November 4, 2019

Mr. Tyler Saecho
Office of Environmental Health Hazard Assessment
Proposition 65 Implementation Program
P.O. Box 4010, MS-12B
Sacramento, California 95812-4010

Via electronic submission

Re: Request for Public Comment on Hazard Identification Materials for Acetaminophen

Dear Mr. Saecho,

This information is submitted on behalf of the Consumer Healthcare Products Association ("CHPA") in response to the September 20, 2019 Office of Environmental Health Hazard Assessment's ("OEHHA") notice: Announcement of the Carcinogen Identification Committee Meeting Scheduled for December 5, 2019, Notice of Availability of Hazard Identification Materials for Acetaminophen and Notice of Public Comment Period. CHPA, founded in 1881, is a member-based association representing the leading manufacturers and distributors of non-prescription (or over-the-counter; OTC) medicines and dietary supplements. CHPA appreciates the opportunity to provide comments on the Hazard Identification Materials for Acetaminophen. It is our understanding that our comments will be provided to the Carcinogen Identification Committee (CIC) for their review prior to the December 5, 2019 meeting. We have also requested presentation time on December 5, 2019 and look forward to sharing our perspective with the Committee in person.

We hope the information provided herein will prove helpful to OEHHA and the CIC as they prepare for the December 5, 2019 meeting.

Sincerely,

Barbara A. Kochanowski

Barbara A. Kochanowski, Ph.D.

Sr. Vice President, Regulatory & Scientific Affairs

CHPA CONSUMER HEALTHCARE PRODUCTS ASSOCIATION
founded 1881

An Integrated Weight of Evidence Assessment of the Carcinogenicity Hazard Potential of Acetaminophen

Information for the California Carcinogen Identification Committee

Submitted by the:

Consumer Healthcare Products Association

November 4, 2019

Table of Contents

1 Executive Summary

The objectives of this document are to provide the California Carcinogen Identification Committee (CIC) members with a scientifically rigorous weight of evidence assessment of the available animal carcinogenicity, genotoxicity, epidemiology and mode of action data and provide them with the necessary information to inform their decision on the carcinogenic hazard potential of acetaminophen. **This weight of evidence assessment clearly demonstrates that acetaminophen is not a carcinogenic hazard to animals or humans at any dose level.** Mechanistic studies evaluating therapeutic, supratherapeutic and overdose exposures in animals and humans show that there are cellular protective mechanisms in place that make it implausible for acetaminophen or its reactive metabolites to induce stable genetic damage that would be indicative of a genotoxic or carcinogenic hazard.

High level summaries of the Animal Carcinogenicity, Genotoxicity, Mode of Action and Epidemiology data are presented in the sections that follow. Specific observations related to scientific accuracy and completeness in the Hazard Identification Document (HID) submitted by the Office of Environmental Health Hazard Assessment (OEHHA) are addressed in various places in the text of this document and in detail in the Appendices.

1.1 Epidemiology Studies

Weight of Evidence Assessment: *Epidemiologic studies do not support a causal association between acetaminophen use and cancer and therefore do meet the "clearly shown" standard*

IARC (1999) reviewed the epidemiologic studies of acetaminophen and concluded that there is "inadequate evidence" in humans of carcinogenicity (IARC, 1999). Many additional epidemiologic studies have been published since IARC's 1999 review, totaling over 130 studies that were summarized in the HID, a few of which would be expected to be statistically significant purely by random variation. For most forms of cancer, the results of these studies suggest no alteration in risk associated with acetaminophen use.

For three cancer types - renal cell cancer, liver cancer, and some forms of lymphohematopoietic cancer - there are suggestions in some studies of an increased risk. A limited number of studies for these three cancer types report relative risks (RRs) >1. It is important to note, however, that these same cancers have studies that show no increase in RR. A cautious interpretation of positive findings for these cancers is warranted for several reasons, and those reasons are presented briefly here and in a more detailed summary of the weight of the evidence for each cancer type within the body of this document. This response will lay out a framework for evaluating studies for these cancers of interest, highlighting specific characteristics of the natural

history and the unique nature of acetaminophen that call into question the ability to conclude that use of acetaminophen leads to an increased risk of these cancers.

In reviewing these studies, one must pay attention to the usual issues commonly addressed in well-conducted epidemiologic studies, e.g., the risks of selection bias in clinic- or hospital-based case-control studies and the potential for appropriate control of confounding by important covariates. In addition, with any study of cancer etiology, latency is an issue, i.e., one needs to collect data on exposures that started years before the outcome. Exposure assessment is always difficult in pharmacoepidemiology studies, but is even more of a concern for studies of acetaminophen given that it is available over-the-counter (OTC) in most countries. Electronic health records and claims data would miss most of the exposures, but patient recall is likely to be quite incomplete, especially in a setting where long-term use needs to be quantified (Lewis et al., 2006; West et al., 1997).

Further, because acetaminophen is available both by prescription and over-the-counter and it is used to treat pain and fever, these unique aspects of acetaminophen use further complicate assessment of latent effects like cancer occurrence, because they contribute to several additional sources of bias that arise in other contexts, but can be more severe in this setting, including protopathic bias, channeling bias, and recall bias. These characteristics of acetaminophen use in clinical practice make the potential biases particularly relevant to renal, liver and lymphohematopoietic cancer.

Protopathic bias results from medication use for the treatment of early signs and symptoms of disease prior to the diagnosis of the disease. It is a potential, or even likely, source for bias for acetaminophen because pain or fever can be a sign of undiagnosed cancer and a history of febrile illnesses is a risk factor for some cancers, most notably the lymphohematopoietic cancers. For example, symptoms of lymphoma may include fever. Protopathic bias can be addressed by looking at timing between the initiation of drug use and diagnosis. If the association has a biologic basis, the association should get weaker as the interval between drug initiation and cancer diagnosis shrinks; if protopathic bias is present, then the association should get stronger as that interval shrinks. Unfortunately, few studies measured exposure in a fashion that would enable accounting for timing in this way. However, in one study of lymphohematopoietic cancers (Walter et al., 2011b), dropping diagnoses in the two years after the start of follow-up (when the short-term period of exposure is unlikely to be actively carcinogenic) reduced the relative risk from 1.84 to 1.50, suggesting protopathic bias. Importantly, though, even that estimate of 1.5 is still likely to be biased by residual confounding, as evidenced by one of the experiments we performed, which is described below.

Channeling bias (described by its synonym as “Confounding by Indication” in the HID) is the use of one drug to a greater extent compared to another drug in certain patient populations, in a way that influences the relative risk. This is a large problem when comparing users to non-users, and

can be helped by looking at alternative OTC analgesics as the comparator. However, even in the context of an active comparator, there can be problematic channeling, e.g., acetaminophen being used in patients at higher risk of cancer versus users of nonsteroidal anti-inflammatory agents (NSAIDs). Further, while channeling is a potential concern for all cancers, for certain cancers, patients with specific conditions or at risk of specific conditions, such as chronic renal disease or liver disease, are directed to take acetaminophen and not NSAIDs. This particular form of channeling stems in part from the product labels for NSAIDs, which direct patients to ask a physician before using the drug if they have had a stroke or gastrointestinal bleeding, or have heart disease, liver cirrhosis or renal disease. Empirical evidence that channeling is not just a theoretical concern was reported in a large electronic health records database when those patients receiving a first prescription for acetaminophen were more likely to have a history of myocardial infarction, stroke, renal disease, and gastrointestinal bleeding, relative to those receiving a first prescription of ibuprofen (Weinstein et al., 2017). Patients with end stage renal disease or even less severe chronic kidney disease are at increased risk for renal cancer (Lowrance et al., 2014; van de Pol et al., 2019), so an apparent positive association could arise from the underlying renal disease, not from exposure to acetaminophen. Similarly, individuals with elevated liver enzymes or cirrhosis would be directed by the drug labels and their health care professional to acetaminophen. Because cirrhosis is such a strong risk factor for liver cancer (eighty to ninety percent of liver cancers have a history of cirrhosis (El-Serag, 2012)), channeling bias is a concern for this cancer as well.

Recall bias can occur in the context of case-control studies, when cases with a serious diagnosis may remember exposures and events better than controls, especially if being asked about OTC use, and use in the distant past, where recall is typically poor and incomplete, especially for analgesics. As a result, this may bias the relative risk to show an artificial, positive association. In contrast to these other sources of bias, failure to capture OTC use might, if nondifferential with respect to cancer occurrence, result in bias toward the null.

Finally, another important consideration in this context is the potential for selective publication of positive results, especially for liver and lymphohematopoietic cancers. The HID describes 10 cohorts in detail that can be used to evaluate the association of acetaminophen and various cancers. However, at present there are published data available for lymphohematopoietic cancers from just one cohort study (Walter et al., 2011b) (which reported a positive association with use of acetaminophen). Analogously, there are published data for liver cancer from just one cohort study (Friis et al., 2002)(which reported a relative risk of 1.8 that is not statistically significant). There is a real possibility that the results of the published studies for these particular forms of cancer are not representative of the available results in the other nine cohort studies, were it possible to identify those results as well.

In order to understand the magnitude of the potential biases described here, two experiments were performed in the same UK database used in several studies reviewed in the HID (Kaye et al., 2001; McGlynn et al., 2015; Yang et al., 2016). Specifically, the experiments used a series of negative control outcomes, i.e., outcomes for which we have high confidence there is no association with acetaminophen exposure (also called falsification hypothesis testing). In that same database, using a case-control design, relative risks of the three cancers of interest were estimated. The relative risks for the control outcomes, which should have relative risks of 1.0, were of the same size as the relative risks for the cancer outcomes.

In a second experiment, one of the cohort studies in the review (Walter et al., 2011b) was reproduced and it showed that the high exposure group differs substantially from the non-users, with respect to a number of variables, including exposure to concomitant medications or presence of specific conditions that could also influence cancer risk. These analyses support the possibility that, in the context of these three cancers, if not accounted for, the unique potential sources of bias with acetaminophen could result in more studies with RR > 1 than expected by chance alone.

As a final note, we consider the application of Hill's considerations for determining causality based on epidemiologic studies (Hill, 1965). For renal cancer, in particular, while some studies show an increased relative risk, there are also many studies that do not show elevated risks; thus, renal cancer does not meet the criterion of consistency. In addition, biological plausibility is called into question by the lack of support from the animal studies for the three cancers of interest. A few studies show a dose-response, i.e., an increase in risk with increasing exposure, but other studies show no such increase, and some show a decrease (pointing toward protopathic bias).

Taken together, the epidemiologic data do not support a causal association between acetaminophen use and cancer and therefore do meet the "clearly shown" standard.

1.2 Carcinogenicity Studies in Animals

Weight of Evidence Assessment: *Carcinogenicity studies in animals show no meaningful evidence of the potential for acetaminophen to cause cancer, which strongly weighs against a "clearly shown to cause cancer" finding.*

Supporting Data:

- Fourteen preclinical rodent carcinogenicity studies examined the carcinogenic potential of acetaminophen. These studies evaluated conditions of chronic dosing up to, and above, a maximum tolerated dose (MTD).
- The International Agency for Research on Cancer (IARC) has evaluated all 14 studies and concluded that there is "inadequate evidence" of carcinogenicity in animals.

- In the NTP (1993) cancer bioassay, NTP concluded "no evidence" of carcinogenicity in male and female mice and male rats and "equivocal evidence" in female rats due to an increase in mononuclear cell lymphomas. This tumor type occurs spontaneously in this strain of rats with a highly variable background incidence; for this reason and because no increase was seen in male rats, NTP considered this finding as "equivocal evidence." NTP no longer uses this strain of rats (F344/N) for carcinogenicity studies in large part due to concerns about the relevance of this endpoint.
- Studies by other investigators (Amo and Matsuyama, 1985; Hagiwara and Ward, 1986; Hiraga and Fujii, 1985; Johansson et al., 1974) reported no significant increases in tumors in any organ systems in the acetaminophen-treated versus control animals.
- Increases in tumors were reported in studies in mice and rats by the same group of investigators (Flaks and Flaks, 1983; Flaks et al., 1985), but these studies are seriously flawed, and do not represent "scientifically valid testing according to generally accepted principles" for many reasons, as detailed in Section 4 (Carcinogenicity Studies in Animals). In addition, these study results have not been confirmed in other studies in mice and rats, including the NTP cancer bioassay.
- Six tumor promotion studies and two tumor initiation studies in animals with compromised liver function did not show meaningful evidence of tumor initiation or promotion; all of these studies were reviewed by IARC when it concluded "inadequate evidence" in experimental animals.
- Consistent with the reviews and conclusions of multiple health authorities, including IARC and the United States Food & Drug Administration (FDA), the weight-of-evidence assessment of the animal carcinogenicity studies clearly demonstrates an absence of a carcinogenic hazard potential for acetaminophen.

A more detailed assessment of the 14 studies along with corresponding data tables can be found in Section 4.

1.3 Genetic Toxicology Studies

Weight of Evidence Assessment: *Data from genetic toxicology studies show no meaningful evidence of potential for acetaminophen to cause genetic toxicity that could lead to cancer and therefore weigh against a "clearly shown to cause" finding.*

Supporting Data:

- The genotoxicity data related to acetaminophen has been extensively reviewed and analyzed by a number of research groups and institutions, including IARC (1990, 1999), NTP (1993) and Bergman et al. (1996). Data has been generated in more than 70 genetic toxicology studies with varying degrees of relevance to humans, quality and conformance to accepted standards, and therefore, the data requires a Weight of Evidence (WoE) approach.
- Acetaminophen showed no evidence of induction of point or gene mutations *in vitro* in bacterial and mammalian cell systems or *in vivo*.

- In studies that evaluated toxicity, acetaminophen also demonstrated no evidence of clastogenicity (micronucleus test and chromosomal aberration assay) in reliable, well-controlled test systems at non-cytotoxic concentrations up to 1 mM *in vitro* or at non-toxic doses *in vivo*. In the *in vitro* and *in vivo* test systems, clastogenic effects were only observed in unstable, p53-deficient cell systems or at toxic and/or excessively high concentrations that adversely affect cellular processes (e.g. mitochondrial respiration) and cause cytotoxicity and, as a consequence, are not expected to produce stable, genetic damage in humans.
- There is no clear evidence that acetaminophen causes DNA damage (Unscheduled DNA Synthesis and COMET) in the absence of toxicity. In well-controlled human clinical studies, there was no meaningful evidence of chromosomal damage, including following multiple dosing at therapeutic doses and in an acute overdose scenario.

