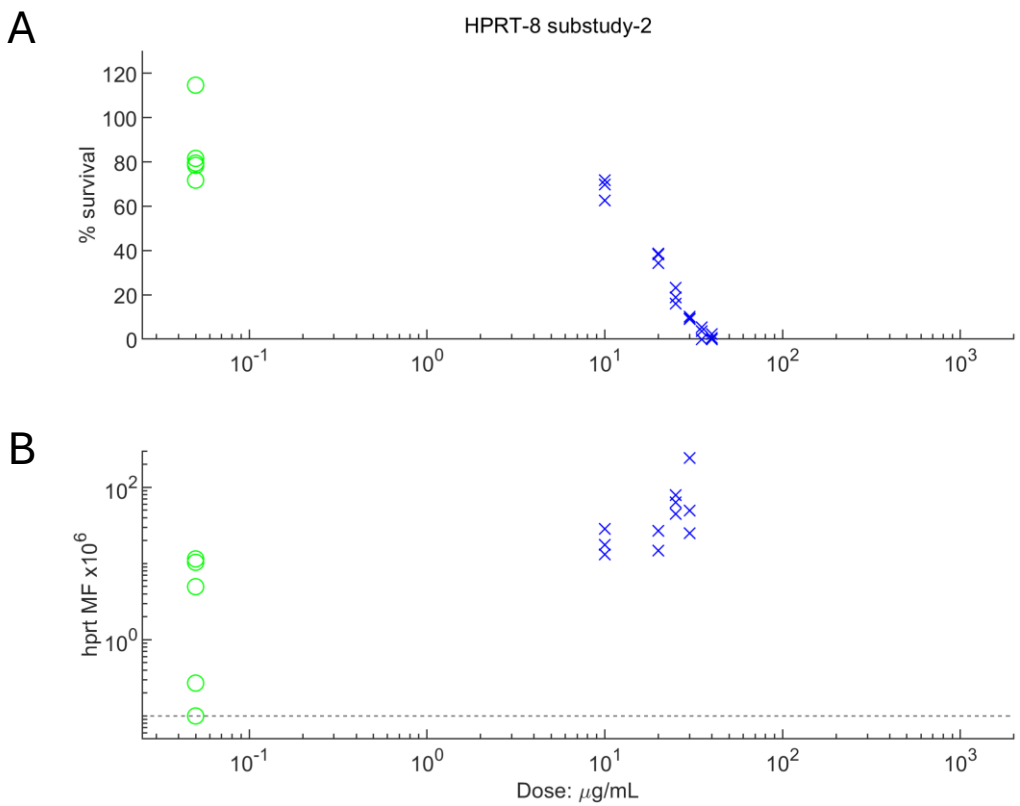



**Figure 12 – Survival and Mutant Frequency Dose-Response Relationships for DMSO and methyl methanesulphonate (substudy "HPRT-8").** (A/B) Scatter plots on double-log scales showing (A) cell survival percentages (relative to vehicle control) and (B) *Hprt* mutant-frequency responses (MF $x10^{-6}$) for each test concentration. (A/B) Green circles depict replicate vehicle control data whilst blue crosses denote responses to the test compound. (B) Data-points placeholdered on the dashed horizontal line indicate zero mutant-frequency responses that fell below the limit-of-detection of the assay.

### 7.3 Tool Nitrosamines

**N-Nitrosodimethylamine (NDMA): HPRT-10**

NDMA is a short alkyl-chain nitrosamine that was identified as an impurity product in several marketed drugs resulting in various product recalls. It is classified by the International Agency for Research on Cancer (IARC) as group 2A: probably carcinogenic to humans. It is a known rodent carcinogen (Toth et al., 1964), with proven *in vitro* mutagenic activity in the Ames test (Thomas et al., 2024).

Results are shown in Table 13 and Figure 13. Prior to the main test, a DRF experiment was conducted to assess toxicity which may have limited the top concentration. As no excess toxicity (*i.e.,* < 20% survival) was observed up to 10 mM, 740.8 µg/mL was used as the top concentration, which was the maximum test concentration in alignment with current OECD guidelines. In the main test, three lower doses of 100, 250 and 500 µg/mL were also selected for mutagenicity assessment to gain insight of activity over a wide concentration range. Six vehicle control samples were treated, and each dose group was evaluated in triplicate, as per the study design outlined in the validation section. Positive control treatments included two concentrations of BaP (0.5 and 1 µg/mL). Cell survival decreased in a dose-dependent manner across all dose groups and at the top dose mean survival was reduced to 31.8%. Increases in mutant frequency were observed in all dose groups, and a statistically significant response (15-fold over concurrent control) was identified at 500 µg/mL only. Notably, there was wide variation in the vehicle control group, with one dose producing a zero response, as well as one replicate in the 250 µg/mL group. It is likely that variation within dose groups resulted in the top dose being deemed non-significant, due to one replicate yielding a lower MF of $4.95 \times 10^{-6}$. Pairwise statistical tests (*e.g.,* Dunnett's Test) rely on comparing the difference between the mean of each dose-group relative to the mean of a single control-group. In this test, one of the three replicates is in-line with control-group responses, so therefore statistical testing suggests there

is <95% confidence of a significant response and thus the dose-group does not constitute a 'positive' overall. The positive control at 1 µg/mL showed a clear increase in mutant frequency, however as it was completed as a single replicate statistical significance could not be assessed.

Overall, the positive *hprt* result obtained aligned with published mutagenicity and carcinogenicity data and showed that the assay was capable of detecting the mutagenicity of NDMA.

**Table 13 – Survival and Mutant Frequency Data for NDMA (HPRT-10)**

| Compound | Concentration (µg/mL) | % Survival | Mutant Freq. (x10$^{-6}$) | Mean Mutant Freq. (per million cells) | Fold Increase | P-Value |
|---|---|---|---|---|---|---|
| Water | 0 | 102.27 | 3.50 | 4.6 | NA | NA |
| Water | 0 | 103.78 | 0.16 | | | |
| Water | 0 | 100.48 | 0.00 | | | |
| Water | 0 | 78.61 | 9.71 | | | |
| Water | 0 | 110.53 | 7.34 | | | |
| Water | 0 | 86.67 | 7.02 | | | |
| NDMA | 100 | 88.99 | 44.06 | 18.2 | 3.9 | 0.3 |
| NDMA | 100 | 117.21 | 7.77 | | | |
| NDMA | 100 | 89.16 | 2.84 | | | |
| NDMA | 250 | 78.58 | 9.69 | 18 | 3.9 | 0.4 |
| NDMA | 250 | 71.14 | 0.00 | | | |
| NDMA | 250 | 67.91 | 44.35 | | | |
| NDMA | 500 | 49.27 | 77.69 | 69.2 | 15 | 0.002 |
| NDMA | 500 | 55.95 | 89.13 | | | |
| NDMA | 500 | 65.17 | 40.79 | | | |
| NDMA | 740.8 | 37.16 | 28.54 | 31.4 | 6.8 | 0.09 |
| NDMA | 740.8 | 26.41 | 60.64 | | | |
| NDMA | 740.8 | 31.83 | 4.95 | | | |
| BaP | 0.5 | 102.85 | 5.02 | 5.0 | 1.1 | NA |
| BaP | 1 | 35.26 | 55.57 | 55.6 | 12 | NA |

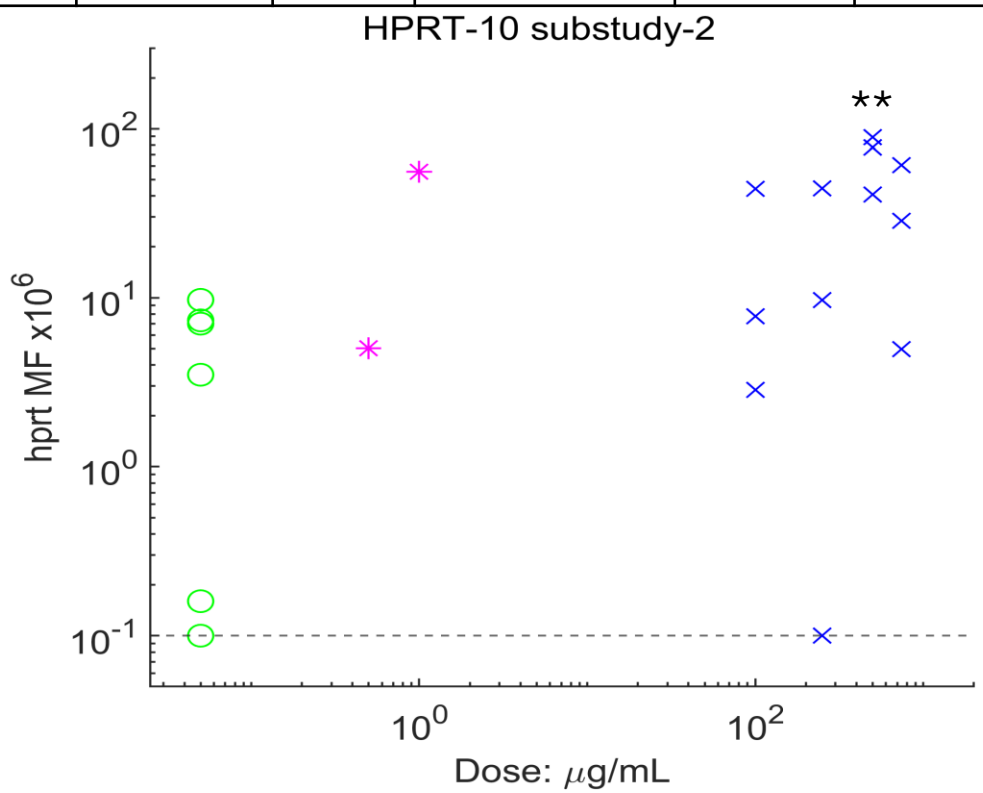



**Figure 13 –Mutant Frequency Dose-Response Relationships for Water and NDMA (substudy "HPRT-10").** *Hprt* mutant-frequency responses (MF x10$^{-6}$) for each test concentration. (A/B) Green circles depict replicate vehicle control data whilst blue crosses denote responses to the test compound and magenta stars indicate positive control responses. Data-points placeholdered on the dashed horizontal line indicate zero mutant-frequency responses that fell below the limit-of-detection of the assay. Pairwise testing was carried out according to the decision-tree described by Johnson et al. [2014] with response significance determined by the one-sided Dunnett's test indicated by the asterisks ( where: *p<0.05, **p<0.01, ***p<0.001).

**N-Nitrosodiethylamine (NDEA): HPRT-11**

NDEA is another example of a short alkyl chain nitrosamine, also categorised by IARC as group 2A. Again, it is a known rodent carcinogen (Gray et al., 1991) with proven in vitro mutagenic activity in the Ames test (Thomas et al., 2024).

*Hprt* assay results are shown in Table 14 and Figure 14. Similarly to NDMA, a DRF experiment was conducted prior to the main test to assess toxicity. As limited toxicity was observed, 1012 µg/mL (10 mM) was used as the top dose, with lower doses of 400, 600 and 800 µg/mL also assessed for mutagenicity. Based on results from the prior experiment (HPRT-10), BaP was treated at 1 µg/mL in triplicate to allow for full statistical analysis of the positive control. NDEA exhibited low toxicity relative to the vehicle controls, with the top dose survival decreasing slightly to 67.6%. Large increases in mutant frequency were observed in all dose groups, with statistically significant responses yielded at 600, 800 and 1012 µg/mL (maximum 19-fold increase over concurrent control). Interestingly, in-group variability was lower in this experiment, yet the genotoxic effect was more potent. Furthermore, none of the treatment groups showed zero responses in MF, supporting the hypothesis that the previous zero responses were likely below the assay's detection limit rather than truly zero. This indicated that heightened chemical potency required less assay sensitivity in order to detect an effect. The BaP positive control did not quite reach the threshold for a significant response ($p = 0.09$) – however it showed clear increases in MF response across all three replicates.