In conclusion, acetaminophen overwhelmingly produces negative results (i.e. is not a genotoxic hazard) in reliable, robust high weight genotoxicity studies. Some genotoxic effects (clastogenicity) are seen in moderate weight studies, particularly in cell types susceptible to misleading positive results. However, when considering data from relevant, robust test systems, clastogenic effects are only seen at excessively high or under cytotoxic conditions and associated with cell lethality. Therefore, from all available data it is not plausible that acetaminophen induces stable, genetic damage that would be indicative of a genotoxic or carcinogenic hazard in humans.

A more detailed assessment of these studies along with corresponding data tables can be found in Section 5.

1.4 Mode of Action Studies (Pathways, Pharmacology and KCC Considerations)

Weight of Evidence Assessment: *Mechanistic data clearly demonstrate that acetaminophen causes cellular toxicity before it exerts any DNA effects and therefore does not initiate or promote cancer; thus, this data set actually is reassuring and weighs against a finding that acetaminophen has been "clearly shown to cause cancer."*

Supporting Data:

- Based on its mode of action, acetaminophen only causes adverse DNA effects in relevant, well-controlled test systems at exposures that result in cell death, which preclude it from having potential to cause any carcinogenic effects.
- Acetaminophen's formation of reactive metabolites does not have the potential to be a carcinogenic hazard. Careful examination of the known pathways for acetaminophen metabolism and disposition under therapeutic, supratherapeutic and overdose conditions demonstrate that potential for oxidative stress and DNA effects resulting from formation of a reactive metabolite occur in a very precise sequence that result in cellular toxicity before it can become a carcinogenic hazard (Table 1).

- From a mode of action (MoA) perspective, all available data from relevant, well-controlled tests support that, at therapeutic doses, acetaminophen has no effects on nuclear DNA.
- Following supratherapeutic doses or on overdose, acetaminophen only causes DNA damage at exposures that result in cell death, making it implausible for acetaminophen to induce stable, genetic damage that would be indicative of a genotoxic or carcinogenic hazard in humans.
- The mechanism of DNA damage is endonuclease-mediated DNA fragmentation, which is set up to degrade the nucleus. This is not repairable and is fundamentally different to a potential DNA modification that could give rise to a cancer cell.
- Furthermore, there is no meaningful evidence that acetaminophen has the potential to cause cancer by non-genotoxic mechanisms and there are some studies showing that it may have anti-proliferative effects on tumors.
- The activity observed in High Throughput Screening data and results from the ToxCast/Tox21 assays do not show any effects supporting carcinogenic potential.
- Simulations have been performed to evaluate the potential for acetaminophen to be a hazard in patient sub-populations and in overdose patients using a Quantitative Systems Toxicology Platform called DILIsym that has been developed and validated using acetaminophen. These simulations support that there is also not a carcinogenicity hazard in patients with susceptibility for liver injury. The methodology and results of these simulations can be found in a separate supplementary document that has been made available to the CIC.

Table 1: Summary of the effects of acetaminophen on different hepatocellular parameters under therapeutic, supratherapeutic and acute overdose conditions (Bajt et al., 2006; Cover et al., 2005b; Heard et al., 2011; Hu et al., 1993; Kang et al., in press; McGill and Jaeschke, 2013; McGill et al., 2013; McGill et al., 2012b; McGill et al., 2011; Xie et al., 2015a; Xie et al., 2014).

Hepatocellular Parameter	Therapeutic < 4 g/day	Supratherapeutic > 4 – 10 g/day	Acute Overdose > 10 -15 g
1. Glutathione (GSH) Depletion	No	Isolated hepatocytes*	Yes
2. Cytosolic Protein SH Group Depletion	No	Isolated hepatocytes*	Yes
3. Mitochondrial Adduct Formation	No	Isolated hepatocytes*	Yes
4. Mitochondrial Oxidative/Nitrosative Stress	No	Isolated hepatocytes*	Yes
5. JNK Pathway Activated	No	Isolated hepatocytes*	Yes
6. Amplification of the Mitochondrial Oxidative/Nitrosative Stress	No	Isolated hepatocytes*	Yes
7. Loss of Mitochondrial Membrane Potential → ATP↓	No	Isolated hepatocytes*	Yes
8. Release of Endonucleases from mitochondria	No	Isolated hepatocytes*	Yes
9. Translocation of Endonucleases to the Nucleus	No	Isolated hepatocytes*	Yes

Hepatocellular Parameter	Therapeutic < 4 g/day	Supratherapeutic > 4 – 10 g/day	Acute Overdose > 10 -15 g
10. Nuclear DNA fragmentation and Cell Death	No	Isolated hepatocytes*	Yes

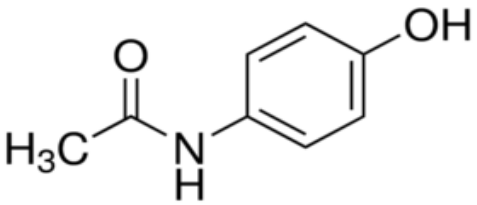
*At supratherapeutic doses, there can be isolated cells in the centrilobular region of the liver in preclinical models and humans with cellular effects seen in overdose which may result in isolated hepatic cell death that do not have any clinical relevance.

In conclusion, the evidence does not support a finding that acetaminophen has "been clearly shown through scientifically valid testing, according to generally accepted principles to cause cancer", as required by Proposition 65. A more detailed assessment of these studies can be found in Section 6.

2 Introduction and Background

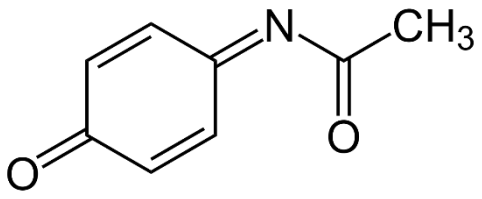
Acetaminophen (4-hydroxyacetanilide, or N-acetyl-*p*-aminophenol, or APAP; CAS No. 103-90-2) is an antipyretic and analgesic drug that was first prepared as long ago as 1877 and was introduced worldwide in the 1950s (Figure 1). It is widely purchased over the counter (usually as paracetamol, Panadol or Tylenol), but can also be used on a prescription basis for treatment of chronic pain. The structure of acetaminophen is shown in Figure 1.

Figure 1: Structure of acetaminophen (paracetamol)



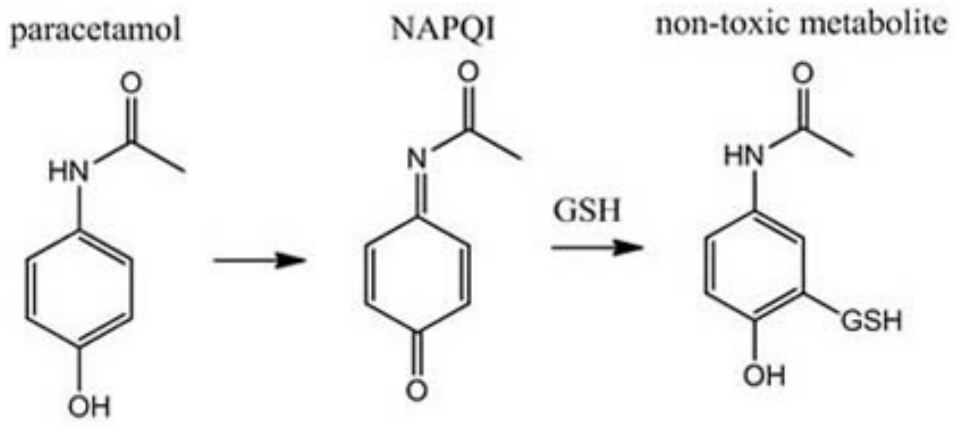
Acetaminophen has been determined by health authorities around the world to be safe at recommended daily doses of 4 g/day and less, but higher doses may lead to hepatotoxicity, and possibly liver failure. In humans, the threshold for acute liver damage is approximately 250 mg/kg for an acute overdose (i.e. 15 g for a 60 kg adult) with 350 mg/kg usually associated with severe hepatotoxicity (Thomas, 1993). The hepatotoxic effects of acetaminophen require that it be metabolically activated. The major detoxification pathways for acetaminophen are the formation of sulfate and glucuronide conjugates with the parent compound. Whereas the sulfation pathway is saturated after an overdose, glucuronidation has proven to be not saturated even after severe overdoses (Xie et al., 2015a). However, acetaminophen can be converted to a reactive electrophile and oxidizing agent, N-acetyl-p-benzoquinone imine or NAPQI (Dahlin et al., 1984; Guengerich and Liebler, 1985) by liver microsomal cytochrome P45Os (2E1, 1A2. 3A4) (Raucy et al., 1989; Thummel et al., 1993). The structure of NAPQI is shown in Figure 2.

Figure 2: Structure of N-acetyl-p-benzoquinone imine (NAPQI)



The pathway for conversion of acetaminophen to NAPQI and the detoxification of NAPQI by conjugation with glutathione is shown in Figure 3.

Figure 3: Metabolic pathway for NAPQI formation and detoxification by conjugation with glutathione



Under conditions of overdose where liver toxicity is induced (e.g. a human taking >15 g/day, or 250 mg/kg for a 60 kg person (Thomas, 1993)), the sulfation conjugation pathway is saturated but the glucuronidation pathway is substantially enhanced (Xie et al., 2015a), whereupon NAPQI depletes glutathione (Mitchell et al., 1973), reacts with cellular macromolecules (primarily to soft nucleophilic sites in proteins - SH groups), and initiates cell death due to mitochondrial damage, increased oxygen/nitrogen stress and DNA fragmentation. The associated molecular signaling mechanisms of the cell death and transcriptomics that accompanies this pathway has been reviewed (Chang et al., 2004; Ramachandran and Jaeschke, 2018, 2019; Stamper, 2015). Overdosage may also lead to acute renal tubular necrosis, which is also believed to involve the formation of a reactive intermediate (probably NAPQI), formed via cortical cytochrome P450-mediated oxidation (Hoivik et al., 1995; Hu et al., 1993), although acetaminophen-induced nephrotoxicity has been much less studied than hepatotoxicity.

It is notable that there are marked species differences in acetaminophen-induced hepatotoxicity, with mice being much more sensitive than rats (Davis et al., 1974). The oral LD_{50} in mice is 338 mg/kg, whereas in rats it is 1944 mg/kg. Thus, doses which far exceed the LD_{50} in mice cause only

minimal necrosis in rat liver (McGill et al., 2012b). These differences are due to differences in the rate of metabolism of acetaminophen to NAPQI (Blair et al., 1980; Tee et al., 1987) and mitochondrial dysfunction (McGill et al., 2012b). The relative sensitivity of freshly isolated hepatocytes from mouse, rat and hamster reflected the hepatotoxicity seen *in vivo.* The sensitivity of primary human hepatocytes to acetaminophen-induced cell injury was similar to mouse hepatocytes. However, the time course of cell death was delayed compared to mouse hepatocytes but was comparable to the development of liver injury in overdose patients in vivo (Xie et al., 2014). Thus, toxic effects (and any genotoxicity resulting from such toxicity) would be expected at similar doses in mice and in humans. The rat is generally considered a poor model for the human pathophysiology (McGill and Jaeschke, 2019; McGill et al., 2012a).

3 Epidemiology Studies

Acetaminophen and Cancer Occurrence: Epidemiologic studies do not support a causal association between acetaminophen use and cancer and therefore do meet the "clearly shown" standard

IARC (1999) reviewed the epidemiologic studies of acetaminophen and concluded that there is "inadequate evidence" in humans of carcinogenicity (IARC, 1999). Many additional epidemiologic studies have been published since IARC's 1999 review. As noted in the HID, however, specific characteristics of acetaminophen use make it challenging to accurately assess any cancer risk through generally accepted epidemiologic methods. Our comments elaborate on these challenges and provide an evaluation of the existing epidemiologic data that considers specific biases and methodological issues in assessing the likelihood of causality. We conclude that the current evidence is not sufficient to establish an association between acetaminophen use and cancer; and, it does not demonstrate what would be necessary to draw a "clearly shown to cause cancer" conclusion.

3.1 Important Methodological Considerations in Evaluating Acetaminophen Studies

3.1.1 Specific Biases and Methodological Issues Related to Acetaminophen Use and Cancer Occurrence

In addition to the usual issues commonly addressed in well-conducted epidemiologic studies, there are specific challenges associated with the unique characteristics of acetaminophen use that must be considered when evaluating epidemiologic evidence, including:

(1) Channeling,

(2) Protopathic Bias, and

(3) Exposure Measurement and Recall Bias.

3.1.2 Channeling Bias

Channeling bias (described by its synonym as "Confounding by Indication" in the HID) is the use of one drug to a greater extent compared to another drug in certain patient populations, in a way that influences the relative risk.

Channeling bias is a critical issue for acetaminophen in particular as it is an important pain relief alternative to nonsteroidal anti-inflammatory drugs (NSAIDs). For example, individuals with coexisting conditions recognized as contraindications/warnings for NSAIDs (e.g., gastrointestinal (GI) bleeding and stomach issues) will be channeled to acetaminophen use.

Label warnings for non-prescription NSAIDs (US-based aspirin, ibuprofen and naproxen) direct consumers to ask a doctor before use if they have various medical conditions, including stomach bleeding, liver cirrhosis, or kidney disease. These conditions, which channel patients toward acetaminophen, can be associated with increased risk of cancer. For example, the HID noted that patients with liver disease (and at higher risk for liver cancer) may be channeled toward acetaminophen. Channeling can also occur when a patient with renal insufficiency is recommended to take acetaminophen for pain relief as the patient is contraindicated for NSAIDs. As a consequence, an increased number of patients with chronic renal insufficiency may take acetaminophen versus other drugs, and patients with chronic renal insufficiency have a greater risk for renal cancer (Lowrance et al., 2014; Suzuki et al., 2016), not due to exposure to acetaminophen. This may influence the relative risk to show an artificial, positive association. Many additional examples of other types of cancer affected by channeling are shown in Table 2. Factors leading to channeling bias in acetaminophen use.

Evidence of channeling bias has been reported in the prescribing of acetaminophen versus ibuprofen in an UK electronic health records database (Weinstein et al., 2017). Those with a prescription for acetaminophen were more likely to have a history of renal disease at the time of prescription (7.4%) compared to those who had a prescription for ibuprofen (2.8%). This example showing the disproportionate use of acetaminophen in patients with chronic renal disease, a risk factor for renal cancer, highlights the impact channeling bias can have on studies of acetaminophen and cancer by creating an artificial, positive association between acetaminophen and renal cancer, which is instead due to channeling to chronic renal disease patients.

Lastly, channeling also occurs when more vulnerable populations (e.g., the elderly) and those with chronic conditions, who are being seen regularly by a physician, are more likely to be prescribed acetaminophen than younger, healthier populations. This is channeling based on severity of illness and these patients are more likely to be diagnosed with cancer, based on age and/or comorbidities alone.

Table 2. Factors leading to channeling bias in acetaminophen use

Reason for being channeled to/ recommended acetaminophen (i.e., contraindications for NSAIDs)	**Cancer risk or association**
Chronic kidney insufficiency	Renal/urinary system
Rheumatoid arthritis (RA): use of disease modifying antirheumatic drugs (DMARDs), steroids and other RA medications	Lymphoma
GI bleeding risk	Stomach Cancer
Cigarette smoker (at risk of GI bleeding)	Kidney and bladder, lung, head, neck, pancreas, esophagus, stomach, cervical
Type 2 Diabetes	Liver
Liver disease including HBV/HCV, cirrhosis, non-alcoholic fatty liver disease (NAFLD), metabolic syndrome, metabolic disorders, obesity	Liver

3.1.3 Protopathic Bias

Acetaminophen is used for pain and pain is a symptom for many cancers. Protopathic bias, the treatment of early signs and symptoms of disease prior to diagnosis, is a potential source for bias for acetaminophen because pain or fever can be a sign of undiagnosed cancer and a history of febrile illnesses is a risk factor for some cancers. If not adequately controlled, this bias may influence the relative risk to show an artificial, positive association. Several studies mention protopathic bias and perform at least modest sensitivity analyses to try to adjust for this form of bias. The ability to adjust for this bias is dependent on how exposure is captured, measured, and analyzed. Protopathic bias can preferably be addressed by looking at timing between the initiation of drug use and diagnosis. If the association has a biologic basis, the association should get weaker as the interval between drug initiation and cancer diagnosis shrinks; if protopathic bias is present, then the association should get stronger as that interval shrinks. Unfortunately, few studies measured exposure in a fashion which would enable accounting for timing in this way.

3.1.4 Exposure Measurement and Recall Bias

Estimates of exposure to acetaminophen have major limitations. Acetaminophen is available as both over-the-counter (OTC) and prescription (Rx) formulations. Some epidemiologic studies estimate exposure based on prescription records, which do not account for OTC exposure. As noted in the HID, even prescription records are not a "perfect measure of exposure because a prescription does not necessarily guarantee that the patient took the medication."

Studies using electronic claims records fail to capture OTC use since medication use is identified from pharmacy records of a prescription being filled (Lewis et al., 2006; West et al., 1997; West et al., 1995). In addition, most pain medications are taken on an as-needed basis. Capture of OTC acetaminophen use therefore relies on self-reporting and individual recall since electronic medical record and claims databases do not capture OTC use.

However, case-control studies that rely on patients accurately remembering and reporting medication use during a specific time period can suffer from recall bias. Recall bias can occur when cases with a recent diagnosis may over report or remember better than controls, especially if being asked about use over multiple years (e.g.,10 or more years in the past); as a result, this bias may influence the relative risk to show an artificial, positive association. Self-reported use also requires users to have accurate knowledge of which medications contain acetaminophen.

Patterns of use (e.g., infrequent use on a consistent basis, episodic short-term frequent use and daily use with a low or high frequency) can vary greatly over the full duration of an epidemiologic study, creating significant challenges for quantifying cumulative acetaminophen exposure. Frequent subject recall during a prospective study may allow for more precise quantification of exposure, including an estimate of the total cumulative dose over time.

The HID (p. 17) states that there should be no recall bias since acetaminophen is not suspected of causing cancer. To the contrary, recall bias can occur whether specific exposures are suspected or not. Memory is enhanced for all events that might have played a role in a recent serious diagnosis like cancer (Rothman et al., 2008).

Hence, recall bias is a very strong concern with case-control study designs in which exposure assessment is determined by subject recall after the cancer has occurred. If daily exposure is relatively constant over time (e.g., cigarette smoking) then determining exposure by subject recall may be a valid measure. However, when exposure is episodic or infrequent over the course of many years (typically the case with acetaminophen), the validity of basing cumulative exposure upon recall is questionable unless recall is assessed frequently throughout that period of time.

Finally, the accurate study of most cancer occurrence should include exposures over a 20-year period consistent with the latency period of the cancer. For the reasons detailed above, the estimated measurement of acetaminophen exposure is of questionable validity in most studies even over relatively short periods. Estimates of exposure to acetaminophen over a 20-year period are tenuous at best.

3.1.5 Summary of Methodologic Considerations

In summary, there are many challenges in conducting a scientifically valid epidemiologic study of acetaminophen and cancer. Rigorous study designs and analytic methods are required to appropriately study the etiological association between acetaminophen exposure and cancer incidence. Such rigorously designed studies would need the following characteristics:

- Adequate control for channeling, protopathic bias, and recall bias
- Robust data on both OTC and prescription acetaminophen use (i.e. frequency, dose, indication).
- Analysis of patient attributes that may be linked with cancer, including body weight, smoking, alcohol use, comorbid conditions, and medical history.
- Quantification of exposure that is defined by protocol before data analysis.
- Multiple measures of cumulative use with a pre-specified primary measure.
- Time-to-event analysis to allow for analysis of time duration since first exposure.

3.2 Study Characteristics Essential to Determining Causality

The interpretation of study findings by cancer site includes evaluating the likelihood that the results reflect biases in design or conduct, confounding, chance, or the role of causality.

The following review of studies by cancer site with respective forest plots incorporates an objective assessment of three key study characteristics that must be present to account for potential sources of bias and confounding and reflect overall study quality (i.e. study validity). Definitions of key study characteristics that were included in this review and in the forest plots are:

- Adjustment for channeling: This is true, at least to some degree, if the study included a propensity score or other method to balance the groups being compared at baseline. Another approach would be to use an active comparator that is another analgesic (e.g., ibuprofen or aspirin) rather than nonuse of acetaminophen while also then controlling analytically for the different indications and contraindications for use of acetaminophen vs NSAIDs.
- Protopathic bias analyzed or accounted for: In case-control studies, this is true, at least to some degree, if the study disregarded exposure closest in time prior to diagnosis of cancer, if the disregarded time was sufficient to be meaningful for the time course of pain associated with the cancer being studied. For cohort studies, this is true if exposure time and cancers diagnosed in the appropriate time after start of follow-up were excluded. The time excluded should reflect the cancer type and the time-course of symptoms that may precede it. Ideally, if biological cause, one should see a decreased risk as initiation of therapy approaches the date of cancer diagnosis.
- Exposure data collected without reliance on subject recall: This is true if exposure collection was based solely on electronic records or databases. It is not the case if exposure collection relied on subject recall from interview or questionnaire. Of course, as noted above, one also needs to collect both OTC and Rx use of NSAIDs and the former is very incomplete in most electronic records or claims databases.

Forest plots were also created for visual ease of interpretation of the various relative risks (RR) and cancer outcomes. The forest plots include one-point estimate and 95% confidence interval for each study for each cancer type as well as columns displaying the presence (✓) or absence (X) of key study characteristics. The estimate in most cases is the RR of any acetaminophen use versus no acetaminophen use or nonuse of acetaminophen. For studies that did not provide a RR forever vs never use, either regular use or the highest exposure category use was used.

3.3 Analysis of Epidemiologic Evidence by Cancer Site: Urinary System, Lymphohematopoietic Neoplasms and Liver Cancer

3.3.1 Urinary System: Urinary Tract, Renal, and Bladder

Acetaminophen is the active metabolite of phenacetin, a drug that was taken off the market in the US in 1983 due to an association with cancer of the renal pelvis (48FR 45466). Early studies examining acetaminophen use in connection with urinary system cancers (n=22, see Section 8.2 for the list) included assessments of phenacetin, many without explicitly or adequately accounting for phenacetin as a source of confounding. As a result, some of these studies will have an artificially increased RR.

(i) Urinary Tract Cancers

Assessment of Evidence: The weight of the evidence does not support a conclusion that acetaminophen is clearly shown to cause urinary tract cancers.

A total of 2 cohort studies and 5 case-control studies reported on acetaminophen use and urinary tract cancers. The studies included in this section are cancers of urinary origin, but the specific organ was not specified. Some studies specified transitional cell carcinomas or urothelial cancers which includes bladder, ureter or renal pelvis. Where mentioned, the sites included urinary tract, ureter, and renal pelvis cancer cases. (Figure 4)

Assessments of key study characteristics show none of these studies adequately account for critical sources of biases/methodological issues, thus making it difficult to establish whether the observed results represent a causal association or are due to measurement/design issues. Furthermore, the studies addressing the most methodological issues reported no increased risk (Friis et al., 2002; Linet et al., 1995; Rosenberg et al., 1998; Walter et al., 2011a). Notably, the study reporting the highest risk (a 2-fold increased risk) accounted for the fewest issues (McCredie and Stewart, 1988). A more detailed review of each study is provided in Section 8.1.1.

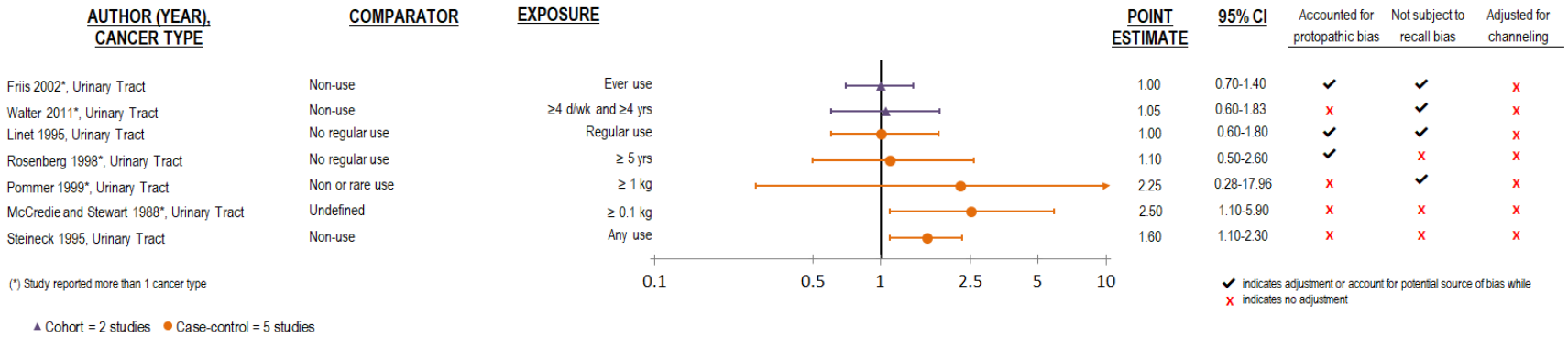
- **McCredie and Stewart** assessed the RR for both >0.1 kg and >1.0 kg of lifetime acetaminophen use in a case-control study and while the OR was increased for the lower exposure group, it was not for the higher exposure group. At least in part,

because of this inconsistency, the authors urged readers to be cautious regarding this outcome. (McCredie and Stewart, 1988)

- Similarly, in the case-control study by **Steineck** et al 1995, the authors characterized the observed RR of 1.6 as moderate and as potentially subject to confounding. And, they further state: ”However, it might be that conventional epidemiological methods and a moderately sized study are too crude to delineate the association, if it exists.” (Steineck et al., 1995)

Given that most of the studies (both cohort studies and 3 out of 5 case-control studies) did not report an increased RR, and the cited methodological limitations above, it cannot be concluded that acetaminophen use is clearly shown to cause increased risk for urinary tract cancer.

Figure 4. Forest plot: urinary tract cancers



(ii) Renal Cancer

Assessment of Evidence: The weight of the evidence does not support a conclusion that acetaminophen is clearly shown to cause renal cancer.

Individuals with renal disease (end stage renal disease and chronic kidney disease, specifically) and are at higher risk of renal cancer (Lowrance et al., 2014) and are also more likely to take acetaminophen for pain (i.e., channeling bias), thus artificially inflating the RR estimates. Note also that renal cancer can take years to become clinically evident, with the risk of pain from as-yet-undiagnosed renal cancer, leading to an association due to protopathic bias.

The studies on renal cancer include renal cell carcinoma and renal pelvis cancer; early reports showed that phenacetin was associated with cancers of the renal pelvis, and so this site is presented separately from renal cell carcinoma. A total of 4 cohort studies and 17 case-control studies assessed the association between acetaminophen use and renal cancer. Five case-control studies assessed the association for cancer of the renal pelvis. Most of the studies on renal cell carcinoma (3 of the 4 cohort studies, 8 of the 12 case-control studies) did not show an increased RR (Figure 5); similarly, all of the renal pelvis cancer studies (5 case-control studies) did not show an increased RR. (Figure 6)

Assessments of key study characteristics show none of these studies adequately account for critical issues, thus making it difficult to establish whether the observed results represent a causal association or are due to methodological errors in measurement/design. Furthermore, the studies addressing the most methodological issues, report no increased risk (3 cohort (Chow et al., 1994; Friis et al., 2002; Walter et al., 2011a); 1 case-control (Rosenberg et al., 1998)). Three additional case control studies addressing some issues reported an increased risk; however, these RRs are likely to be artificially inflated due to lack of proper confounder adjustment and bias issues (Derby and Jick, 1996; Kaye et al., 2001; McCredie et al., 1993). One cohort study reporting an increased risk, but relied on individual recall and used little adjustment for confounding. A more detailed review of each study is provided in Section 8.1.1.