The positive *hprt* result for NDEA therefore aligned with published mutagenicity and carcinogenicity data.

**Table 14 – Survival and Mutant Frequency Data for NDEA (HPRT-11)**

| Compound | Concentration (µg/mL) | % Survival | Mutant Freq. (x10$^{-6}$) | Mean Mutant Freq. (per million cells) | Fold Increase | P-Value |
|---|---|---|---|---|---|---|
| Water | 0 | 88.96 | 14.51 | 8.9 | NA | NA |
| Water | 0 | 97.87 | 0.41 | | | |
| Water | 0 | 95.63 | 15.61 | | | |
| Water | 0 | 89.12 | 0.00 | | | |
| Water | 0 | 76.21 | 17.17 | | | |
| Water | 0 | 79.89 | 5.80 | | | |
| NDEA | 400 | 76.17 | 41.64 | 36.2 | 4.1 | 0.1 |
| NDEA | 400 | 72.26 | 45.34 | | | |
| NDEA | 400 | 82.97 | 21.48 | | | |
| NDEA | 600 | 53.76 | 120.17 | 82.6 | 9.3 | 0.004 |
| NDEA | 600 | 74.26 | 82.94 | | | |
| NDEA | 600 | 73.37 | 44.59 | | | |
| NDEA | 800 | 65.54 | 115.36 | 148.3 | 16.6 | 0.0001 |
| NDEA | 800 | 61.63 | 161.87 | | | |
| NDEA | 800 | 75.58 | 167.60 | | | |
| NDEA | 1012 | 70.21 | 188.15 | 172.7 | 19.4 | 0.000009 |
| NDEA | 1012 | 65.79 | 44.01 | | | |
| NDEA | 1012 | 66.92 | 286.04 | | | |
| BaP | 1 | 47.64 | 29.81 | 39.8 | 4.5 | 0.09 |
| BaP | 1 | 49.10 | 38.81 | | | |
| BaP | 1 | 38.90 | 50.63 | | | |

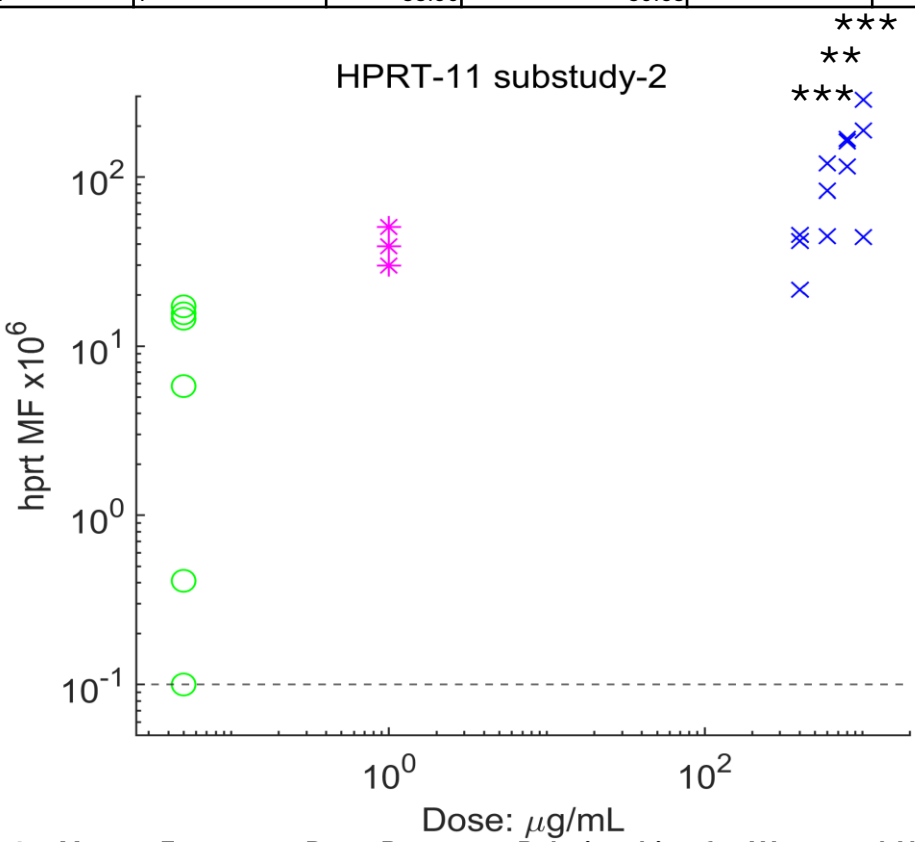



**Figure 14 – Mutant Frequency Dose-Response Relationships for Water and NDEA (substudy "HPRT-11").** *Hprt* mutant-frequency responses (MF x10$^{-6}$) for each test concentration. (A/B) Green circles depict replicate vehicle control data whilst blue crosses denote responses to the test compound and magenta stars indicate positive control responses. Data-points placeholdered on the dashed horizontal line indicate zero mutant-frequency responses that fell below the limit-of-detection of the assay. Pairwise testing was carried out according to the decision-tree described by Johnson et al. [2014] with response significance determined by the one-sided Dunnett's test indicated by the asterisks ( where: *p<0.05, **p<0.01, ***p<0.001).

**N-Nitrosodiisopropylamine (NDIPA): HPRT-12**

NDIPA is a further example of a nitrosamine which is known to have weak mutagenic activity in the Ames test (Thomas et al., 2024) as well as being recognised as a rodent carcinogen (Lijinsky, 1987).

Results are shown in Table 15 and Figure 15. A DRF was conducted prior to the main test to assess toxicity. As little toxicity was observed, 1301.9 µg/mL (10mM) was used as the top dose, with lower doses of 500, 700 and 900 µg/mL also assessed for mutagenicity to explore a dose response. Based on the positive control results from HPRT-11, the BaP concentration was increased to 1.5 µg/mL in order to achieve a robust positive response which should be statistically significant. NDIPA showed no toxicity, with replicates at the top dose resulting in similar survival rates to the vehicle controls. One replicate at 900 µg/mL had a notably lower survival (52%), however this was not reflected across the whole dose group, and the overall mean of the group was still at control level. There were no increases in mutant frequency at any of test concentrations assessed. Again, high in-group variability was observed: for example, replicates for the top dose spanned a MF range of 0.8 – 14 $x10^{-6}$. The positive control yielded a statistically positive response ($p = 0.006$), indicating a valid test.

NDIPA is considered a weak mutagen and rodent carcinogen, with legacy Ames-test data often being conflicting (Thomas et al., 2024). Thomas et al., (2025) reported that the enhanced Ames test was able to detect NDIPA as a weak mutagen when using water as a solvent and only in the presence of hamster S9-mix. This is thought to be because of the association of steric hinderance with its structure. This occurs because the branched alkyl side chain can partially obstruct α-hydroxylation by cytochrome P450 enzymes, preventing the formation of a reactive derivative (Cross and Ponting, 2021). It is possible that differences in species S9 could have an effect on the metabolism of NDIPA resulting in differences between rat and hamster tests. As this HPRT assay was only conducted in the presence of rat S9-mix, it is unsurprising that it was

not able to detect increases in MF and therefore correlates with the **rat S9** result seen in the Ames test. Correlation with outcomes from the MLA will be discussed in a later section.

**Table 15 – Survival and Mutant Frequency Data for NDIPA (HPRT-12)**

| Compound | Concentration (µg/mL) | % Survival | Mutant Freq. ($x10^{-6}$) | Mean Mutant Freq. (per million cells) | Fold Increase | P Value |
|---|---|---|---|---|---|---|
| Water | 0 | 99.78 | 0.00 | 4.3 | NA | NA |
| Water | 0 | 87.79 | 18.05 | | | |
| Water | 0 | 87.51 | 0.16 | | | |
| Water | 0 | 85.02 | 0.00 | | | |
| Water | 0 | 92.52 | 0.31 | | | |
| Water | 0 | 84.31 | 7.47 | | | |
| NDIPA | 500 | 80.26 | 8.42 | 6.1 | 1.4 | 0.3 |
| NDIPA | 500 | 85.70 | 0.75 | | | |
| NDIPA | 500 | 80.99 | 9.04 | | | |
| NDIPA | 700 | 98.97 | 3.84 | 7.8 | 1.8 | 0.1 |
| NDIPA | 700 | 73.89 | 11.33 | | | |
| NDIPA | 700 | 79.06 | 8.36 | | | |
| NDIPA | 900 | 106.84 | 4.69 | 2.6 | 0.6 | 0.7 |
| NDIPA | 900 | 89.92 | 3.03 | | | |
| NDIPA | 900 | 51.45 | 0.00 | | | |
| NDIPA | 1301.9 | 80.41 | 13.98 | 5.9 | 1.2 | 0.5 |
| NDIPA | 1301.9 | 86.02 | 1.03 | | | |
| NDIPA | 1301.9 | 93.53 | 0.78 | | | |
| BaP | 1.5 | 50.10 | 47.42 | 56.9 | 13.1 | 0.006 |
| BaP | 1.5 | 45.15 | 99.69 | | | |
| BaP | 1.5 | 34.80 | 23.58 | | | |

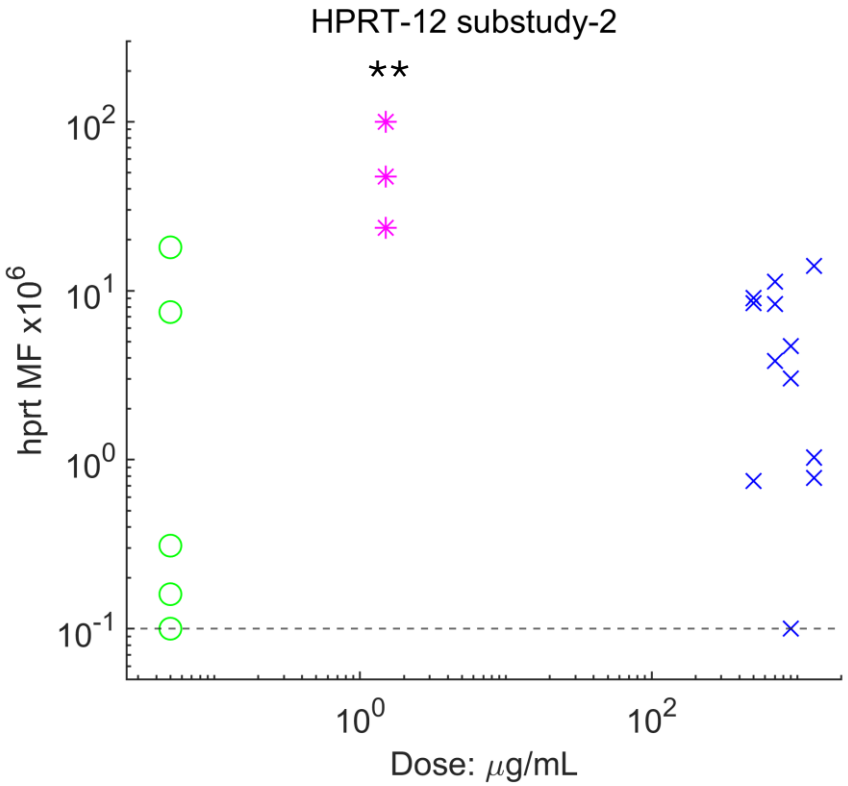



**Figure 15 – Mutant Frequency Dose-Response Relationships for Water and NDIPA (substudy "HPRT-12").** *Hprt* mutant-frequency responses (MF $x10^{-6}$) for each test concentration. (A/B) Green circles depict replicate vehicle control data whilst blue crosses denote responses to the test compound and magenta stars indicate positive control responses. Data-points placeholdered on the dashed horizontal line indicate zero mutant-frequency responses that fell below the limit-of-detection of the assay. Pairwise testing was carried out according to the decision-tree described by Johnson et al. [2014] with response significance determined by the one-sided Dunnett's test indicated by the asterisks (where: *p<0.05, **p<0.01, ***p<0.001).