- The increased RR observed in the study by Derby and Jick 1996 had a magnitude >2.0, but no confounders were adjusted for to arrive at the effect estimate (Derby and Jick, 1996).
- Karami et al 2016 reported a positive association between acetaminophen use and increased risk for renal cancer. However, the study design did not include sufficient follow-up time to allow for cancer latency and was so short that it risked solely observing protopathic bias. Furthermore, the RR did not increase for increasing duration. (Karami et al., 2016)
- Gago-Dominguez et al 1999 was a case-control study that reported a positive association for regular use of acetaminophen, as well as aspirin, phenacetin and non-

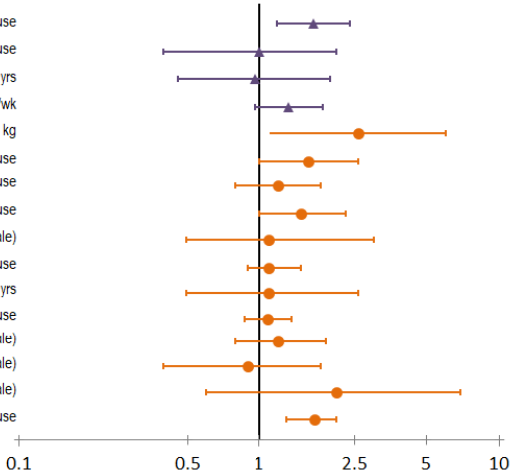
aspirin NSAIDs. Trends in use were also increased for all analgesics examined. It was also subject to possible recall bias. The increased risks for all analgesics are not consistent with biologic plausibility, but are consistent with bias, and thus calls into question the validity of the findings. (Gago-Dominguez et al., 1999)

- Although McCredie et al 1993 was a case-control study with borderline increased RR for renal cell carcinoma. This study also reported an increased risk in those with allergic asthma but not among rheumatoid arthritis patients, both of which were not expected. The dose-response trend for acetaminophen was not increased. These unexpected findings are reason to question the biologic meaningfulness of the overall findings. (McCredie et al., 1993)
- Kaye et al 2001 assessed only a short period of exposure (1-5 years prior to the index date). There may also be residual confounding because only a few covariates were controlled. (Kaye et al., 2001)

After a review of all renal cancer studies, it cannot be concluded that acetaminophen use is associated with an increased risk for renal cancer.

Figure 5. Forest plot: renal cancer

AUTHOR (YEAR), CANCER TYPE	COMPARATOR	EXPOSURE	POINT ESTIMATE	95% CI	Accounted for protopathic bias	Not subject to recall bias	Adjusted for channeling
Karami 2016, Kidney (RCC)	No regular use	Regular use	1.68	1.19-2.39	✓	✓	x
Friis 2002*, Kidney (RCC)	Non-use	Ever use	1.00	0.40-2.10	✓	✓	x
Walter 2011*, Kidney (RCC)	Non-use	≥4 d/wk and ≥4 yrs	0.96	0.46-1.98	x	✓	x
Cho 2011, Kidney (RCC)	No regular use	≥2/wk	1.32	0.96-1.84	x	✓	x
Derby and Jick 1996*, Kidney (RCC)	Non-use	≥ 1 kg	2.60	1.10-6.00	✓	✓	x
Kaye 2001*, Kidney (RCC)	Non-use	Ever use	1.60	1.00-2.60	✓	✓	x
McCredie 1988, Kidney (RCC)	Non-user	Regular use	1.20	0.80-1.80	✓	x	x
McCredie 1993*, Kidney (RCC)	Non-use	Any use	1.50	1.00-2.30	✓	x	x
Mellemgaard 1994, Kidney (RCC)	Never user	Any use (male)	1.10	0.50-3.00	✓	x	x
McCredie 1995, Kidney (RCC)	Never or irregular use	Regular use	1.10	0.90-1.50	✓	x	x
Rosenberg 1998*, Kidney (RCC)	No regular use	≥ 5 yrs	1.10	0.50-2.60	✓	x	x
Karami 2016, Kidney (RCC)	No regular use	Regular use	1.09	0.87-1.37	✓	x	x
McLaughlin 1985*, Kidney (RCC)	Never use	Ever use (female)	1.20	0.80-1.90	x	x	x
Kreiger 1993, Kidney (RCC)	Never or irregular use	Any use (male)	0.90	0.40-1.80	x	x	x
Chow 1994, Kidney (RCC)	Non-use	Regular use (female)	2.10	0.60-6.90	x	x	x
Gago-Dominguez 1999, Kidney (RCC)	Non/irregular use	Regular use	1.70	1.30-2.10	x	x	x



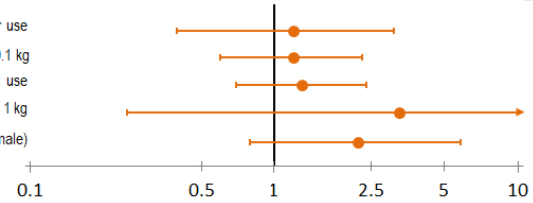
(*) Study reported more than 1 cancer type

▲ Cohort = 4 studies ● Case-control = 12 studies

✓ indicates adjustment or account for potential source of bias while
x indicates no adjustment

Figure 6. Forest plot: renal pelvis cancer

AUTHOR (YEAR), CANCER TYPE	COMPARATOR	EXPOSURE	POINT ESTIMATE	95% CI	Accounted for protopathic bias	Not subject to recall bias	Adjusted for channeling
Kaye 2001*, Kidney (Renal pelvis)	Non-use	Ever use	1.20	0.40-3.10	✓	✓	x
McCredie and Stewart 1988*, Kidney (Renal pelvis)	Undefined	≥ 0.1 kg	1.20	0.60-2.30	✓	x	x
McCredie 1993*, Kidney (Renal pelvis)	Non-use	Any use	1.30	0.70-2.40	✓	x	x
Pommer 1999*, Kidney (Renal pelvis)	Non or rare use	≥ 1 kg	3.27	0.25-43.02	x	✓	x
McLaughlin 1985*, Kidney (Renal pelvis)	Never use	Ever use (female)	2.20	0.80-5.80	x	x	x



(*) Study reported more than 1 cancer type

● Case-control = 5 studies

✓ indicates adjustment or account for potential source of bias while
x indicates no adjustment

(iii) Bladder Cancer

Assessment of Evidence: The weight of the evidence does not support a conclusion that acetaminophen is clearly shown to cause bladder cancer.

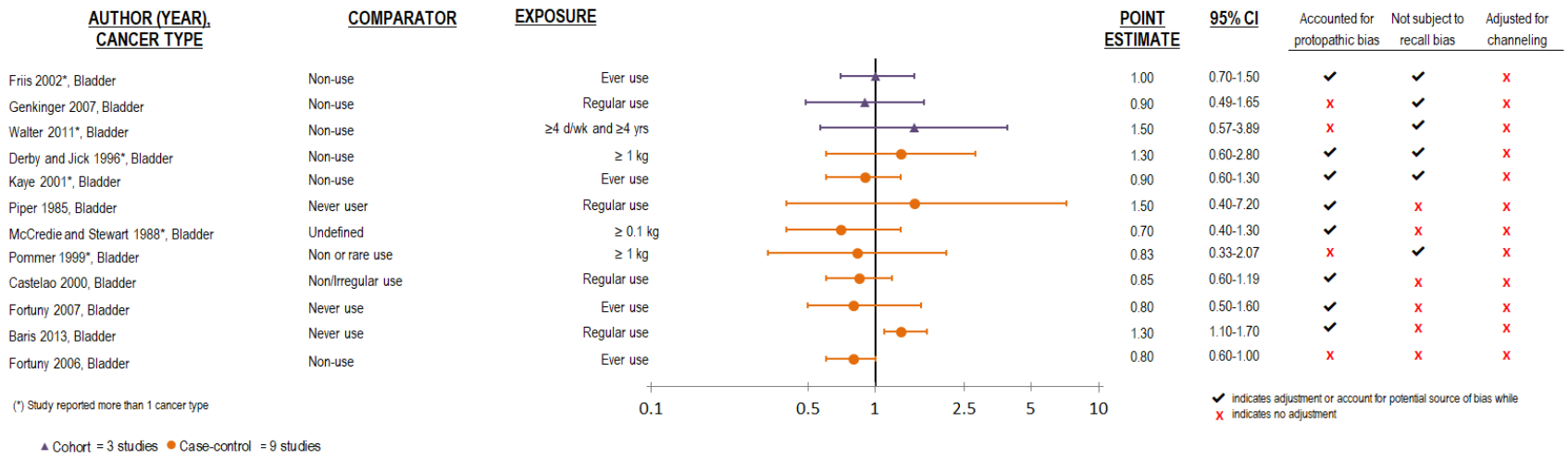
A total of 3 cohort studies and 9 case-control studies assessed the association between acetaminophen use and bladder cancer (Figure 7); almost all of the studies (8 of the 9 case-control studies; all 3 of the cohort studies) did not show an increase in RR for bladder cancer.

Assessments of key study characteristics show none of these studies adequately account for critical methodological issues, thus making it difficult to establish whether the observed results represent a causal association or are due to methodological errors in measurement/design. A more detailed review of each study is provided in Section 8.1.1.

Although almost all studies on bladder cancer reported no association (11 out of 12 studies), one case-control study (Baris et al., 2013) reported an increase for regular use of acetaminophen; however, the trends in the ORs with increasing cumulative lifetime exposure were flat, weakening the evidence for causal association.

Given the limitations of the studies above and that almost all studies did not show an increase in RR, it cannot be concluded that acetaminophen is clearly shown to cause bladder cancer.

Figure 7. Forest plot bladder cancer



3.3.2 Lymphohematopoietic Neoplasms

Assessment of Evidence: The weight of the evidence does not support a conclusion that acetaminophen is clearly shown to cause lymphohematopoietic cancer.

The assessment of studies for lymphohematopoietic neoplasms below is separated by type, including: Lymphoma, Non-Hodgkin Lymphoma (NHL), NOS and its subtypes (Figure 8); Hodgkin Lymphoma (HL, Figure 9), Multiple Myeloma (MM, Figure 10), Leukemia (adult, Figure 11), and Leukemia (childhood, Figure 12). Weiss (2016) discussed the possibility of pain as a source of protopathic bias in several studies for plasma cell carcinoma (Weiss, 2016). Signs and symptoms of cancer may arise years prior to diagnosis and may lead to the choice of acetaminophen for fever and pain relief throughout that time. Since lymphomas are immune-cell malignancies, they may be preceded by immune-related illness. For multiple myeloma (MM) for example, bacterial and viral infections as well as autoimmune diseases have been recognized as potential early signs of risk (Brown et al., 2008; Lindqvist et al., 2011; Lindqvist et al., 2017). For these reasons, steps to mitigate against the risk of protopathic bias should be taken.

(i) Lymphoma, Non-Hodgkin Lymphoma (NHL), NOS and its subtypes

A total of 3 cohort studies and 4 case control studies reported on acetaminophen use and various lymphomas, excluding Hodgkin lymphoma. Of these, 1 cohort and 3 case-control studies reported increased RRs. Assessments of key study characteristics show none of these studies adequately account for critical sources of biases/methodological issues, thus making it difficult to establish whether the observed results represent a causal association or are due to measurement/design issues. Furthermore, the studies addressing the most issues reported no increased risk (Friis et al., 2002; Kato et al., 2002). Notably, the study reporting the highest risks accounted for fewer issues (Baker et al., 2005; Becker et al., 2009; Walter et al., 2011b). A more detailed review of each study is provided in Section 8.1.2.

- **Walter** et al 2011b did not adjust for protopathic bias and because lymphoma is an immunological disease associated with increased risk of fever and febrile infections prior to diagnosis; such adjustment is necessary. (Walter et al., 2011b)
- **Becker** et al 2009 did not adjust for important confounders in the analysis of acetaminophen use. The study did not, for example, adjust for rheumatoid arthritis (RA) prescription medications. It also did not find that RA was a significant risk factor for lymphoma. The lack of consistency of this finding with the rest of the literature is another reason to question the validity of the study in general. (Becker et al., 2009)
- **Baker** et al 2005 did not address protopathic bias and reported inconsistent results by gender and various trends. Specifically, they found no association among males and all

analyses by duration, frequency of use, and cumulative acetaminophen use were not associated. These issues are a concern for a causal interpretation. (Baker et al., 2005).

Assessments of key study characteristics show none of these studies adequately account for critical methodological issues, thus making it difficult to establish whether the observed results represent a causal association or are due to methodological errors in measurement/design.

(ii) Hodgkin Lymphoma (HL)

One cohort and 1 case control study reported on acetaminophen use and Hodgkin lymphoma. The one cohort study reported no association between acetaminophen use and Hodgkin Lymphoma. Assessments of key study characteristics show none of these studies adequately account for critical issues, thus making it difficult to establish whether the observed results represent a causal association or are due to methodological errors in measurement/design. Furthermore, the study addressing the most methodological issues, reports no association (Friis et al., 2002). The case-control study had the higher risk and accounted for the fewest sources of potential bias (Chang et al., 2004).

- Chang et al 2004 reported an association between acetaminophen use and Hodgkin Lymphoma. This study was designed to examine whether aspirin use was associated with a reduced risk of Hodgkin lymphoma, did not adjust for protopathic bias and only obtained exposure data for the last 5 years prior to diagnosis. It also relied on self-reported analgesics use and is therefore subject to recall bias. (Chang et al., 2004)

Given the methodological limitations above and inconsistent results, it cannot be concluded that acetaminophen use is clearly shown to cause increased risk for Hodgkin lymphoma.

(iii) Multiple Myeloma (MM)

Bacterial and viral infections as well as autoimmune diseases have been recognized as potential early signs of risk for MM ((Brown et al., 2008; Lindqvist et al., 2011; Lindqvist et al., 2017). Weiss (2016) noted bone pain as an early sign of MM (Weiss, 2016). Signs and symptoms of cancer may arise years prior to diagnosis and may lead to the choice of acetaminophen for fever and pain relief throughout that time. Pneumonia, a personal history of sinusitis, meningitis, septicemia, herpes zoster, infectious mononucleosis, and myocarditis have been associated with a significantly increased risk of MM (Brown et al., 2008; Lindqvist et al., 2011).

One cohort and 1 case control study reported on acetaminophen use and MM. One cohort study reported no association between acetaminophen use and MM. Assessments of key study characteristics show neither of these studies adequately account for critical issues, thus making it difficult to establish whether the observed results represent a causal association or are due to methodological errors in measurement/design. Furthermore, the study addressing the most

methodological issues, reported no association (Friis et al., 2002). The case-control study had the higher risk and accounted for the least sources of potential bias (Moysich et al., 2007).

- Moysich et al 2007 did not account for protopathic bias which is important since MM patients experience infections and immune-related disease prior to diagnosis. In this hospital-based case control study at a cancer hospital, the analysis only adjusted for age, smoking and year of questionnaire completion. The authors note: "Our results warrant further investigation in population-based case-control and cohort studies and should be interpreted with caution in light of the limited sample size and biases inherent in hospital-based studies." (Moysich et al., 2007)

Given the methodological limitations noted above and inconsistent results, it cannot be concluded that acetaminophen use is clearly shown to cause increased risk for MM.

(iv) Leukemia (adult)

A total of two cohort and 8 case-control studies reported on acetaminophen use and leukemias among adults. Assessments of key study characteristics show none of these studies adequately account for critical issues, thus making it difficult to establish whether the observed results represent a causal association or are due to methodological errors in measurement/design. Furthermore, the studies addressing the most methodological issues, report no association (Friedman, 1982; Friis et al., 2002). The cohort study and 2 case-control studies with the higher risk accounted for fewer sources of potential bias (Ross et al., 2011; Walter et al., 2011b; Weiss et al., 2006).

- **Walter** et al 2011b did not adjust for protopathic bias for this estimate of RR. (Walter et al., 2011b)
- **Weiss** et al 2006 assessed self-reported acetaminophen use and therefore is subject recall bias. Analyses by subtype likely suffered loss of precision and power due to small numbers of cases. (Weiss et al., 2006)
- **Ross** et al 2011 assessed self-reported acetaminophen use in a population-based case-control study in the US (N=670 cases and 701 controls). No associations were seen for myeloid leukemia and subtypes among males. (Ross et al., 2011)

Given the methodological limitations above and inconsistent results, it cannot be concluded that acetaminophen use is clearly shown to cause increased risk for leukemia in adults.

(v) Leukemia (childhood)

A total of 2 case-control studies reported on maternal use of acetaminophen and pediatric leukemias among adults. Since the exposure period of interest was relatively short and recent, the potential for misclassification of exposure is reduced. Protopathic bias is also not a

consideration in these studies. Recall bias is still a risk. However, neither of the studies reported an increased risk (Couto et al., 2015; Ognjanovic et al., 2011).

Both case control studies that assessed the association between acetaminophen use and childhood leukemia reported no increase RRs. Although the studies had limitations, since the latency and exposure periods are relatively short, the risks for bias and confounding are also reduced. One cannot conclude that acetaminophen use is clearly shown to cause increased risk for pediatric leukemia.

A final observation regarding the published studies of lymphohematopoietic cancers is the lack of studies in the known cohorts that are listed in the HID, for example, the Nurses' Health Study and the Health Professionals Follow-up Study. Weiss (2016) mentions the possibility of selective publication of positive results for plasma cell disorders and leukemia in known cohorts(Weiss et al., 2006). While numerous studies have been published on a wide variety of outcomes, including cancer none have been seen for these cancers. Since the Weiss publication, no additional studies on adult lymphohematopoietic cancers have been published, although here was one publication in 2015 on maternal exposure to acetaminophen and leukemia in the offspring (Couto et al., 2015).

Figure 8. Forest plot: Lymphoma, Non-Hodgkin Lymphoma (NHL), NOS and its subtypes

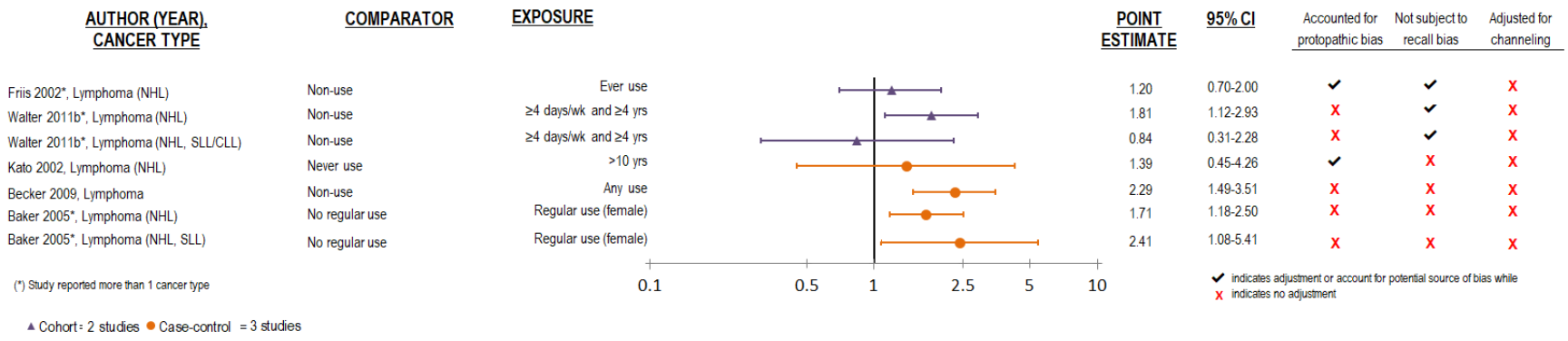


Figure 9. Forest plot: Hodgkin Lymphoma (HL)

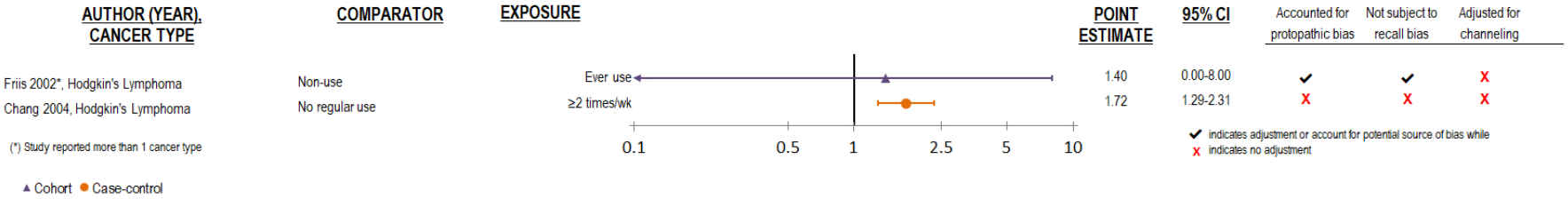


Figure 10. Forest plot: Multiple Myeloma

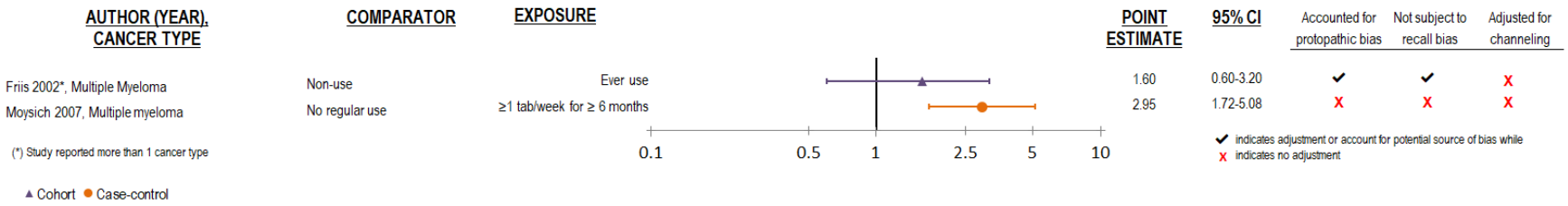


Figure 11. Forest plot: Leukemias (Adult)

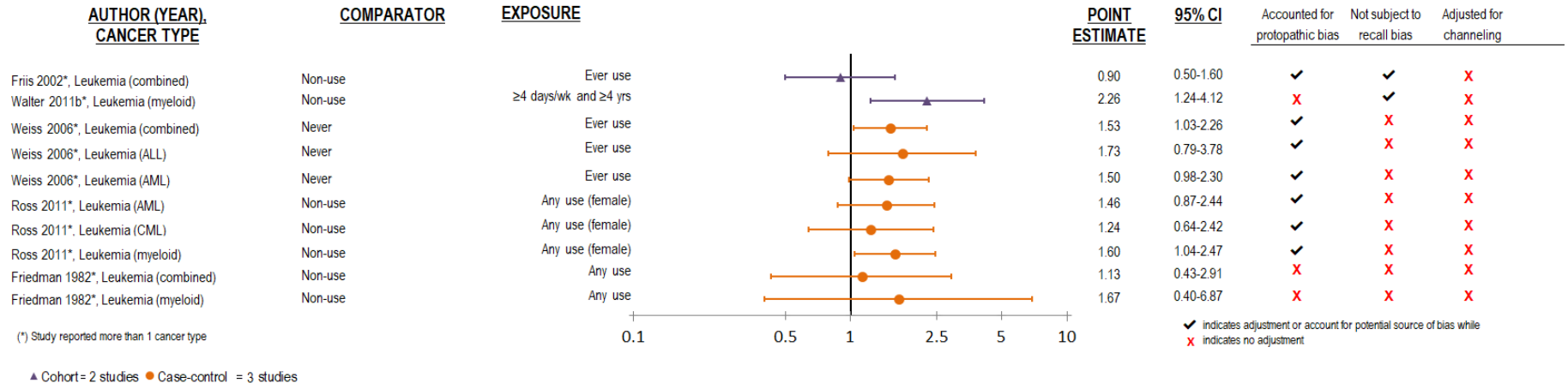
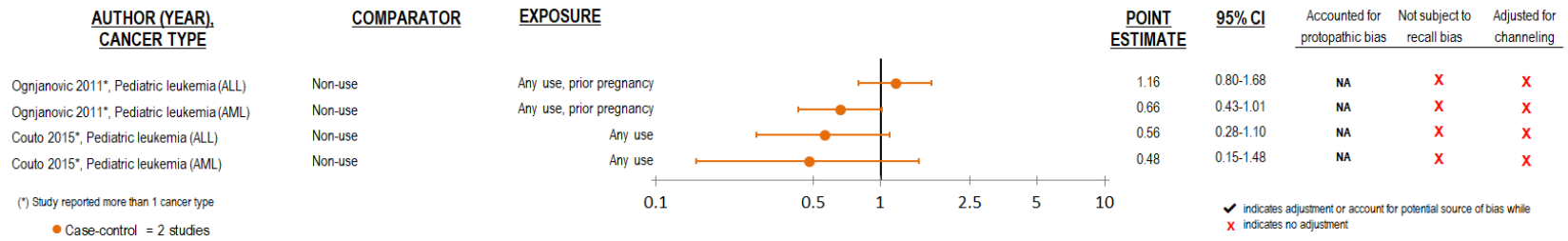


Figure 12. Forest plot: Leukemias (childhood)



3.3.3 Liver Cancer

Assessment of Evidence: The weight of the evidence does not support a conclusion that acetaminophen is clearly shown to cause liver cancer.

A total of 2 cohort studies and 2 case-control studies assessed the association between acetaminophen use and liver cancer (Figure 13). Of these, 1 cohort study and 2 case-control studies reported an increased relative risk. Assessments of key study characteristics show none of these studies adequately account for critical methodological issues, thus making it difficult to establish whether the observed results represent a causal association or are due to methodological errors in measurement/design. Also pre-existing but not yet diagnosed liver cancer could well cause GI symptoms, which would lead to the avoidance of NSAIDS, and channeling to acetaminophen. A more detailed review of each study is provided in Section 8.1.3.

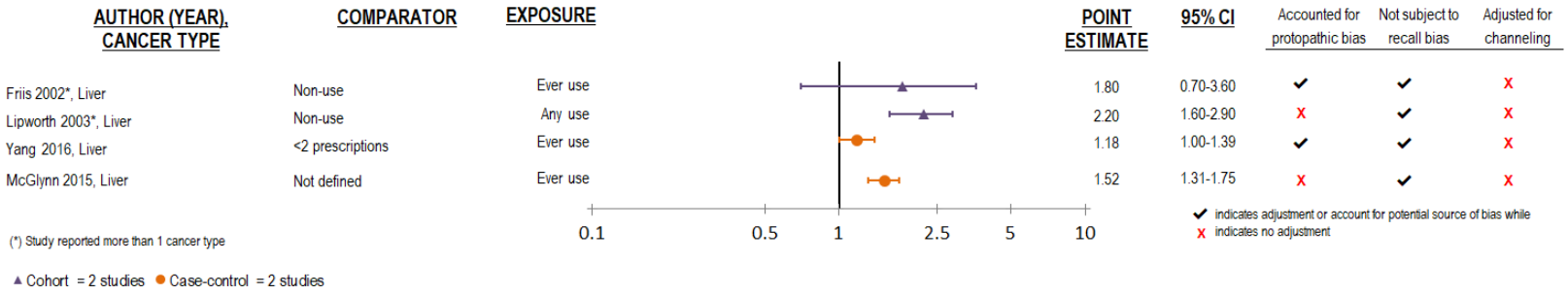
Although Lipworth et al 2003 (Lipworth et al., 2003), McGlynn et al 2015 (McGlynn et al., 2015), and Yang et al 2016 (Yang et al., 2016) reported an association between acetaminophen use and increased risk for liver cancer, the observed associations were likely affected by bias and confounding:

- The cohort study by Lipworth et al 2003 examined mortality as an outcome and only controlled for age and sex as possible confounders. Mortality studies have a limited role in the assessment of acetaminophen use and cancer occurrence since 1) if there was a diagnosis of liver cancer prior to death from liver cancer, the use of acetaminophen may have been for pain after disease onset and 2) many incident cases will be missed since not all patients with the disease will die from liver cancer. (Lipworth et al., 2003)
- McGlynn et al 2015 was an unadjusted analysis. (McGlynn et al., 2015)
- Yang et al 2016 (nested case-control in CPRD) likely suffers from residual confounding due to undiagnosed chronic liver disease (CLD). (Yang et al., 2016)
    - The HID states the following regarding a subgroup analysis by CLD:
        - "…results did not change materially when restricting the analyses to individuals without chronic liver disease, supporting that confounding by indication was not an explanation for these positive findings." (HID, Sect 3.1.2.3 Liver cancer, p92)
    - However, chronic liver disease is asymptomatic until later stages, therefore it often goes undiagnosed (Runyon 2011). Further, the proportion of liver cancer patients with cirrhosis is 80-90% (El-Serag, 2012). In a cohort of nearly 630 liver cancer patients in the UK, only 20% had no known chronic liver disease (Dyson et al., 2014). The definition for CLD was not provided in Yang et al but the proportion of cases with CLD was very low (14 percent) (Yang et al., 2016).

Undiagnosed chronic liver disease among the cases in Yang et al is likely to have caused residual confounding.

- In the analysis restricted to individuals who, in the data, were not identified as having chronic liver disease, an increased OR was reported for liver cancer. The OR for those who were diagnosed with CLD was however very large (OR 32.8, 95% CI 20.6–52.1) and, therefore, even if there was a small proportion of the cases with undiagnosed liver disease, the OR would be inflated in the subgroup analysis of those without a chronic liver disease diagnosis in the data.

Figure 13. Forest plot: liver cancer



3.4 Other Cancers: Hormone-related Cancers, Skin, Colorectal, Brain, Respiratory Tract, Gastrointestinal Tract, Pancreatic, All Sites Combined

The HID states: "For cancers of the breast, ovary, uterine endometrium, prostate, skin, and colorectum, the association with acetaminophen use was either decreased, null, or inconsistent. The data from cohort and case-control studies from a number of other cancer sites were too sparse to evaluate thoroughly, namely the brain, respiratory tract, gastrointestinal tract (stomach, esophagus, oral/pharyngeal cancer), pancreas, cervix, and all cancers combined." We agree these studies do not provide evidence that acetaminophen is clearly shown to cause any of these forms of cancer. Forest plots for each form of cancer are available in Section 8.3.