**1-Cyclopentyl-4-nitrosopiperazine (CPNP): HPRT-14**

CPNP also belongs to the nitrosamine compound class, however it is a piperazine derivative. It is considered a possible human carcinogen and it is structurally similar to a known rodent carcinogen (1-nitrosopiperazine) (Love et al., 1977) as well as a potent mutagen in the Ames test, often being used as a robust positive control (Thomas et al., 2025).

Results are shown in Table 16 and Figure 16. Based on previous test (HPRT-10, 11 and 12) DRFs aligning with previously-generated toxicity data in the MLA, the data for CPNP was reviewed and used to inform on top concentration. As no toxicity was observed in the MLA up to 10mM, 1832.6 µg/mL was used for the top dose group. Concentrations of 500, 1000 and 1500 µg/mL were also assessed to inform on response across the dose range. Again, BaP was treated in triplicate at 1.5 µg/mL. CPNP showed slight toxicity compared to concurrent vehicle controls, with the top dose group survival being reduced to 54%. There were slight increases in MF at the top concentration of 1832.6 µg/mL, however this was not statistically significant, again likely due to the variation between replicates in that group. The mean group MF at the top concentration was 11.7 $x10^{-6}$, which compared to the mean of the vehicle controls is a 7-fold increase, clearly indicating potential mutagenic activity. In this experiment, the mean MF of the vehicle controls seemed lower than that seen in previous experiments (*i.e.,* 1.7 compared to 4.6, 8.9 and 4.3 in HPRT-10, 11 and 12, respectively), indicating that spontaneous background MF was particularly low. The positive control yielded a statistically positive response ($p = 0.005$), indicating a valid test.

Because CPNP is such a potent mutagen in the Ames test, it was surprising that this was not reflected in the *hprt* result. Clear increases in two replicates at the top dose well above the mean of the vehicle control suggests mutagenic activity, however the response in the third replicate was in-line with the vehicle controls. In this way, the p-value at this dose did not exceed 0.05 and therefore a positive outcome could not be called.

**Table 16 – Survival and Mutant Frequency Data for CPNP (HPRT-14)**

| Compound | Concentration (µg/mL) | % Survival | Mutant Freq. (x10$^{-6}$) | Mean Mutant Freq. (per million cells) | Fold Increase | P Value |
|---|---|---|---|---|---|---|
| Methanol | 0 | 92.77 | 0.30 | 1.7 | NA | NA |
| Methanol | 0 | 78.57 | 3.95 | | | |
| Methanol | 0 | 90.79 | 5.12 | | | |
| Methanol | 0 | 78.25 | 0.44 | | | |
| Methanol | 0 | 79.15 | 0.00 | | | |
| Methanol | 0 | 81.50 | 0.15 | | | |
| CPNP | 500 | 71.79 | 0.00 | 0 | 0 | 1.0 |
| CPNP | 500 | 67.38 | 0.00 | | | |
| CPNP | 500 | 65.07 | 0.00 | | | |
| CPNP | 1000 | 68.11 | 11.61 | 4.5 | 2.7 | 0.6 |
| CPNP | 1000 | 66.83 | 1.94 | | | |
| CPNP | 1000 | 54.87 | 0.00 | | | |
| CPNP | 1500 | 59.84 | 8.47 | 2.9 | 1.8 | 0.8 |
| CPNP | 1500 | 50.75 | 0.35 | | | |
| CPNP | 1500 | 69.76 | 0.00 | | | |
| CPNP | 1832.6 | 68.15 | 20.15 | 11.7 | 7.1 | 0.2 |
| CPNP | 1832.6 | 50.18 | 14.64 | | | |
| CPNP | 1832.6 | 44.54 | 0.35 | | | |
| BaP | 1.5 | 50.10 | 47.42 | 56.9 | 34.3 | 0.005 |
| BaP | 1.5 | 45.15 | 99.69 | | | |
| BaP | 1.5 | 34.80 | 23.58 | | | |

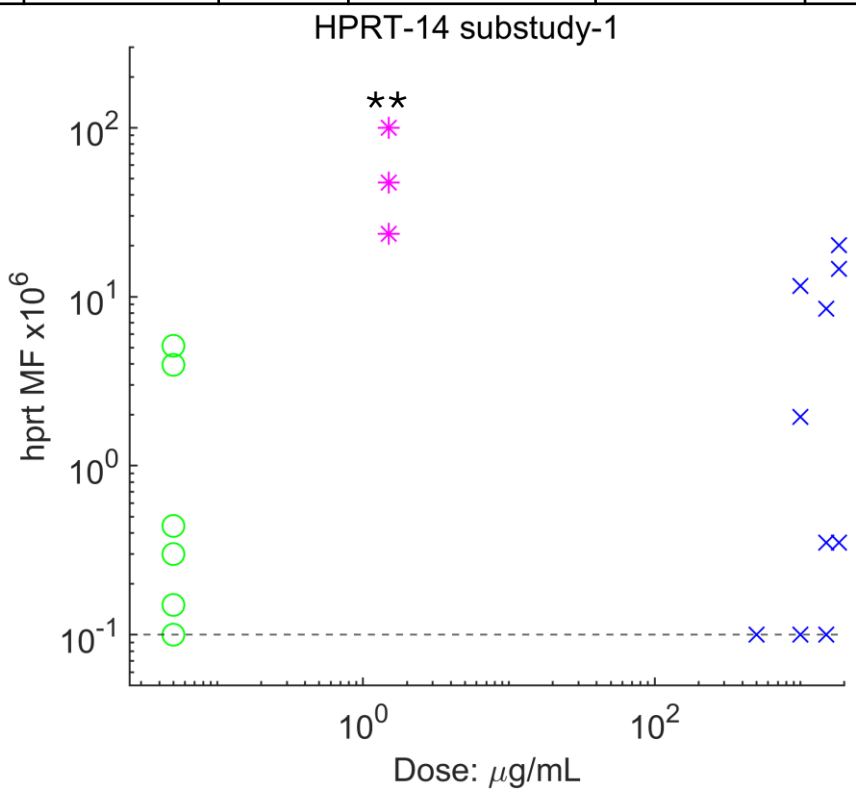



**Figure 16 – Mutant Frequency Dose-Response Relationships for Methanol and CPNP (substudy "HPRT-14").** *Hprt* mutant-frequency responses (MF x10$^{-6}$) for each test concentration. (A/B) Green circles depict replicate vehicle control data whilst blue crosses denote responses to the test compound and magenta stars indicate positive control responses. Data-points placeholdered on the dashed horizontal line indicate zero mutant-frequency responses that fell below the limit-of-detection of the assay. Pairwise testing was carried out according to the decision-tree described by Johnson et al. [2014] with response significance determined by the one-sided Dunnett's test indicated by the asterisks (where: *p<0.05, **p<0.01, ***p<0.001).

### 7.4 In Vitro Test Data Comparison

Nitrosamines are a chemical class that exhibit a broad range of mutagenic and carcinogenic potencies which have gained significant attention in recent years in the field of toxicology. Their presence in various consumer products and pharmaceuticals necessitates accurate and reliable genotoxicity testing methods to ensure safety compliance and enable regulatory decision-making regarding justifiable acceptable intake limits. Among the suite of assays used to evaluate genotoxic potential, two approaches - the MLA and *hprt* assays – are pivotal components of the standard battery of genotoxicity tests. Whilst both assays provide critical insights into mutagenic potential, discerning their suitability for nitrosamine risk assessment remained a key objective of this research.

This section presents a comparative analysis of the *hprt* and MLA assay outputs to evaluate their efficacy in detecting genotoxic effects induced by nitrosamines. The *hprt* assay, utilising the L5178Y cell line, measures mutations that confer resistance to 6TG, with a focus on mutation specifically. On the other hand, the MLA also utilises L5178Y cells however is able to detect both point mutations and chromosomal aberrations, representing a broader scope of genomic changes (Wang et al., 2009).

By contrasting assay performance in terms of sensitivity, specificity, practicality and applicability to nitrosamine risk assessment, this analysis aims to elucidate which assay best captures the genotoxic properties of these compounds in a robust manner. Whereas the MLA assays were conducted using both rat or hamster S9-mix, the comparison is made to those results obtained in the presence of rat S9-mix only to present a direct comparison to the test conditions used in the *hprt* studies.

**Table 17** – Comparison of MLA and *hprt* Results for Tool Nitrosamines

| ***Compound*** | ***MLA Result*** | ***hprt Result*** |
|---|---|---|

| | | |
|---|---|---|
| ***NDMA*** | Positive at 10 µg/mL | Positive at 500 µg/mL |
| ***NDEA*** | Positive at 800 µg/mL[1] | Positive at 600 µg/mL |
| ***NDIPA*** | Negative | Negative |
| ***CPNP*** | Positive at 100 µg/mL[1] | Negative |

[1]*Lowest concentration tested*

Table 17 shows the comparison between both assays, with respective lowest observed genotoxic effect concentrations where applicable, for the nitrosamines tested. The agreement of both assays was successful for three out of four nitrosamines tested. Both assays identified NDMA as genotoxic, albeit at different concentration thresholds. The MLA detected a positive result at a lower concentration (10 µg/mL), while the *hprt* assay gave a positive result at a substantially higher concentration (500 µg/mL) which suggests higher sensitivity in the MLA. In the case of NDEA, the MLA exhibited a genotoxic response at 800 µg/mL, while the *hprt* assay reported positive results at a slightly lower concentration of 600 µg/mL. These results demonstrate comparable sensitivity between the two assays in detecting NDEA's mutagenic effects, with the *hprt* assay showing marginally greater sensitivity. Testing of NDIPA in both assays returned negative results under standard rat S9 conditions, which aligns with the earlier explanation that highlights NDIPA's weak mutagenic properties due to the challenge of successful metabolic activation. Interestingly, the MLA assay does record positive results for NDIPA in the presence of hamster S9, suggesting the potential influence of metabolic activation conditions on NDIPA mutagenicity detection, which could be investigated in future *hprt* experiments. A divergence in assay results was noted for CPNP, with the MLA assay reporting a positive outcome at 100 µg/mL, while the *hprt* assay did not achieve statistical significance at any concentration up to 10 mM. Despite this, notable increases in mutant frequency in the *hprt* assay suggest underlying mutagenic activity warranting further investigation through follow-up tests.

It is important to note that the way these assays are evaluated differs. The discordance between the MLA and *hprt* result for CPNP, as well as the variance in lowest-observed genotoxic concentrations for NDMA between both assays highlight how differences in interpreting data can influence overall result. The MLA employs a straightforward global evaluation factor (GEF) approach, calculated as the mean mutant frequency plus 126, to determine the threshold for a positive response. This approach provides a consistent benchmark that may be more sensitive to variations in mutant frequency across different concentrations, potentially capturing positive results even with smaller changes in frequency. In contrast, the *hprt* assay utilises more nuanced and statistically rigorous methods, evaluating significance based on pairwise statistical tests that account for both the mean mutant frequency and variability within replicates. Consequently, borderline or modest increases in mutant frequency, such as those observed for CPNP or where variance between replicates is high, may fail to demonstrate statistical significance within the *hprt* framework, resulting in its classification as 'negative', when in fact, as in the example of CPNP, the mean MF of the highest dose group is over 7-fold higher than the vehicle control. In effect – whereas a p-value of less than 0.05 provides confidence in a robust positive outcome, p values greater than 0.05 do not mean the response is the same as that observed for the vehicle controls.

An additional point to note is that the MLA has the capability of detecting additional damage that the *hprt* assay cannot, such as chromosome breaks, due to its heterozygous *tk* gene as explained in Section 1.4. This broader detection capability could explain the greater sensitivity in the MLA as it detects a broader spectrum of damage than the *hprt* assay. As shown in Section 1.7, nitrosamines initiate mutations through several mechanisms, predominantly in the form of DNA alkylation (Li and Hecht, 2022). Processing of these DNA base modifications via DNA repair mechanisms such as base excision repair (BER) can lead to strand breaks and replication fork collapse, which can additionally give rise to large deletions and chromosomal rearrangements if mis-repaired (Gasser, 2026), even further compounding the amount of genetic damage caused.

Additionally, nitrosamines are known to generate reactive oxygen species (ROS) (de la Monte and Tong, 2009), which compromise the DNA backbone and cause up to 2300 single strand breaks per cell per hour in mammalian cells (Chatterjee and Walker, 2017). These strand breaks, large deletions and chromosomal rearrangements are all types of genetic damage that the MLA is effective at detecting over the *hprt*, which due to its lack of breadth, may underestimate mutagenicity and offer an explanation for the differences in assay sensitivity.

### 7.5 Evaluation of the In-House *hprt* Assay Protocol

*DRF Design*

The Dose Range Finder (DRF) experimental design serves as a preliminary assessment tool that aids in optimising the concentration range for subsequent testing. Given the substantial number of replicates required in the main test, deploying a DRF using single culture replicates enables efficient gauging of the % survival at various concentrations, thereby facilitating targeted investigations within the critical toxic range (20-30% survival), sparing resources such as materials and human effort in the lab. The % survival for DRF experiment versus main test was compared where both types of tests were performed (Figure 17). The general survival trend observed across compounds in preliminary DRF experiments correlates sufficiently with outcomes noted in the main test. This correlation supports the notion that a single-replicate DRF can satisfactorily guide concentration selections for compounds exhibiting low to moderate toxicity profiles. However, when dealing with increasingly toxic compounds, variability between DRF results and main test outcomes poses a significant challenge in predicting toxic concentrations. An example of large variation in the four examples presented is in NDMA (Figure 17 (B)), where in the DRF the % survival for the top concentration was 68%, whereas in the main test it was reduced to 32% - less than half of the survival originally generated. Although they did not have an impact on the evaluation of the main test MFs, such fluctuations could cause potential imprecision when extrapolated to compounds with more pronounced cytotoxicity. The

fluctuations could result in misidentifying doses within the biologically relevant toxic range, thereby compromising the main test *i.e.,* not reaching a limit dose due to % survival increasing or decreasing from DRF to main test, which could result in repeat experiments having to be conducted. For example in the MMS comparisons (Figure 17 (A)) the DRF survival at 30 µg/mL was 24% - the ideal level of cytotoxicity to detect biologically relevant increases in MF, however in the main test the survival was reduced to 9.5% at the same concentration, which would render any increases in MF at this level as a cytotoxic response and not biologically relevant and therefore more work may have to be done to find a biologically relevant response.

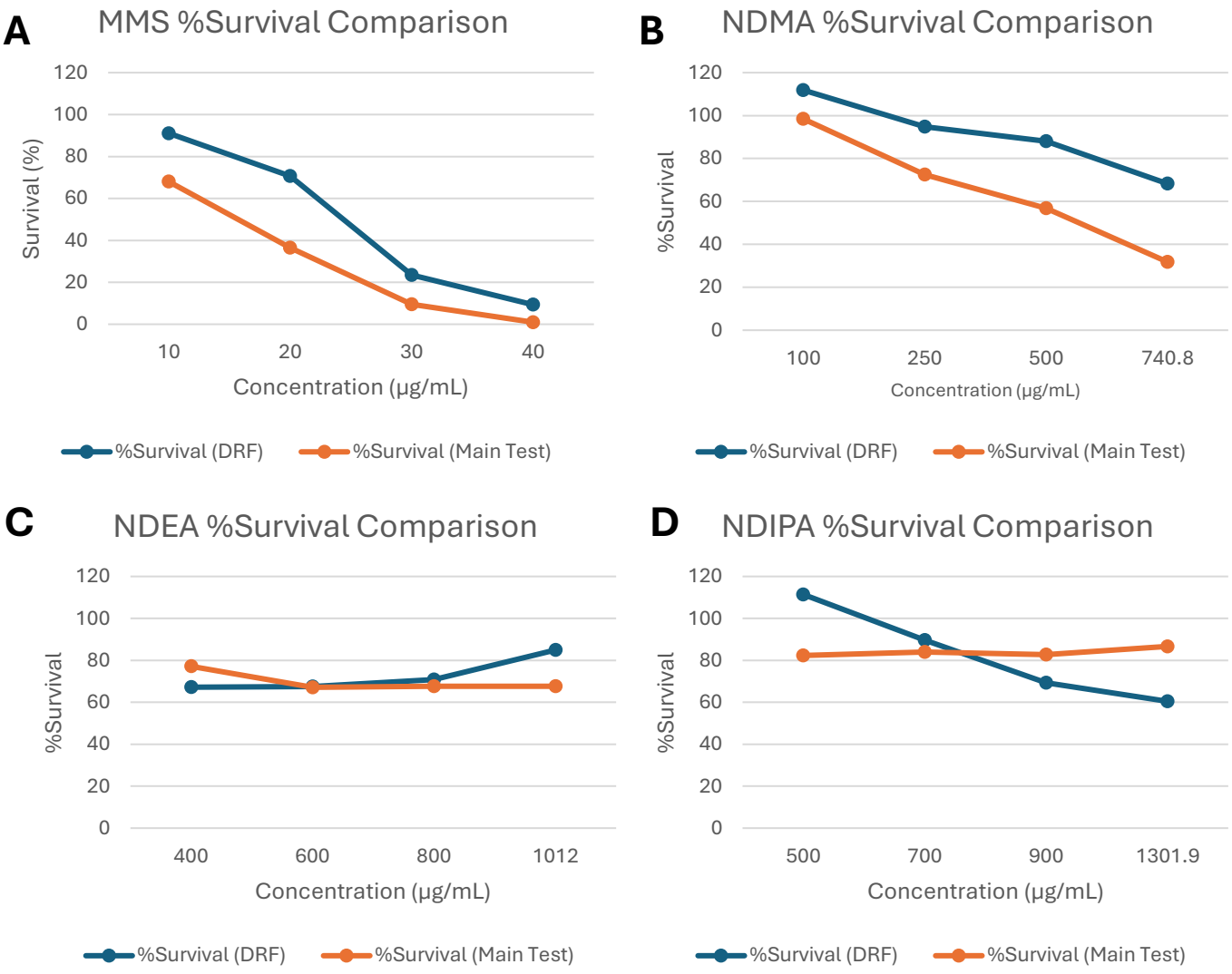



**Figure 17 – Comparison Between DRF and Main-Test Survival Values.** Comparison of % survival plotted against dose for DRF experiments (blue) and main test experiments (orange) for (**A**) MMS (HPRT-8), (**B**) NDMA (HPRT-10), (**C**) NDEA (HPRT-11) and (**D**) NDIPA (HPRT-12).

Therefore, the DRF design has proven to be an effective tool for preliminary toxicity assessments within the L5178Y cell culture environment, particularly for compounds that are non-toxic or exhibit little toxicity – as it allows generation of survival data without the necessity for extensive replicates. This streamlined approach allows progression into the main testing phase with a concise set of dose groups, typically just four or five, thereby focusing resources effectively. However, the DRF design may be more constrained for compounds exhibiting higher toxicity, and the variation in survival observed between DRF and main test highlights the need for caution in dose selection. To address this, when dealing with these types of compounds, a more conservative approach should be adopted which could include considering an increased number of dose-groups to ensure accurate capture of the survival range and mitigate the risk of missing the target range whilst accommodating potential variability in survival response.

As the MLA utilises a single-replicate methodology, survival data can be generated as part of the large main test which allows a wide dose range to be assessed without the preliminary step and highlights a fundamental difference in experimental design and application scope when assessing laboratory resource requirements between the two assays.

*Replicate Variability in MF Data*

Consistently throughout the results section it is evident that the *hprt* assay often exhibits high variability between dose-group replicates which negatively impacts the utility of the assay through affecting the following:

- **Signal to noise ratio**: The main objective of the *hprt* assay is to detect meaningful increases in MF between dose groups. When variability within dose groups is high, the noise (*i.e.,* random variation within groups) may overshadow the signal (*i.e.,* the true effect or difference between groups). Therefore, the larger the variability, the harder it is to discern a significant difference between the mean of the dose group and that of the vehicle control.

- **Statistical testing**: The Dunnett's test is designed to compare multiple treatment groups against a single control group. It employs a pooled variance estimate to manage the variability across groups, alongside utilisation of multiple comparisons correction using adjusted critical values that become higher with the increasing number of comparisons made to maintain robust significance levels and avoid false-positive outcomes. When there is wide variation within replicates of the treatment dose groups, this variability inflates the pooled variance and results in larger standard errors in the statistical calculations. Consequently, the test's sensitivity to detect a positive response is lessened. With stringent critical thresholds required for significance, high within-group variability makes it more challenging to achieve significant results unless the compound exerts a pronounced mutagenic effect.
- **P-values and statistical power**: As described above, high in-group variability impacts the statistical testing. This in turn has an effect on the statistical power, reducing the sensitivity to detect a real difference between groups even if a true one exists.

Overall, to precisely detect small biologically relevant increases in MF within this assay for accurate hazard and risk assessment, it is crucial to manage and reduce within-group variability. One way to do this would be to further increase the number of replicates in each group to improve the accuracy and reliability of the mean estimates of each group. However, this strategy would pose a significant burden on time and laboratory resources.

*Cell Type*

The L5178Y *tk*$^{+/-}$ cell line is well characterised and widely accepted by regulators. Because they are suspension cells with a low doubling time, they are very practical for use in the *hprt* assay, as well as having established published spontaneous MF control ranges to compare laboratory data against. However, the choice of cell line may have been a limitation of the studies outlined

in this work, as both the underlying biology of the *hprt* locus and the broader characteristics of this cell line have potential to constrain the sensitivity and reliability of the results.

The *hprt* gene is located on the X chromosome and is therefore heterozygous in L5178Y cells. Whilst this simplifies mutant selection, it has consequences for cell recovery against a broader range of DNA damage. As discussed in Section 7.4, this means that damage such as large deletions that inactivate the *hprt* gene will lead to cell death and as a result these types of damage are completely lost from analysis, creating a bias towards the detection of only small intragenic changes (McGregor et al., 1996).

Secondly, L5178Y cells possess strain-dependent variability in mutation at the *hprt* locus. The L5178Y $tk^{+/-}$ 3.7.2C variant of the cell line was used for this work which is derived from the L5178Y-S (LY-S) strain (ATCC). Research suggests that differences in strain (LY-S versus LY-R) can show >500-fold differences in UV-induced *hprt* mutant frequencies, with LY-S often giving mutant frequencies at or below the limit of detection even at high doses (Jacobson et al., 1984), as well as marked reduced MFs at the *hprt* locus for other genotoxins such as N-Nitroso-N-methylurea (NMU) and ethyl methanesulfonate (EMS) (Evans et al., 1986). This strong emphasis on differences in strain dependence means that *hprt* mutagenicity in L5178Y cells can be very low or highly variable, which is what has been observed in this work, reducing the overall assay sensitivity and robustness.