3.5 Studies Conducted to Quantify Bias in Epidemiologic Studies on the Association Between Acetaminophen and Cancer

We conducted a review of the study design characteristics that were employed in prior publications that examined association between acetaminophen and cancer. We observed that most publications used a case-control design and there were some commonalities in analysis choices within the case-control design (including general use of age and gender as matching criteria), and also some differences. The other publications applied a cohort design, where patients exposed to acetaminophen were compared to non-users. In all of these publications, the studies failed to examine new users of acetaminophen and demonstrate a balanced comparison with some alternative treatment at the start of follow-up, thereby risking threats to validity due to various sources of bias (e.g. channeling) and confounding (due to an imbalance of patient population characteristics at baseline).

In order to quantify the extent of bias that could be present in these studies, two analyses were performed to replicate the studies (See Section 8.4 for the full study results and a link to the pre-specified protocol on-line). In these, variations were used in the study designs seen in the literature, and for each design, measured the residual systematic error through the use of a sample of negative control outcomes for which we a priori expect to observe no association.

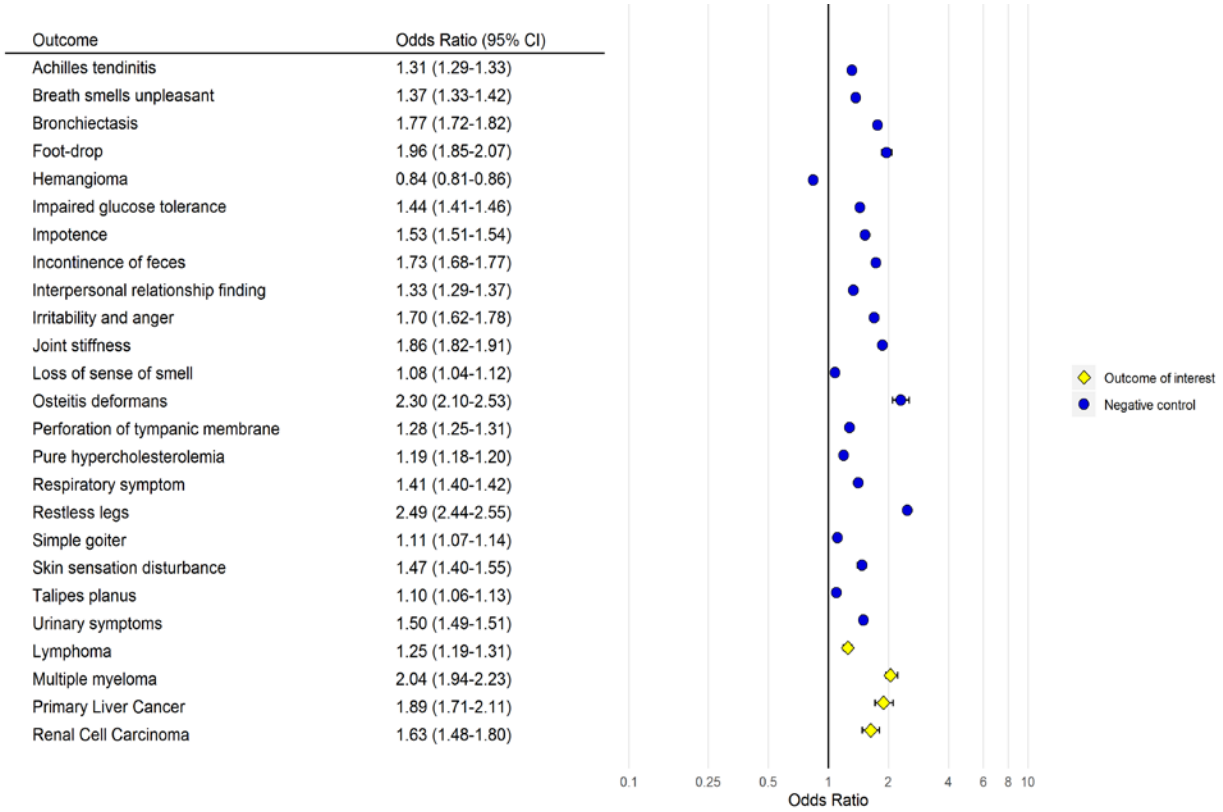
3.5.1 Study Assessing the Clinical Practice Research Datalink

The findings of this systematic error were reported to provide context around how much confidence one can draw from an unknown effect estimate (versus known for the negative controls) produced for an outcome of interest (here, the cancer outcomes) using these epidemiologic study designs. We carried out this analysis using the existing observational data in the Clinical Practice Research Datalink (the same data source in several of the case controls studies (Kaye et al., 2001; Yang et al., 2016)).

The results show that the study designs which were applied in the literature to examine the association of acetaminophen and cancer have substantial bias that are likely producing spurious statistically significant estimates upward of RR=2 even when no true effect exists. These findings suggest that observed estimates in the literature are in range of what would be expected due to study bias alone and should not be interpreted as clear evidence of a true causal effect. (Figure 14)

Figure 14: Replication of epidemiology study designs from studies reviewed in the HID in the CPRD database using negative control outcomes (e.g. restless leg) for which there is high certainty of no association between the endpoint and cancer or acetaminophen use. The results show that the study designs which were applied in the literature to examine the association of acetaminophen use (versus no use) and cancer have substantial bias and have a high likelihood of producing spurious statistically significant estimates upward of RR=2 even when no true effect exists. These findings suggest that observed estimates in these studies are in the range of what would be expected due to study bias alone and should not be interpreted as demonstrating a causal effect.

Analysis 1: Sampling, all time prior, adjusted for age, sex & year

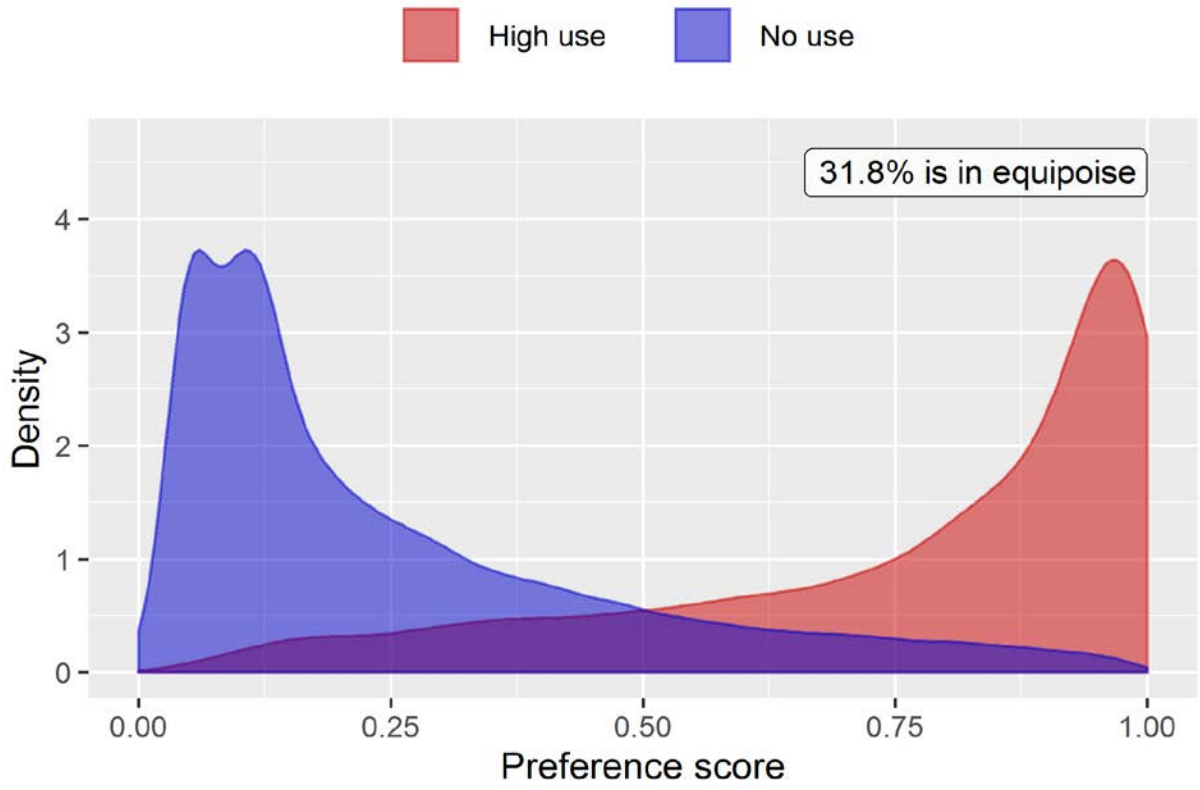


3.5.2 Propensity analysis

In the emulation of study designs seen in the literature, we replicated the cohort design used in the Walter et al. (2011b) study and specifically, the high use group (Walter et al., 2011b). We also fitted a propensity model to evaluate to what extent the 2 exposure groups are comparable. This model was fitted by included a very large set of covariates (all prior drugs, drug classes, diagnoses, procedures, etc.), and using a regularized logistic regression.

Figure 15 below shows the preference score distribution. The preference score is a transformation of the propensity score to account for the different sizes of the 2 exposure groups (Walker et al., 2016).

Figure 15. Preference score distribution of high acetaminophen users versus nonusers



The propensity score plot tells us that the cohorts are very different and there is very little overlap of characteristics in the individuals in the database. For most people their treatment assignment was highly predictable. This means the data can reliably predict who will be prescribed (channeled to, in this case) acetaminophen or not. This reinforces the notion of channeling to the drug based on existing comorbidities/medications/treatments.

The region around the preference score value of 0.5 is where individuals are equally likely to receive a prescription of acetaminophen (defined as clinical equipoise). Ideally, the region in

clinical equipoise, between 0.25 and 0.75 on the graph, would have the highest density of patients in both groups, or at least 50% of the patients. However, the large peaks of probability at either end of the plot show that this is clearly not the case, since only 32% of the cohorts fall in the region of clinical equipoise. Instead, the plot shows that the 2 groups are very different. Thus, the potential for bias is quite high.

Note that a rule-of-thumb is that all covariates must have a standardized difference of mean 0.10 for us to consider 2 groups 'balanced'. There are 1,312 covariates that do not meet our rule-of-thumb for balance. All these unbalanced indicate that the high-use group is already 'sicker' at baseline on all these dimensions. For example, the high-users are more often exposed to antibiotics, diuretics, drugs for acid-related disorders, and antidepressants than non-users. (These drug classes, have the largest standardized difference of mean, likely because they represent entire disease areas.)

In conclusion, the bias observed in our emulations of the designs used in prior observational studies was considerable, despite the fact that these designs attempt to adjust for confounding and other forms of bias in several ways. The demonstrated magnitude of the bias is such that it could explain the few positive findings reported in the literature as solely reflecting bias. This is not proof that no causal link between acetaminophen and cancer exists, just that these studies cannot inform us on whether it does.

3.6 Conclusions

This evaluation reviewed published scientific reports to answer an etiologic question about exposure to acetaminophen and cancer incidence using available evidence from more than 130 epidemiology studies. Most of these studies across a multitude of cancer types reported no association with acetaminophen use. The weight of the epidemiologic evidence does not support a finding that acetaminophen is clearly shown to cause cancer due to the lack of several important factors for determining causal association including: consistent results with a variety of study designs, a strong association, an association with increasing cumulative exposure that is consistent with known biology of cancer (i.e. latency), an association not subject to bias or confounding and, finally, publications of all outcomes from studies in several of the known cohorts with available data.

In addition, specific characteristics of acetaminophen use make it challenging to accurately assess cancer risk through generally accepted epidemiologic methods. Channeling bias with the use of acetaminophen is likely to be present in several cancer outcomes; no studies included in this review successfully adjusted for channeling, and as demonstrated by the quantification of bias study, considerable error exists in these studies. Furthermore, since acetaminophen is an analgesic and cancer can cause pain, the risk of protopathic bias is a concern which must be mitigated against in study design or analysis. Although many studies recognized the potential for

protopathic bias, most studies did not demonstrate sufficient control. For example, evidence shows that lymphohematopoietic neoplasm precursors could be present for many years prior to disease onset or diagnosis, leading to the choice of acetaminophen for fever and pain relief during that time. Finally, because acetaminophen is available both through OTC and prescription access, complete and accurate measurement of exposure is virtually impossible. For studies included in this review, exposure measurement was variable and generally poor; studies that captured OTC use relied on self-report. Recall bias is a strong concern for case-control studies as exposure assessment occurs after cancer diagnosis, likely influencing the relative risk to show an artificial, positive association. Database studies relying on prescriptions miss considerable exposure through OTC access. Few studies sought to measure cumulative exposure or to evaluate latency, which are important to determine causal association.

The limits of the body of evidence in the published epidemiologic data were confirmed in a study to quantify the extent of bias in case-control and cohort studies. The analysis of negative controls under a variety of study variations resulted in widely varying RRs among outcomes where the RRs were expected to be 1.0. The RRs for outcomes of interest (cancers) were within the range of error observed in the negative controls in all study designs. The implication of this result is that the RRs reported in the reviewed epidemiologic studies of acetaminophen use and cancer could have equally been achieved due to bias. As such, RRs of 2.5-3 could have just as easily occurred through the systematic error (bias) reflected in the negative controls, and therefore should not be used to infer causation from the data.

In view of the challenges related to assessment of acetaminophen use, as well as the influence of several important biases and methodological limitations, the epidemiologic evidence on acetaminophen use does not meet the standard of clearly shown to cause cancer.

4 Carcinogenicity Studies in Animals

4.1 Summary of Animal Carcinogenicity Results

Table 3 contains a summary of the results and weight of evidence assessment of the long-term studies that evaluated acetaminophen carcinogenicity. These results clearly show that acetaminophen is not a carcinogenic hazard.