L5178Y cells are a tumour-derived cell line and are therefore inherently genomically unstable (Paeratakul, 1986). As well as the strain differences as mentioned above, this can also mean that within a single experiment, cultures can contain multiple subclones with distinct genetic and phenotypic properties (e.g., differences in growth rate, DNA repair capacity and mutation susceptibility). During routine culture, especially over extended passages, certain subclones can gain an advantage over others and progressively dominate the population which results in a sub clonal shift and can have a range of consequences in the *hprt* assay (Evans et al., 1986).

Firstly, it can affect the background spontaneous mutant frequency and cause variation, depending on which subclone is seeded in the mutation plates, like observed in this work for example, with big differences between vehicle control cultures of the same experiment. It can also cause changes to the apparent sensitivity to mutagens, since subclones may differ in their ability to be resistant to 6TG, for example (Myhr and Caspary, 1988). This can therefore result in large variance in the responses to mutagenic compounds. An example of this is in the treatment with CPNP at the highest concentration whereby the MFs in this treatment group were 20, 15 and 0.35 $x10^{-6}$. Here, in the case of the third outlier value, the culture could have experienced sub clonal shift toward a clone with lower *hprt* mutability and therefore appears much lower than the other replicates.

An additional limitation of the L5178Y model is its deficiency in functional p53, a key tumour suppressor involved in DNA damage responses. L5178Y cells have been shown to lack a wild-type Trp53 allele, resulting in impaired p53 activity and absence of normal checkpoint control mechanisms (Clark et al., 2004). Consequently, these cells do not undergo effective G1/S arrest following DNA damage, allowing replication to proceed in the presence of DNA lesions (Szumiel et al., 2000). This has important implications for the *hprt* assay, as the loss of p53-mediated apoptosis permits survival and proliferation of genetically damaged cells that would otherwise be eliminated by DNA damage repair pathways. As a result, DNA damage is more likely to be fixed as mutations, potentially leading to an overestimation of mutagenic effects. Furthermore, the absence of proper cell cycle control and repair pathways may alter the spectrum of mutations observed, reducing the physiological relevance of the assay. When combined with the inherent bias of the heterozygous *hprt* locus towards small mutations, p53 deficiency further skews the detectable mutational profile by becoming even further selective towards these types of mutations.

Taken together, these points mean that L5178Y cells are not an optimal choice for a *hprt* mutagenicity assay when the goal is a sensitive, broadly informative, and well-standardised in vitro gene-mutation test. For that purpose, CHO or V79 cell lines may be better suited as they are well established for use in this assay, or alternatively the human lymphoblastoid TK6 cell line, which is a human-relevant, p53 competent and more genetically stable choice.

### 7.6 Scope for Further Assay Development

#### 7.6.1 Automated Plate Scoring: Colony Identification using Pixel Classification Machine Learning

In order to overcome the issues encountered regarding inter-replicate variability, more replicates per-dose could be added to each experiment. This would however result in generation of additional 96-well plates for survival, viability and mutation analysis which subsequently relies on time-consuming manual scoring of colonies (as described in the methods section).

Work to automate the scoring of these individual colonies began with reviewing automation of scoring for other genetic toxicology assays such as the micronucleus assay (Wills et al., 2021) and comet assay (Collins et al., 2023) which were achieved using machine learning image analysis methods, highlighting the potential to adapt these methods to enable scoring of *hprt* plates. The image analysis tool 'ilastik' was chosen for this primary investigation as it is a user-friendly and well supported package for scientists with little programming expertise (Berg et al., 2019) which could have the capability to fulfil the aim of this preliminary investigation, which was to see if colonies could be automatically identified and segmented using pixel classification machine learning (Durkee et al., 2021). Well-plate images were acquired using a GelCount imager (Oxford Optometrics). Feature extraction of input images (Figure 18) producing image variations with transformations designed to enable colony identification (*e.g.*, edge enhancement) were then carried out in the ilastik software using the pixel-classification

workflow. Figure 19 shows how brush tools were used to carry out sparse annotation on the set of training images with the goal to train a model able to identify two categories - either 'colony' (yellow) or 'other' (blue). After training on a small set of images, the random forest algorithm was shown able to classify previously unseen images predicting the probability of each pixel belonging to a colony represented by the areas of white in Figure 20. These probability images were then subjected to morphological cleaning (thresholding, and removal of any regions less than 25 pixels) to provide a final colony prediction which can be seen by the blue overlay (Figure 20).

This initial work demonstrates proof-of-concept that automated well-plate imaging and pixel classification machine has the potential to automate colony identification which provides the potential for a new technology to replace the manual effort of scoring. Moreover, with additional work, the ilastik predictions could be used to expedite building of a large number of labelled training images such that a more powerful artificial intelligence algorithm (*e.g.*, a convolutional neural network) could be trained to achieve the scoring task. With further work this could result in an end-to-end automated scoring which could then allow increased numbers of replicate doses to be scored without increasing the human resource required to manually score them. Currently, manual scoring of a 96-well plate can take an average of 60 seconds for a positive control plate as time must be taken to carefully decipher cell colonies from debris. In this way automated scoring could increase throughput providing a potential solution to increasing the number of scorable replicates which would consequently increases statistical power and reduce in-group variability without increasing the human resource. Additionally, using this method digitalises all data and provides consistency across plates, removing the potential for scoring bias.

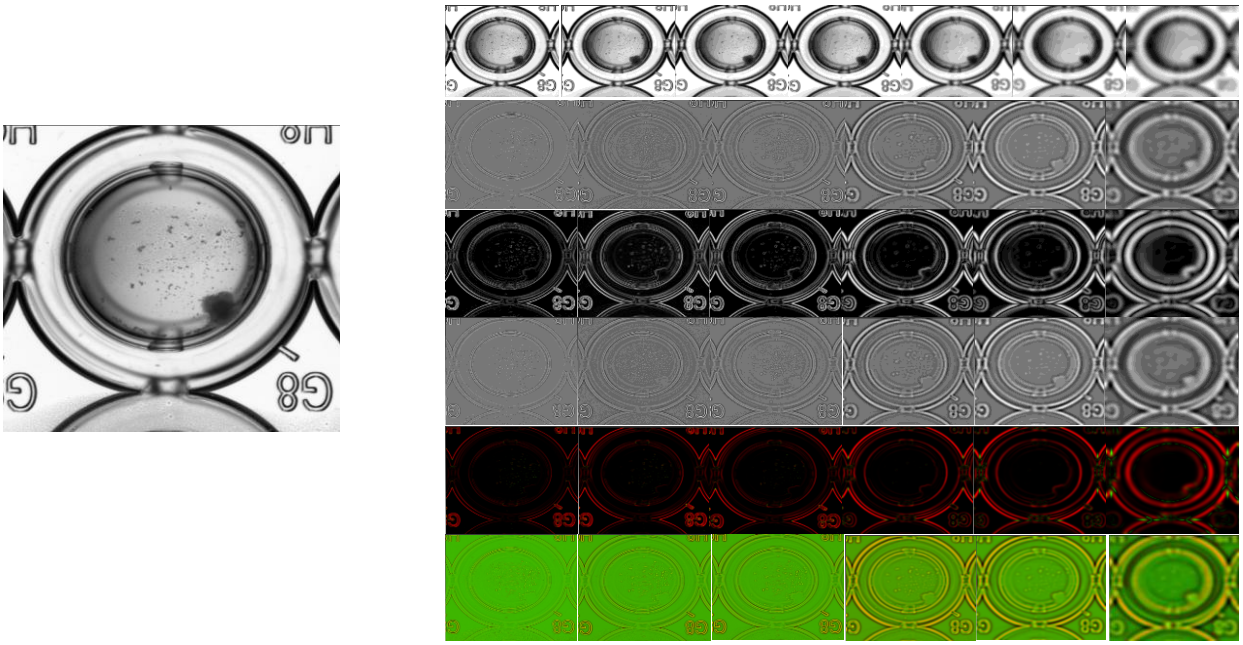


**Figure 18 – Feature Extraction of Brightfield Colony Input Image.** The image on the left was taken using Oxford Optronics GelCount and input into Ilastik software to produce feature extraction using many different channels (right).

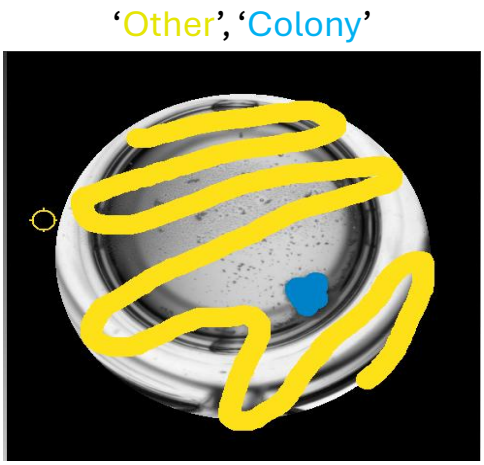



**Figure 19 – Sparse Annotation of a Masked, Colony-Containing Image.** A mask was applied to the image to remove background artefacts such as etching on the 96-well plate. Pixel annotations were made in the Ilastik software providing training data to distinguish between "cell colony" and "other" categories.

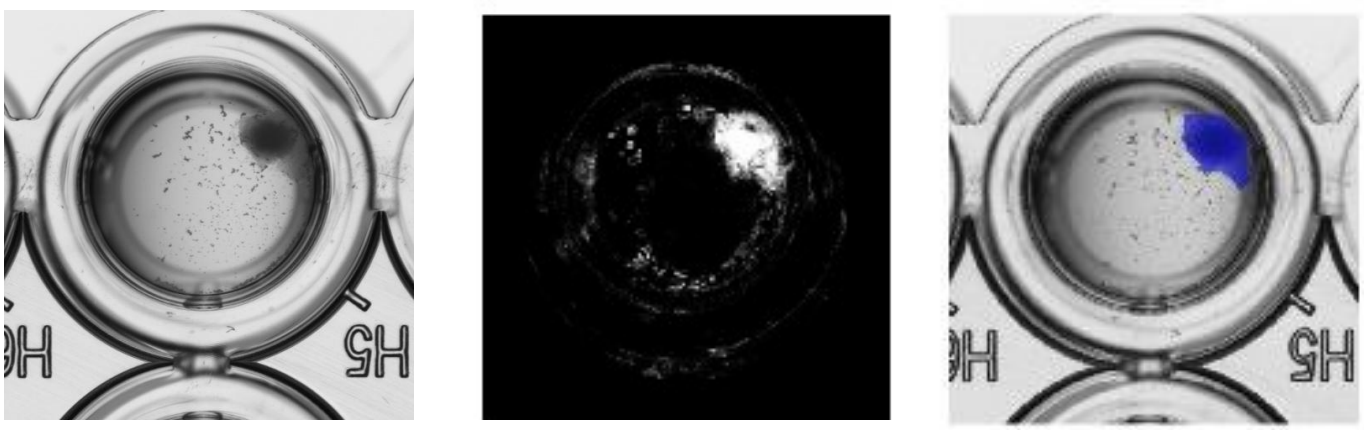


**Figure 20 – Automated scoring proof-of-concept with unseen image-data.** Left to right: Input image, algorithm prediction of the probability of each pixel belonging to a colony before morphological clean up, input image with an overall (blue) of the final, automated colony prediction.

### 7.6.2 Combined Benchmark Dose (BMD) Modelling

Another route to combatting variability between *hprt* assay replicates could utilise combined-covariate BMD modelling. The FDA technical guidance on BMD modelling states that datasets that share statistical and biological features can be combined prior to dose-response modelling to give increased confidence in calculated BMDs (EPA, 2012). This supports utilisation of all available data and in particular where responses are variable and data harder to interpret, it allows a BMD to be calculated. This is demonstrated in the example below (Figures 21 and 22).

As seen for CPNP (HPRT-13, Figure 16), simple, positive or negative test outcomes determined by pairwise statistical testing methods have limitations, and as demonstrated in this case, clear elevations in MF do not satisfy a positive statistical test due to variability between replicates. A potential strategy to combat this is combined BMD modelling which is now being widely used to quantitatively define points of departure and enable robust potency rankings (Wills et al., 2016a, Wills et al., 2016b). BMD modelling works by fitting mathematical functions to dose-response data in order to interpolate the dose that can be expected to cause a specific increase in response over the control (*i.e.,* 50%) (Wills et al., 2016a). Combined BMD modelling can also be used to simultaneously assess multiple dose-response curves (*e.g.*, arising across different compounds) for a shared endpoint, such as MF. Importantly, because we have demonstrated in Figure 21 that the 'shape and steepness' parameters (*c* and *d*, respectively) are similar enough to be estimated as constants within an endpoint (Slob and Setzer, 2014) this approach can improve precision in each individual BMD estimate for the compounds included in the combined analysis. This is because these constant parameters can be estimated from all data combined, leading to more accurate estimations. To investigate the potential of this approach with the *hprt* assay, five positive data sets were chosen for proof-of-concept HPRT-4 (MMS), HPRT-8 (BaP), HPRT-10 (NDMA), HPRT-11 (NDEA) and HPRT-14 (CPNP). Although on first glance the dose responses for each study differ based on the variance within controls, Figure 21 (right

side panel) demonstrates how, the dose-response data can be well-described using curves with constant shape and steepness parameters.

Figure 22 shows the predicted BMD confidence intervals for each compound when they are analysed one at a time using independent BMD analysis (A) versus as the product of a combined-covariate analysis (B). The individual analyses run for MMS, BaP and NDEA obtained bounded BMD estimates (*i.e.,* with upper and lower confidence intervals) however produced confidence intervals (CI) that were generally quite uncertain and spanned a wide dose-range. In the independent BMD analyses, despite use of single replicates per dose group, the dose response data for BaP best captured the complete dose-response relationship curve due to the wide range of test-concentrations analysed. This results in a CI that was much smaller compared to those for MMS and NDEA. Here the BMD CIs are much wider because there is less information about the shape of the response in the low-dose regions close to the BMR 50% interpolation size. For NDMA and CPNP no BMD could be predicted due to the data being too sporadic resulting in the model being unsure if a true monotonic response was established.

In contrast, when the datasets for all compounds were analysed together harnessing information from all dose response curves in the combined-covariate method, this significantly improved model confidence in BMD prediction. Bounded BMDs were now achieved for all compounds – even for NDMA and CPNP – which could not be previously generated using independent analysis. Moreover, the narrower BMD confidence intervals obtained started to allow differences in the potencies of the different compounds to be determined *i.e.,* NDMA and NDEA potencies are quite similar (within 30-fold), and that tool genotoxic positive controls (MMS and BaP) are significantly more potent (~ 100 to 1000 times more potent) than NDEA and NDMA in this assay. Importantly, combined BMD modelling also resolved a probable monotonic trend for CPNP allowing its potency to be compared to NDMA suggesting that it could be ~30 –

100-fold less potent than NDMA / NDEA if the true BMD lies close the predicted BMDU (*i.e.,* the least conservative interpretation of the confidence intervals).

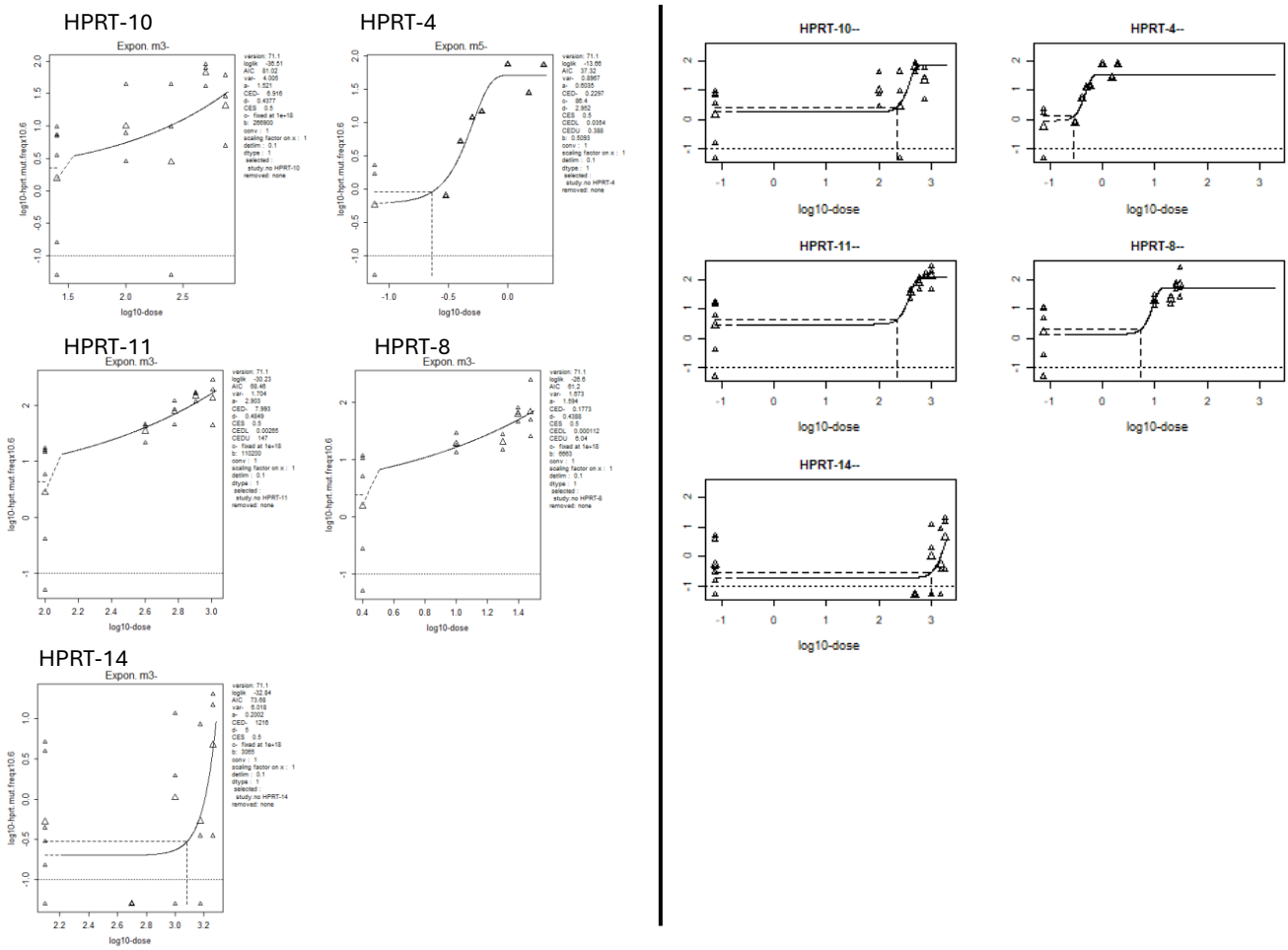



**Figure 21 - Benchmark Dose Modelling to Determine Critical Effect Doses for a Subset of Test Compounds.** This shows the fitted models and interpolation (horizontal lines) to determine the critical effect doses using a critical effect size of 50% for NDMA (HPRT-10), BaP (HPRT-4), NDEA (HPRT-11), MMS (HPRT-8) and CPNP (HPRT-14) when analysed independently (left), or using the combined-covariate BMD method (right).

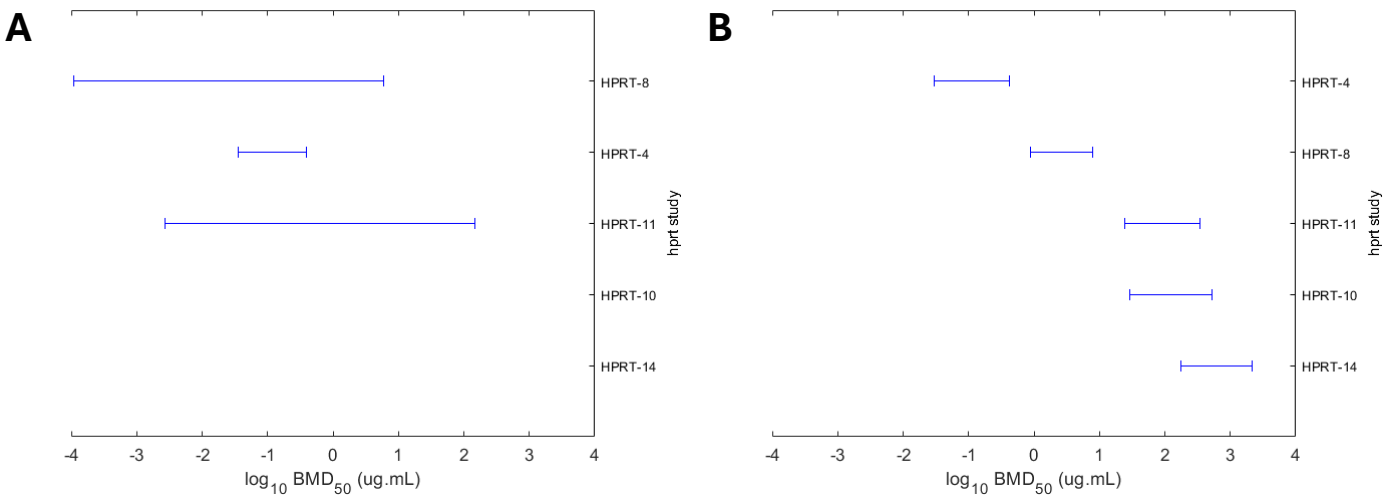



**Figure 22 – BMD Confidence Interval comparison between Independent and Combined Analysis approaches.** Plots show the 90% BMD confidence intervals of the $BMD_{50}$ for NDMA (HPRT-10), BaP (HPRT-4), NDEA (HPRT-11), MMS (HPRT-8) and CPNP (HPRT-14) when the five data sets were analysed independently (left; A) or in combination, using test compound as the covariate (right; B). For the combined analysis, parameters for maximum response (c) and log-steepness (d) were modelled as constants using all data across all compounds. The combined analysis shows that combined BMD analysis for these five data sets produces more precision and was even able to predict BMDs for NDMA (HPRT-10) and CPNP (HPRT-14) which was not possible when the datasets were analysed individually.

This pilot analysis suggests that combined BMD analysis – harnessing all historical dose-response data available for the *hprt* assay – has significant potential to improve the utility of *hprt* assay data for compound potency ranking. With more dose-response data added, *e.g.,* running tool mutagens through the assay over a wide concentration range, the data can be used to further strengthen the model and provide even more confidence to calculated BMDs. This could in-turn require less experimental data for new compounds in the future (*i.e.,* less replicates or less concentrations for analysis) yet achieve outputs containing compound potency information. In this way, combined covariate BMD modelling could contribute to less resource required to run the assay whilst still enabling robust, quantitative potency estimates to support drug development and risk assessment processes.

**Conclusions**

The nitrosamine impurity crisis in pharmaceuticals is constantly evolving, with regulatory bodies such as the FDA and EMA now requesting *in vitro* mammalian cell test data to bolster the overall risk assessment and potentially justify increases in acceptable intake limits, when supported by the outcomes of *in vivo* studies.