Table 3: Results and weight of evidence assessment of the long-term animal carcinogenicity studies demonstrating that acetaminophen is not a carcinogenic hazard

Study	Species (Strain)	Increased Tumors Identified by Authors[a]	Weight of Evidence Assessment	Comments
(NTP, 1993)	Mouse (B6C3F1)	None	Supports no Hazard	• "No evidence" in males and females
(NTP, 1993)	Rat (F344/N)	None in male rats; equivocal evidence in female rats	Supports no Hazard	• "No evidence" in males; "Equivocal evidence" in females • Increase in Mononuclear Cell Lymphoma (MCL) in females but not males • Highly variable background incidence of MCL in F344 rats • Tumor type not considered relevant to human hazard assessment (Maronpot et al., 2016)
(Amo and Matsuyama, 1985)	Mouse (B6C3F1)	None	Supports no Hazard	• Considered a negative study in both sexes by the authors and by IARC; the HID re-analyzed the data and reported increased liver and pituitary gland tumors in high dose females, but not males; no increase in tumors was seen in mice in the NTP bioassay.
(Hagiwara and Ward, 1986)	Mouse (B6C3F1)	None	Supports no Hazard	
(Hiraga and Fujii, 1985)	Rat (F344/Du CrJ)	None	Supports no Hazard	
(Johansson, 1981a)	Rat (SD)	None	Supports no Hazard	
(Flaks and Flaks, 1983)	Mouse (IF)	Liver tumors (mostly benign)	Not relevant	• Not scientifically valid based on significant mortality (up to 55%) and large decreases in weight gain • Doses greatly exceeded MTD • Highly questionable relevance based on chronic hepatotoxic exposures
(Flaks et al., 1985)	Rat (Leeds)	Benign bladder tumors and liver adenomas and carcinomas	Not relevant	• Not scientifically valid according to generally accepted principles • No tumors of any type in the control animals (lack of credibility noted by IARC) • Many other study deficiencies • All tumors were benign and CIC listing criteria focus exclusively on malignant tumors. • Lack of dose related increase in benign bladder tumors

[a] Increased tumors identified and considered treatment-related by the study authors

A more detailed evaluation of these studies is provided in the sections that follow.

4.2 Assessment of Animal Carcinogenicity Results

IARC's finding of "inadequate evidence" in animals is reassuring and supports that acetaminophen does not meet the "clearly shown to cause cancer" standard.

Since its approval in 1951, multiple health authorities and agencies around the world have evaluated and re-evaluated the safety of acetaminophen and unanimously concluded that the drug is not a carcinogenic hazard (FDA, 2010; IARC, 1999; NTP, 1993). IARC conducted a comprehensive evaluation of the long-term animal carcinogenicity studies of acetaminophen in 1990 and 1999. In 1990, IARC concluded that there is "limited evidence" in experimental animals. In 1999, IARC updated its evaluation of acetaminophen, incorporating new studies, including the NTP bioassay; IARC (1999) downgraded the level of evidence in animals: "There is *inadequate evidence* in experimental animals for the carcinogenicity of paracetamol [i.e., acetaminophen]." IARC's (1999) evaluation is relevant because IARC evaluated the complete set of available long-term carcinogenicity studies of acetaminophen, as well as all of the short-term tumor promotion studies and classified acetaminophen as Group 3. There are no long-term carcinogenicity studies or tumor promotion studies of acetaminophen that were not considered by IARC in 1999.

NTP (1993) cancer bioassay shows that acetaminophen does not cause cancer in mice or rats.

The most current and most comprehensive cancer study of acetaminophen in animals is the NTP cancer bioassay in which rats and mice were administered 0, 600, 3000, or 6000 ppm of acetaminophen in the diet (NTP, 1993). NTP concluded:

- B6C3F1 Male mice: "*no evidence of carcinogenicity activity*"
- B6C3F1 Female mice: "*no evidence of carcinogenic activity*"
- F344 Male rats: "*no evidence of carcinogenic activity*"
- F344 Female rats: "*equivocal evidence of carcinogenic activity*"

According to NTP, there was not a single instance of "*clear evidence*" or even "*some evidence*" of carcinogenic activity in either sex of either species in the NTP cancer bioassay. NTP describes evidence in four ways; in descending order of the strength of evidence, NTP's descriptors of carcinogenic activity are: "*clear evidence," "some evidence," "equivocal evidence,"* and *"no evidence*." NTP describes "*some evidence*" as follows: "*Some evidence* of carcinogenicity activity describes studies that are interpreted as showing a chemically related increased incidence of neoplasms (malignant, benign, or combined) in which the strength of the response is less than that required for clear evidence." NTP describes "equivocal evidence" as follows: "Equivocal

evidence of carcinogenic activity describes studies that are interpreted as showing a marginal increase of neoplasms that may be chemically related."

The NTP's lone "*equivocal evidence*" finding was due to increased incidences of mononuclear cell leukemia (MCL) in female rats (NTP, 1993). It is highly doubtful that the "*equivocal evidence*" in female rats in the NTP bioassay is treatment-related. NTP noted: (1) "the relatively high spontaneous rate of mononuclear cell leukemia and its highly variable incidence in controls (6-40%) increases the likelihood that such differences in neoplasm incidence among groups could occur by chance" (NTP, 1993), and (2) "the incidence of mononuclear cell leukemia in male rats decreased with dose and therefore did not support the increase observed in females" (NTP, 1993). In 2006, the NTP discontinued using the F344 rat in its cancer bioassay program due largely to the high background control incidences of MCL and Leydig cell tumors (King-Herbert et al., 2010). In its evaluation of acetaminophen, the IARC Working Group noted "the high and variable incidence of mononuclear-cell leukemia between and within studies with Fischer rats and considered that this was not a treatment-related effect" (IARC, 1999). Further, no increase in MCL was observed in male or female rats in three other carcinogenicity studies (Flaks et al., 1985; Hiraga and Fujii, 1985; Johansson, 1981b) of acetaminophen, including a study (Hiraga and Fujii, 1985) in which F344 female rats were given a higher dose of acetaminophen (13,000 ppm in the diet) than were the F344 females in the NTP bioassay (6000 ppm in the diet).[1] And finally, in a 2016 review of the legacy of the F344 rat as a cancer bioassay model by the former Chief of the Laboratory of Experimental Pathology at the National Toxicology Program/NIEHS, Maronpot et al. wrote: "Therefore, the commonly occurring F344 rat MNCL [mononuclear cell leukemia] cannot be considered a relevant predictor of human disease" (Maronpot et al., 2016).

The results of the NTP cancer bioassay are reassuring. The finding of no increased tumor incidence in the NTP cancer bioassay of acetaminophen is particularly relevant to the issue of human hazard identification since, as noted in the HID, "The metabolism of acetaminophen is largely similar in humans and laboratory animals, with many of the same metabolites detected in both humans and animals" (OEHHA, 2019). If acetaminophen were to cause cancer, it would be expected to produce clear evidence of carcinogenicity in the NTP cancer bioassay, and it did not come close.

---

[1] The fact that there was a negative study in F344 rats at a higher dose of acetaminophen Hiraga, K., Fujii, T., 1985. Carcinogenicity testing of acetaminophen in F344 rats. Jpn J Cancer Res 76, 79-85. was not considered by NTP when it assigned "equivocal evidence" to female rats exposed to acetaminophen in its cancer bioassay Maronpot, R.R., Nyska, A., Foreman, J.E., Ramot, Y., 2016. The legacy of the F344 rat as a cancer bioassay model (a retrospective summary of three common F344 rat neoplasms). Critical Reviews in Toxicology 46, 641-675.

According to Dr. James Huff (NIEHS):

"For those [known human carcinogens] that can be studied experimentally, the qualitative concordance between humans and animals approaches unity, and in every case, there is at least one common organ site of cancer in both species. … All 39 human carcinogens that have undergone adequate experimental studies have been shown to cause cancer in animals, and exhibit concordance for tumor sites. … All exposures identified as being carcinogenic to humans that have been studied adequately have been shown to cause cancer in experimental animals at minimally toxic exposures" (Huff, 1993).

Similarly, Dr. Robert Maronpot and colleagues at NIEHS wrote:

"Although we know that all known human carcinogens are also carcinogenic to rodents, it is noteworthy that nearly one-third of these were first identified in animals and only subsequently in humans (Huff, 1993; Tomatis, 1979). … Although not perfect, there is enough concordance between human and rodent carcinogens, in repeatability of bioassay results, and in site-specificity to warrant continued use of existing hazard identification testing approaches until such time as we develop a more suitable means of identifying agents with human carcinogenic potential" (Maronpot et al., 2004).

Given the concordance between known human carcinogens and findings in animal studies, the lack of carcinogenicity of acetaminophen in the NTP cancer bioassay provides compelling evidence that acetaminophen is not carcinogenic. NTP has not listed acetaminophen in its Report on Carcinogens (NTP, 14th Report on Carcinogens, 2016).

Six other carcinogenicity studies of acetaminophen published prior to the NTP cancer bioassay do not indicate it causes cancer in mice or rats.

No treatment-related tumor findings were observed in other long-term animal carcinogenicity studies of acetaminophen conducted prior to the NTP cancer bioassay. These studies include:

- A study in male rats (0, 5350 ppm in the diet) for up to 117 weeks (Johansson, 1981b)
- Hiraga and Fujii (1985) study in male rats (0, 4500, 9000 ppm in the diet) and female rats (0, 6500, 13,000 ppm in the diet) for 2 years and additional 26-week observation,
- Amo and Matsuyama (1985) study in male and female mice (0, 3000, 6000 ppm in the diet) for 31 months, and
- Hagiwara & Ward (1986) study in male mice (0, 5000, 10,000 ppm in the diet) for 70 weeks.

While these studies were not as thorough and comprehensive as the (1993) NTP cancer bioassay, they provide significant additional support for the lack of carcinogenicity of acetaminophen in long-term animal studies that involved daily dosing at levels well above those that would be toxic to humans.

The HID reports "significant tumor findings" in female mice in the Amo and Matsuyama (1985) study, noting "the incidence of hepatocellular adenoma or carcinoma combined was significantly increased in the high-dose group by pairwise comparison with controls, with a significant dose-related trend." However, this is not the interpretation of either the study authors or IARC. Contrary to the HID, Amo and Matsuyama (1985) stated:

"There was no difference in the rates of tumor-bearing mice for the male groups, but those for the 2 female experimental groups were slightly lower than that of the control group (Fig. 7). Tumors were found in the mice of the control and experimental groups, and in various organs: the hematopoietic tissues (bone marrow, thymus, spleen and lymph nodes) lungs, liver, pituitary, digestive tract, uterus, ovaries, breasts, adrenals and skin were involved (Table 1), with no statistical difference in the incidences. These tumors were regarded as spontaneous tumors of the B6C3F1 mice. **The results of the present tests show that feeding the maximum tolerated dose of acetaminophen (0.6% diet) held no carcinogenic hazard for B6C3F1 mice"** (Amo and Matsuyama, 1985)**.** [emphasis added]

In contrast to the analysis in the HID, the evaluation by Amo and Matsuyama took into account the variability in spontaneous tumors observed in B6C3F1 mice. IARC (1990) reached a similar conclusion about this study:

"No difference was found in the incidence of tumours at any site between treated and control mice (Amo and Matsuyama, 1985)."

Since Amo and Matsuyama (1985) and the NTP (1993) cancer bioassay were both conducted using B6C3F1 mice, it is instructive to compare the liver tumor results in the two studies, as summarized in Table 3. In male mice, there were ***decreases*** in the incidences of liver tumors, expressed as adenomas and carcinomas combined, at the high dose in both studies, and the decrease was statistically significant in the NTP (1993) cancer bioassay. In female mice, the small increase in the liver tumors at the high dose in the Amo and Matsuyama (1985) study was not observed in the NTP (1993) cancer bioassay, i.e., the incidence of liver tumors was virtually the same in the control and high dose groups (Table 4). It is also noted that the incidence of liver tumors in the male mouse was *decreased* in the high dose compared to controls in the Amo and Matsuyama (1985) study; the authors also considered this to be a reflection of the variability in spontaneous liver tumors. The B6C3F1 mouse strain is highly susceptible to liver tumors, and liver tumors are the most common type of tumor induced in B6C3F1 mice by exposures to test materials in NTP cancer bioassays. Yet, there is no meaningful evidence that acetaminophen causes liver tumors in male or female B6C3F1 mice in the NTP cancer bioassay.

Table 4: Comparison of the incidence of liver adenoma and carcinoma combined among male and female B6C3F1 mice in Amo and Matsuyama (1985) and NTP (1993).

Study	**Concentration of acetaminophen in the diet**			
	0 ppm	**600 ppm**	**3000 ppm**	**6000 ppm**
Amo and Matsuyama (1985)				
Male mice	13/43 (30%)	ND	12/39 (31%)	6/45 (13%)
Female mice	2/49 (4%)	ND	2/46 (4%)	8/50 (16%)
NTP (1993)				
Male mice	16/50 (32%)	9/50 (18%)	10/50 (20%)	7/50* (14%)
Female mice	3/49 (6%)	4/50 (8%)	7/50 (14%)	3/49 (6%)

*Statistically significant by pair-wise comparison using Fischer exact test, $p<0.05$. Statistical results are those presented by the study authors.

The HID also described a statistically significant increase in benign pituitary gland tumors in the females, but not the males, at the high dose in the Amo and Matsuyama (1985) study (based on its statistical re-evaluation of the pituitary adenoma data (Table 5). Pituitary adenomas are extremely common in this strain and rarely become malignant (Sarich et al., 1997). No malignant tumors of the pituitary gland were observed in this study, and this is important because the CIC listing criteria gives greater weight to increases in malignant tumors than benign tumors (OEHHA, 2001). In comparison, in the NTP (1993) cancer bioassay, no difference in the incidence of pituitary gland adenomas was observed at the same high dose compared to controls among either male or female B6C3F1 mice (Table 5).[2] Considered collectively, these data provide no clear or consistent evidence of an increase in tumors of the pituitary gland in male or female B6C3F1 mice.

Table 5: Comparison of the incidence of pituitary gland adenomas among male and female B6C3F1 mice in Amo and Matsuyama (1985) and NTP (1993).

Study	**Concentration of acetaminophen in the diet**			
	0 ppm	**600 ppm**	**3000 ppm**	**6000 ppm**
Amo and Matsuyama (1985)				
Male mice	0/43 (0%)	ND	1/39 (3%)	1/45 (2%)

[2] In the NTP cancer bioassay, no carcinoma of the pituitary gland was observed in male mice; in female mice, the incidence of carcinoma of the pituitary gland was 1/46, 1/43, 1/42 and 0/45 at 0, 600, 3000, and 6000 ppm, respectively.

Female mice	2/49 (4%)	ND	3/46 (7%)	9/50* (18%)
NTP (1993)				
Male mice	0/48 (0%)	0/39 (0%)	0/39 (0%)	0/46 (0%)
Female mice	14/46 (30%)	16/43 (37%)	7/42 (17%)	14/45 (31%)

*Statistically significant by pair-wise comparison using Fischer exact test, $p<0.05$. Statistical results are those presented by the study authors.