The two overall aims of this research were:

1. To establish and OECD-compliant *hprt* protocol in-house for mutagenicity testing of GSK API related nitrosamine impurities.
2. To assess the suitability of the *hprt* assay for nitrosamine risk assessment, using equivalent MLA data to advise on potential drawbacks of the assay and providing a justification for use of one *in vitro* mammalian assay over the other.

We offer the following concluding statements regarding the validation of an OECD compliant in-house *hprt* protocol and more generally its utility in nitrosamine risk assessment:

1. This work developed an OECD-compliant, in-house testing protocol that is ready for use going forward for to support GSK compound testing. The study design was optimised to address inter-replicate variability and satisfy OECD guidelines around statistical testing. This protocol was able to successfully detect the mutagenicity of tool genotoxicants (BaP and MMS), as well as the small-alkyl nitrosamines (NDMA and NDEA).
2. A DRF study to assess unknown toxicity profiles can be performed upfront to determine a focused concentration range whereby labs can increase replicates within these critical dose groups without additional resource allocation. It was found that for non/low toxicity compounds this allowed testing up to 10mM (the maximum concentration in accordance with current guidelines) however for highly toxic compounds, more concentrations should be assessed to ensure the full toxicity range (20-30% survival) is captured.
3. Despite the general agreement in results for three out of four nitrosamines tested, distinct differences were identified regarding sensitivity between *hprt* and MLA assays, with the *hprt* assay exhibiting high levels of variability, which can significantly affect the reliability and clarity of statistical outcomes, as evidenced through testing of CPNP.
4. The *hprt* assay requires additional resource when compared to the MLA, requiring a DRF study, longer mutation fixing period, and at least triplicate replicates per dose – requiring a longer turnaround time, more materials and increased personnel.
5. Challenges concerning additional resource could be resolved through optimisation of the assay to include an automated plate scoring pipeline, as well as the employment of combined BMD analyses to make best utility of all available data.
6. As both the *hprt* and MLA were able to detect increases in MF, the selection of which assay to perform for nitrosamine assessment may depend on the specific laboratory resources available. The *hprt* may be the favourable option based on the fact that it focusses solely on a mutational endpoint which satisfies regulatory questions around

compounds with specific mechanisms of genotoxicity. However, where resources are constrained, the MLA is also a viable option for nitrosamine risk assessment, with a shorter turnaround time and less labour involved.

This work has enhanced GSK's testing strategies for drug-like nitrosamine impurities through the use of 3Rs-compliant *in vitro* mammalian cell assays, yielding an *hprt* assay protocol that is 'ready now' for immediate deployment. By establishing an OECD-compliant in-house protocol and rigorously comparing the *hprt* and MLA assays, we have refined GSK's hazard and risk assessment toolkit for evaluating nitrosamine mutagenicity and paved the way for future innovation including machine learning scoring automation and the expanded use of combined BMD modelling. These advancements will equip regulatory bodies with more detailed insights into the relative mutagenicity of compounds. With these robust methods in hand, GSK is better positioned to generate data that supports regulatory decision-making while safeguarding pharmaceutical quality and public health.

**References**


(WHO), W. H. O. 2019. *Information Note Nitrosamine impurities* [Online]. Available: https://www.who.int/news/item/20-11-2019-information-note-nitrosamine-impurities [Accessed 8 August 2025 2025].

AIDOO, A., MORRIS, S. M. & CASCIANO, D. A. 1997. Development and utilization of the rat lymphocyte hprt mutation assay. *Mutat Res,* 387**,** 69-88.

ALBERTINI, R. J. 2001. HPRT mutations in humans: biomarkers for mechanistic studies. *Mutat Res,* 489**,** 1-16.

AMUNDSON, S. A. & LIBER, H. L. 1992. A comparison of induced mutation at homologous alleles of the tk locus in human cells. II. Molecular analysis of mutants. *Mutat Res,* 267**,** 89-95.

ANDERSON, L. M., SOULIOTIS, V. L., CHHABRA, S. K., MOSKAL, T. J., HARBAUGH, S. D. & KYRTOPOULOS, S. A. 1996. N-nitrosodimethylamine-derived O(6)-methylguanine in DNA of monkey gastrointestinal and urogenital organs and enhancement by ethanol. *Int J Cancer,* 66**,** 130-4.

ATCC. *L5178Y TK+/- clone (3.7.2C) [TK+/- (clone 3.7.2C)]* [Online]. Available: https://www.atcc.org/products/crl-9518#:~:text=+/%2D%20(clone%203.7.-,2C)%5D,be%20used%20in%20toxicology%20research. [Accessed 2026].

BARTSCH, H., CAMUS, A. & MALAVEILLE, C. 1976. Comparative mutagenicity of N-nitrosamines in a semi-solid and in a liquid incubation system in the presence of rat or human tissue fractions. *Mutat Res,* 37**,** 149-62.

BERG, S., KUTRA, D., KROEGER, T., STRAEHLE, C. N., KAUSLER, B. X., HAUBOLD, C., SCHIEGG, M., ALES, J., BEIER, T., RUDY, M., EREN, K., CERVANTES, J. I., XU, B., BEUTTENMUELLER, F., WOLNY, A., ZHANG, C., KOETHE, U., HAMPRECHT, F. A. & KRESHUK, A. 2019. ilastik: interactive machine learning for (bio)image analysis. *Nat Methods,* 16**,** 1226-1232.

CARLSSON, M. J., HERZOG, N., FELSKE, C., ACKERMANN, G., REGIER, A., WITTMANN, S., FERNÁNDEZ CEREIJO, R., STURLA, S. J., KÜPPER, J. H. & FAHRER, J. 2025. The DNA Repair Protein MGMT Protects against the Genotoxicity of N-Nitrosodimethylamine, but Not N-Nitrosodiethanolamine and N-Nitrosomethylaniline, in Human HepG2 Liver Cells with CYP2E1 Expression. *Chem Res Toxicol,* 38**,** 1134-1146.

CHATTERJEE, N. & WALKER, G. C. 2017. Mechanisms of DNA damage, repair, and mutagenesis. *Environ Mol Mutagen,* 58**,** 235-263.

CHRISTIE, N. T., DRAKE, S., MEYN, R. E. & NELSON, J. A. 1984. 6-Thioguanine-induced DNA damage as a determinant of cytotoxicity in cultured Chinese hamster ovary cells. *Cancer Res,* 44**,** 3665-71.

CLARK, L. S., HARRINGTON-BROCK, K., WANG, J., SARGENT, L., LOWRY, D., REYNOLDS, S. H. & MOORE, M. M. 2004. Loss of P53 heterozygosity is not responsible for the small colony thymidine kinase mutant phenotype in L5178Y mouse lymphoma cells. *Mutagenesis,* 19**,** 263-8.

COLLINS, A., MØLLER, P., GAJSKI, G., VODENKOVÁ, S., ABDULWAHED, A., ANDERSON, D., BANKOGLU, E. E., BONASSI, S., BOUTET-ROBINET, E., BRUNBORG, G., CHAO, C., COOKE, M. S., COSTA, C., COSTA, S., DHAWAN, A., DE LAPUENTE, J., BO, C. D., DUBUS, J., DUSINSKA, M., DUTHIE, S. J., YAMANI, N. E., ENGELWARD, B., GAIVÃO, I., GIOVANNELLI, L., GODSCHALK, R., GUILHERME, S., GUTZKOW, K. B., HABAS, K., HERNÁNDEZ, A., HERRERO, O., ISIDORI, M., JHA, A. N., KNASMÜLLER, S., KOOTER, I. M., KOPPEN, G., KRUSZEWSKI, M., LADEIRA, C., LAFFON, B., LARRAMENDY, M., HÉGARAT, L. L., LEWIES, A., LEWINSKA, A., LIWSZYC, G. E., DE CERAIN, A. L., MANJANATHA, M., MARCOS, R., MILIĆ, M., DE ANDRADE, V. M., MORETTI, M., MURUZABAL, D., NOVAK, M., OLIVEIRA, R., OLSEN, A. K., OWITI, N., PACHECO, M., PANDEY, A. K., PFUHLER, S., POURRUT, B., REISINGER, K., ROJAS, E., RUNDÉN-PRAN, E., SANZ-SERRANO, J., SHAPOSHNIKOV, S., SIPINEN, V., SMEETS, K., STOPPER, H., TEIXEIRA, J. P.,

VALDIGLESIAS, V., VALVERDE, M., VAN ACKER, F., VAN SCHOOTEN, F. J., VASQUEZ, M., WENTZEL, J. F., WNUK, M., WOUTERS, A., ŽEGURA, B., ZIKMUND, T., LANGIE, S. A. S. & AZQUETA, A. 2023. Measuring DNA modifications with the comet assay: a compendium of protocols. *Nat Protoc,* 18**,** 929-989.

CRESPI, C. L. & THILLY, W. G. 1984. Assay for gene mutation in a human lymphoblast line, AHH-1, competent for xenobiotic metabolism. *Mutat Res,* 128**,** 221-30.

CROSS, K. P. & PONTING, D. J. 2021. Developing Structure-Activity Relationships for N-Nitrosamine Activity. *Comput Toxicol,* 20.

DE LA MONTE, S. M. & TONG, M. 2009. Mechanisms of nitrosamine-mediated neurodegeneration: potential relevance to sporadic Alzheimer's disease. *J Alzheimers Dis,* 17**,** 817-25.

DEMARINI, D. M., BROCK, K. H., DOERR, C. L. & MOORE, M. M. 1988. Mutagenicity and clastogenicity of proflavin in L5178Y/TK +/- -3.7.2.C cells. *Mutat Res,* 204**,** 323-8.

DOS SANTOS, C. E. M., DORTA, D. J. & DE OLIVEIRA, D. P. 2022. Setting limits for N-nitrosamines in drugs: A defined approach based on read-across and structure-activity relationship for N-nitrosopiperazine impurities. *Regul Toxicol Pharmacol,* 136**,** 105288.

DURKEE, M. S., ABRAHAM, R., CLARK, M. R. & GIGER, M. L. 2021. Artificial Intelligence and Cellular Segmentation in Tissue Microscopy Images. *Am J Pathol,* 191**,** 1693-1701.

EFSA 2022. Guidance on the use of the benchmark dose approach in risk assessment.

EISENBRAND, G., BUETTNER, A., DIEL, P., EPE, B., FÖRST, P., GRUNE, T., HALLER, D., HEINZ, V., HELLWIG, M., HUMPF, H. U., JÄGER, H., KULLING, S., LAMPEN, A., LEIST, M., MALLY, A., MARKO, D., NÖTHLINGS, U., RÖHRDANZ, E., SPRANGER, J., STEINBERG, P., VIETHS, S., WÄTJEN, W. & HENGSTLER, J. G. 2024. Commentary of the SKLM to the EFSA opinion on risk assessment of N-nitrosamines in food. *Arch Toxicol,* 98**,** 1573-1580.

EMA. 2025. *Nitrosamine impurities: guidance for marketing authorisation holders*. [Online]. European Medicines Agency. Available: Questions and answers for marketing authorisation holders/applicants on the CHMP Opinion for the Article 5(3) of Regulation (EC) No 726/2004 referral on nitrosamine impurities in human medicinal products [Accessed 5 August 2025 2025].

EPA, U. S. 2012. Benchmark Dose Technical Guidance. Washington, DC.

EVANS, H. H., HORNG, M. F. & BEER, J. Z. 1986. Lethal and mutagenic effects of radiation and alkylating agents on two strains of mouse L5178Y cells. *Mutat Res,* 161**,** 91-7.

GASSER, S. M. 2026. The Double Face of Base Excision Repair: Preventing and Triggering Double-Strand Breaks. *Bioessays,* 48**,** e70092.

GRAY, R., PETO, R., BRANTOM, P. & GRASSO, P. 1991. Chronic nitrosamine ingestion in 1040 rodents: the effect of the choice of nitrosamine, the species studied, and the age of starting exposure. *Cancer Res,* 51**,** 6470-91.

GRIESS, J. 1984. On a new series of bodies in which nitrogen is substituted for hydrogen. *The Royal Society,* 154.

GUO, M., CHEN, Y., LIN, L., WANG, Y., WANG, A., YUAN, F., WANG, C., WANG, S. & ZHANG, Y. 2022. The Study on the Clinical Phenotype and Function of HPRT1 Gene. *Child Neurol Open,* 9**,** 2329048x221108821.

HASSAN SN, A. F. 2020. The relevance of antibiotic supplements in mammalian cell cultures: Towards a paradigm shift. *Gulhane Medical Journal.,* 62**,** 224-230.

JACOBSON, E. D., KRELL, K., OLEMPSKA-BEER, Z. & BEER, J. Z. 1984. UV-induced mutagenesis at the hypoxanthine-guanine phosphoribosyl transferase locus in two L5178Y mouse lymphoma cell strains with different UV sensitivities. *Mutat Res,* 129**,** 259-67.

JOHNSON, G. E. 2012. Mammalian cell HPRT gene mutation assay: test methods. *Methods Mol Biol,* 817**,** 55-67.

JOHNSON, G. E., SOETEMAN-HERNÁNDEZ, L. G., GOLLAPUDI, B. B., BODGER, O. G., DEARFIELD, K. L., HEFLICH, R. H., HIXON, J. G., LOVELL, D. P., MACGREGOR, J. T.,

POTTENGER, L. H., THOMPSON, C. M., ABRAHAM, L., THYBAUD, V., TANIR, J. Y., ZEIGER, E., VAN BENTHEM, J. & WHITE, P. A. 2014. Derivation of point of departure (PoD) estimates in genetic toxicology studies and their potential applications in risk assessment. *Environ Mol Mutagen,* 55**,** 609-23.

KANG, J. S., WANIBUCHI, H., MORIMURA, K., GONZALEZ, F. J. & FUKUSHIMA, S. 2007. Role of CYP2E1 in diethylnitrosamine-induced hepatocarcinogenesis in vivo. *Cancer Res,* 67**,** 11141-6.

KEOHAVONG, P., XI, L. & GRANT, S. G. 2005. Molecular analysis of mutations in the human HPRT gene. *Methods Mol Biol,* 291**,** 161-70.

LI, X., HE, X., LE, Y., GUO, X., BRYANT, M. S., ATRAKCHI, A. H., MCGOVERN, T. J., DAVIS-BRUNO, K. L., KEIRE, D. A., HEFLICH, R. H. & MEI, N. 2022. Genotoxicity evaluation of nitrosamine impurities using human TK6 cells transduced with cytochrome P450s. *Arch Toxicol,* 96**,** 3077-3089.

LI, Y. & HECHT, S. S. 2022. Metabolic Activation and DNA Interactions of Carcinogenic N-Nitrosamines to Which Humans Are Commonly Exposed. *Int J Mol Sci,* 23.

LIJINSKY, W. 1987. Carcinogenicity and mutagenicity of N-nitroso compounds. *Mol Toxicol,* 1**,** 107-19.

LLOYD, M. & KIDD, D. 2012. The mouse lymphoma assay. *Methods Mol Biol,* 817**,** 35-54.

LOVE, L. A., LIJINSKY, W., KEEFER, L. K. & GARCIA, H. 1977. Chronic oral administration of 1-nitrosopiperazine at high doses to MRC rats. *Z Krebsforsch Klin Onkol Cancer Res Clin Oncol,* 89**,** 69-73.

LYNCH, A. M., HOWE, J., HILDEBRAND, D., HARVEY, J. S., BURMAN, M., HARTE, D. S. G., CHEN, L., KMETT, C., SHI, W., MCHUGH, C. F., PATEL, K. K., JUNNOTULA, V., KENNY, J., HAWORTH, R. & WILLS, J. W. 2024. N-Nitrosodimethylamine investigations in Muta™Mouse define point-of-departure values and demonstrate less-than-additive somatic mutant frequency accumulations. *Mutagenesis,* 39**,** 96-118.

MCGREGOR, D. B., RIACH, C., CATTANACH, P., EDWARDS, I., SHEPHERD, W. & CASPARY, W. J. 1996. Mutagenic responses of L5178Y mouse cells at the tk and hprt loci. *Toxicol In Vitro,* 10**,** 643-7.

MOORE, M. M., PARKER, L., HUSTON, J., HARRINGTON-BROCK, K. & DEARFIELD, K. L. 1991. Comparison of mutagenicity results for nine compounds evaluated at the hgprt locus in the standard and suspension CHO assays. *Mutagenesis,* 6**,** 77-85.

MYHR, B. C. & CASPARY, W. J. 1988. Evaluation of the L5178Y mouse lymphoma cell mutagenesis assay: intralaboratory results for sixty-three coded chemicals tested at Litton Bionetics, Inc. *Environ Mol Mutagen,* 12 Suppl 13**,** 103-94.

OECD. 2016. *Test No. 476: In Vitro Mammalian Cell Gene Mutation Tests using the Hprt and xprt genes*. [Online]. Available: https://www.oecd.org/en/publications/test-no-476-in-vitro-mammalian-cell-gene-mutation-tests-using-the-hprt-and-xprt-genes_9789264264809-en.html#:~:text=The%20in%20vitro%20mammalian%20cell,xanthineguanine%20phosphoribosyl%20transferase%20(XPRT). [Accessed].

PAERATAKUL, U. 1986. Isolation and characterization of mutants at the APRT locus in the L-5178Y TK+/TK− mouse lymphoma cell line. *Mut. Research,* 160**,** 61-69.

PARRY, J. M., PARRY, E. M., JOHNSON, G., QUICK, E. & WATERS, E. M. 2005. The detection of genotoxic activity and the quantitative and qualitative assessment of the consequences of exposures. *Exp Toxicol Pathol,* 57 Suppl 1**,** 205-12.

SHANK, R. C. 1975. Toxicology of N-nitroso compounds. *Toxicol Appl Pharmacol,* 31**,** 361-8.

SLOB, W. & SETZER, R. W. 2014. Shape and steepness of toxicological dose-response relationships of continuous endpoints. *Crit Rev Toxicol,* 44**,** 270-97.

SNODIN, D. J., TREJO-MARTIN, A., PONTING, D. J., SMITH, G. F., CZICH, A., CROSS, K., CUSTER, L., ELLOWAY, J., GREENE, N., KALGUTKAR, A. S., STALFORD, S. A., TENNANT, R. E., VOCK, E., ZALEWSKI, A., ZIEGLER, V. & DOBO, K. L. 2024. Mechanisms of Nitrosamine

Mutagenicity and Their Relationship to Rodent Carcinogenic Potency. *Chem Res Toxicol,* 37**,** 181-198.

SZUMIEL, I., JAWORSKA, A., KAPISZEWSKA, M., JOHN, A., GRADZKA, I. & SOCHANOWICZ, B. 2000. Differential induction of apoptosis in x-irradiated L5178Y sublines bearing p53 mutation. *Radiat Environ Biophys,* 39**,** 33-40.

THOMAS, D. N., WILLS, J. W., BURMAN, M., WILLIAMS, A. N., HARTE, D. S. G., BUCKLEY, R. A., URQUHART, M. W., BRETONNET, A. S., JEFFRIES, B., WHITE, A. T., HARVEY, J. S., HOWE, J. R. & LYNCH, A. M. 2025. Resolution of historically discordant Ames test negative/rodent carcinogenicity positive N-nitrosamines using a sensitive, OECD-aligned design. *Mutagenesis,* 40**,** 116-125.

THOMAS, D. N., WILLS, J. W., TRACEY, H., BALDWIN, S. J., BURMAN, M., WILLIAMS, A. N., HARTE, D. S. G., BUCKLEY, R. A. & LYNCH, A. M. 2024. Ames test study designs for nitrosamine mutagenicity testing: qualitative and quantitative analysis of key assay parameters. *Mutagenesis,* 39**,** 78-95.

THRESHER, A., FOSTER, R., PONTING, D. J., STALFORD, S. A., TENNANT, R. E. & THOMAS, R. 2020. Are all nitrosamines concerning? A review of mutagenicity and carcinogenicity data. *Regul Toxicol Pharmacol,* 116**,** 104749.

TOTH, B., MAGEE, P. N. & SHUBIK, P. 1964. CARCINOGENESIS STUDY WITH DIMETHYLNITROSAMINE ADMINISTERED ORALLY TO ADULT AND SUBCUTANEOUSLY TO NEWBORN BALB-C MICE. *Cancer Res,* 24**,** 1712-21.

TRAN, D. H., KIM, D., KESAVAN, R., BROWN, H., DEY, T., SOFLAEE, M. H., VU, H. S., TASDOGAN, A., GUO, J., BEZWADA, D., AL SAAD, H., CAI, F., SOLMONSON, A., RION, H., CHABATYA, R., MERCHANT, S., MANALES, N. J., TCHEUYAP, V. T., MULKEY, M., MATHEWS, T. P., BRUGAROLAS, J., MORRISON, S. J., ZHU, H., DEBERARDINIS, R. J. & HOXHAJ, G. 2024. De novo and salvage purine synthesis pathways across tissues and tumors. *Cell,* 187**,** 3602-3618.e20.

WANG, J., SAWYER, J. R., CHEN, L., CHEN, T., HONMA, M., MEI, N. & MOORE, M. M. 2009. The mouse lymphoma assay detects recombination, deletion, and aneuploidy. *Toxicol Sci,* 109**,** 96-105.

WILLS, J. W., JOHNSON, G. E., DOAK, S. H., SOETEMAN-HERNÁNDEZ, L. G., SLOB, W. & WHITE, P. A. 2016a. Empirical analysis of BMD metrics in genetic toxicology part I: in vitro analyses to provide robust potency rankings and support MOA determinations. *Mutagenesis,* 31**,** 255-63.

WILLS, J. W., LONG, A. S., JOHNSON, G. E., BEMIS, J. C., DERTINGER, S. D., SLOB, W. & WHITE, P. A. 2016b. Empirical analysis of BMD metrics in genetic toxicology part II: in vivo potency comparisons to promote reductions in the use of experimental animals for genetic toxicity assessment. *Mutagenesis,* 31**,** 265-75.

WILLS, J. W., VERMA, J. R., REES, B. J., HARTE, D. S. G., HAXHIRAJ, Q., BARNES, C. M., BARNES, R., RODRIGUES, M. A., DOAN, M., FILBY, A., HEWITT, R. E., THORNTON, C. A., CRONIN, J. G., KENNY, J. D., BUCKLEY, R., LYNCH, A. M., CARPENTER, A. E., SUMMERS, H. D., JOHNSON, G. E. & REES, P. 2021. Inter-laboratory automation of the in vitro micronucleus assay using imaging flow cytometry and deep learning. *Arch Toxicol,* 95**,** 3101-3115.

WYATT, M. D. & PITTMAN, D. L. 2006. Methylating agents and DNA repair responses: Methylated bases and sources of strand breaks. *Chem Res Toxicol,* 19**,** 1580-94.

ZHOU, F., XIE, Y., WANG, Y., ZHANG, H., WANG, J., LIAO, X. & CHEN, C. 2025. An overview of the formation mechanisms of endogenous and exogenous N-nitrosamines in human diets. *J Environ Sci (China),* 158**,** 527-541.