The HID erroneously includes an 11-month tumor promotion study by Weisburger et al. (1973) among the long-term carcinogenicity studies of acetaminophen. This study was designed to assess the potential effect of acetaminophen given in combination with potent tumorigenic and genotoxic initiators in mice, rats and hamsters. It was not designed to evaluate the carcinogenic potential of acetaminophen alone given the short duration of the study and small group size. The results of this study and other tumor promotion studies are described in a section below on tumor promotion.

The HID also included a negative study of acetaminophen in male and female ABC-AF mice by Wright (1967). According to the HID: "This study was originally planned as a five-generation study; however, poor survival in the treated F0 males and females and reduced fertility resulted in the discontinuation of acetaminophen studies beyond the F1 generation. Although males were treated, only mammary tumors were accessed [sic], and only in female mice. Due to the survival issues, incidence of female mammary tumors was reported for all three treatment groups combined (8%); control female mammary tumor incidence was 9.2%."The mean lifetime survival was reported to be less than 40 weeks in this study. While this study did not report any statistically significant increase in tumors, it should not be regarded as "scientifically valid testing according to generally accepted principles" for many reasons.

<u>The Flaks and Flaks (1983) and Flaks et al. (1985) studies are inadequate to clearly show acetaminophen causes cancer</u>

Long-term carcinogenicity studies by the same group of investigators reported positive findings of tumorigenicity of acetaminophen in mice and rats (Flaks and Flaks, 1983). However, these studies have serious limitations, and for the reasons described below, they do not represent "scientifically valid testing according to generally accepted principles." Notably, IARC (1999) evaluated both of these studies when it determined that there is "inadequate evidence" of carcinogenicity in animals.

Flaks and Flaks (1983) gave IF mice diets containing 0, 5000 or 10,000 ppm of acetaminophen. Increased incidences of liver tumors (mostly benign) were reported in both male and female mice at the higher dose level (10,000 ppm), which was described by IARC as a "markedly toxic dose"

(IARC, 1999). At the lower dose (5000 ppm), no increase in any tumors was observed in either males or females. The higher dose level, which was the only dose that produced an increase in any tumors in mice, greatly exceeded the Maximum Tolerated Dose (MTD), and it is a widely accepted principle that the high dose in an animal carcinogenicity study should not exceed the MTD (OECD, 2018). US EPA has cautioned against the use of excessively high dose levels, which would confound interpretation of study results (EPA, 2005). According to the US EPA, "significant increases in mortality from effects other than cancer generally indicate that an adequate high dose has been exceeded" (EPA, 2005). The underlying reason for this guidance is that cytotoxicity can occur especially at doses that exceed the MTD. Dr. David Gaylor, formerly of the National Center for Toxicological Research, stated: "Increased carcinogenicity would be expected from increased opportunities for mutagenic activity during regenerative cell replication to compensate for cytotoxicity" (Gaylor, 2005). Under such conditions, an increase in cancer is likely due to one or more nearly universal modes of action, such as, regenerative cell replication, rather than due to some unique carcinogenic property of a chemical.

The death rate at the high dose level in Flaks and Flaks (1983) was 55% (33/60) and 12% (7/60) among the male and female mice, respectively; most of these deaths occurred during the first two days of exposure. The US EPA Guidelines for Carcinogen Risk Assessment state: "With regard to the appropriateness of the high dose, an adequate dose would generally be one that produces some toxic effects ***without unduly affecting mortality*** from effects other than cancer or producing significant adverse effects on the nutrition and health of the test animals (OECD, 1981; NRC, 1993a)" (EPA, 2005). [emphasis added] Exposure to acetaminophen in the diet was terminated for the surviving high-dose mice at 18 months of exposure when the body weights of the males and females were approximately 38% and 31%, respectively, lower than those of the controls. Over the course of 18 months, this amounts to decreases in body weight gain greater than 100% (due to weight loss) and nearly 86% in males and females, respectively (Flaks and Flaks, 1983).[3] In long-term toxicity studies, dose levels that produce greater than a 10% reduction in body weight gain compared to controls are considered to exceed the MTD and inappropriate (OECD, 2018).

The carcinogenicity study of acetaminophen by Hagiwara and Ward at the National Cancer Institute also noted that a dose of 10,000 ppm of acetaminophen in the diet to mice greatly exceeds the MTD when nearly half the mice at this high dose in their study died before 24 weeks: "A dietary level of 10,000 ppm would appear considerably greater than the maximally tolerated dose (MTD), as defined by body weight gain suppression, mortality, and incidence and severity of hepatic lesions" (Hagiwara and Ward, 1986).

---

[3] These estimated reductions in body weight and body weight gain are based on Figure 1 in Flaks and Flaks (1983).

Other limitations of the Flaks and Flaks (1983) study include: inadequate description of the methodology, lack of statistical analysis of the tumor data, no randomized assignment of animals, no observation of clinical symptoms, no measurement of food consumption, no testing of diets to validate the concentration and stability of the test material, frequency of body weight measurements not stated, and only two dose levels of acetaminophen. In addition, the IF mouse is not a strain recommended or used for carcinogenicity testing by any regulatory or scientific organization. In fact, a PubMed search did not identify any other published long-term carcinogenicity study of any substance conducted in IF mice by these or any other investigators; we found no evidence of a historical control database for the IF mouse. Finally, none of the other carcinogenicity studies of acetaminophen, including the NTP cancer bioassay, reported an increase in liver tumors in mice.

For their study of acetaminophen in rats, Flaks et al. (1985) used Leeds rats, an inbred strain of rats. These investigators reported statistically significant increases in benign, but not malignant, bladder tumors in males at the high dose (10,000 ppm) only and in females at the low dose (5000 ppm) only and in liver "neoplastic nodules" in both sexes at the high dose only. First, this study did not report any statistically significant increase in malignant tumors, and the CIC listing criteria places greater weight on malignant tumors than benign tumors (OEHHA, 2019). The HID states: "In males, statistically significant increases in urinary bladder transitional cell papilloma and transitional cell papilloma and carcinoma combined were also seen in the high-dose group, with positive dose-related trends" (OEHHA, 2019). However, it is important to recognize that the statistically significant increase in combined tumors was due to an increase in the benign tumors, not malignant tumors, since there was never more than a single male rat with bladder carcinoma in any dose group. The HID also states: "In females, a statistically significant increase in urinary bladder transitional cell papilloma and carcinoma combined was seen in the mid-dose group." Actually, this sentence refers to the findings at the low dose, since there was no mid-dose group in this study, and once again, the statistically significant increase in combined tumors is attributable to benign, not malignant, tumors since there was never more than a single female rat with bladder carcinoma in any group (OEHHA, 2019).

It is questionable whether the benign bladder tumors reported by Flaks et al. (1985) are actually bladder tumors. The photomicrograph of a benign bladder tumor (bladder papilloma) in the publication by Flaks et al. (1985) was recently reviewed by two highly-respected pathologists, Dr. Samuel Cohen (University of Nebraska Medical Center) and Dr. Jerry Hardisty (EPL, Inc.); they both concluded that that the "bladder papilloma" in the photomicrograph is not a bladder papilloma, but represents papillary hyperplasia. Papillary hyperplasia is a reversible lesion, identical to the histopathology reported by Shirai et al. (Shirai et al., 1986; Shirai et al., 1995) with uracil, which reversed within 1-2 weeks of discontinuing treatment. They were associated with calculi, not seen in humans or in other studies in rats, including in the NTP study.

Second, the IARC Working Group "noted that in the study in rats in which tumours were induced (Flaks et al., 1985) no tumours were found in either male or female controls, which is a highly unusual finding and raises questions about the interpretation of the findings" (IARC, 1999). The HID states that "other publications from the same laboratory corroborate the extremely low spontaneous incidence of liver and bladder neoplasms in Leeds rats" (OEHHA, 2019). However, the HID does not consider the more important point that these investigators did not find any tumors in any tissues in any control group of male or female Leeds rats in the acetaminophen study or in any of their other carcinogenicity studies, which are identified in the HID.[4] It is unheard of to have no background tumors in the control group of a rodent bioassay and therefore defies credibility.

Third, as with the IF strain of mouse used by these investigators the Leeds rat is not a strain used or recommended for carcinogenicity testing by any regulatory or scientific organization. A PubMed search did not identify any other published long-term carcinogenicity study of any substance conducted in Leeds rats by any other investigators; we found no evidence of a historical control database for the Leeds rat. Fourth, other limitations of the Flaks et al. (1985) study include: limited description of methods, no description of the statistical methods, no randomized assignment of animals, no observation of clinical symptoms, no testing of diets to validate the concentration and stability of the test material, and infrequent (monthly) measurements of body weights. Fifth, none of the other carcinogenicity studies of acetaminophen, including the NTP cancer bioassay, reported an increase in bladder or liver tumors in rats. In summary, this highly-questionable study cited within the HID as suggestive of increased carcinogenic hazard risk: (1) does not meet the statutory standard of "scientifically valid testing according to generally accepted principles," (2) reported increases in benign tumors only, and (3) is inconsistent with the results of three other carcinogenicity studies of acetaminophen in rats that did not observe increases in either bladder or liver tumors.

The NTP cancer bioassay and the other animal studies are capable of detecting the types of tumor types showing elevated RRs in some epidemiologic studies.

---

[4] Flaks et al. Flaks, B., Flaks, A., Shaw, A.P., 1985. Induction by paracetamol of bladder and liver tumours in the rat. Effects on hepatocyte fine structure. Acta Pathol Microbiol Immunol Scand A 93, 367-377. did not find a single tumor in 40 control male and 40 control female Leeds rats in their acetaminophen study. The HID notes that no liver tumors were observed among 40 untreated male Leeds rats in an earlier study by Flaks, A., and Flaks B., Flaks, A., Flaks, B., 1982. 3-Methylcholanthrene-inhibition of hepatocarcinogenesis in the rat due to 3'-methyl-4-dimethylaminoazobenzene or 2-acetylamino-fluorene: a comparative study. Carcinogenesis 3, 981-991, ibid.; in fact, no tumors of any type were reported in the 40 negative control rats in this study. And finally, the HID states that no liver tumors were observed in untreated controls in a 20-month study in male Leeds rats Flaks, B., 1978 Flaks, B., 1978. Effects of chronic oral dosing with quinine sulphate in the rat. Pathol Res Pract 163, 373-377, ibid.; once again, no tumors of any type were found among the control rats in this 1978 publication by Flaks. It appears that these investigators have never seen a tumor in any tissue or organ in a control group in any of their cancer studies using Leeds rats. This seems highly improbable.

The tumor types showing elevated RRs in a small number of epidemiologic studies can readily be detected in animal studies. For renal cell carcinoma, leukemia, lymphoma and liver tumors, there are one or more epidemiologic studies of acetaminophen that report an increase in relative risk. Importantly, there was no clear evidence that any of these tumor types was significantly increased in the NTP cancer bioassay or other scientifically valid animal studies. Yet, the animal models chosen in these animal studies are sensitive for detecting increases in the tumor types in the few epidemiologic studies that reported an increase in relative risk. For example, in its cancer bioassays, NTP demonstrated in rats and/or mice that 58 tested chemicals caused increased incidences of kidney tubular cell neoplasms, which are considered to be analogous or similar to renal cell carcinoma in humans (NTP, 2019). Similarly, 39 chemicals were shown to cause lymphoma or leukemia in rats and/or mice in NTP bioassays (NTP, 2019). And, 176 chemicals were demonstrated to produce liver tumors in rats and/or mice in NTP bioassays (NTP, 2019). If acetaminophen were responsible for increasing the risk of any of these tumor types, the NTP bioassay would be expected to show increases in these or other types of tumors and it does not.

Tumor promotion studies do not indicate that acetaminophen is clearly shown to cause cancer.

Six studies have evaluated the potential for acetaminophen to promote tumors initiated by known carcinogens. The vast majority of these studies demonstrate that acetaminophen, given in combination with an initiating carcinogen, had no effect on or, in several cases, reduced the incidence of various types of tumors. All of these tumor promotion studies were reviewed by IARC (1990).

Weisburger et al. (1973) reported tumor reductions in a study of acetaminophen examining tumor promotion in rats, mice, and hamsters. The study administered acetaminophen (11,000 ppm in the diet) and two known genotoxic carcinogens: N-2-fluorenylacetamide (FAA) or N-hydroxy-N-2-fluorenylacetamide (N-OHFAA) (Weisburger et al., 1973). When acetaminophen was given in combination with N-OHFAA, the incidence of bladder tumors was lower than the incidence observed when N-OHFAA was given alone among the male mice. In female rats, acetaminophen decreased the incidence of mammary tumors produced by FAA and, somewhat less, by N-OHFAA. In male and female hamsters, acetaminophen reduced the incidence of forestomach cancers induced by N-OHFAA.

Tsuda et al. (1984) reported that acetaminophen reduced liver tumors and enhanced renal cell adenomas (but not carcinomas) in male Fischer 344 rats given a known carcinogen (N-nitrosoethyl-N-hydroxyethylamine) in drinking water followed by administration of 13,000 ppm of acetaminophen in the diet for 29 weeks. The association of acetaminophen with the renal

tumors is unlikely since the nitrosamine itself produces renal lesions in 1-2 weeks, and the tumor incidences with the nitrosamine alone are highly variable.

Kurata et al. (1986) found no evidence of tumor promotion in male F344 rats given 13,000 ppm of acetaminophen in the diet for 32 weeks in combination with N-nitroso-N-(4-hydoxybutyl)amine, a carcinogen that causes bladder tumors. Hagiwara and Ward (1986) found no acetaminophen treatment-related tumor findings in male B6C3F1 mice given a N-nitrosodiethylamine as an initiator, followed two weeks later with 5000 or 10,000 ppm of acetaminophen in the diet for up to 70 weeks.

In a model of urinary bladder carcinogenesis in male F344 rats (Shibata et al., 1995), acetaminophen (8000 ppm in the diet for 35 weeks) did not significantly increase the incidences of tumors of the renal tubules, renal pelvis, ureter or urinary bladder when compared to the tumor-initiated control group treated with 0.1% N-nitrosodi(2-hydroxypropyl)amine in the drinking water and 3% uracil in the diet for the first four weeks (Shibata et al., 1995). Williams and Iatropoulos (1997) reported that, in male F344 rats, acetaminophen (up to 5000 ppm in the diet for 44 weeks) reduced the intestinal tumors initiated by administration of 3,2'-dimethyl-4-aminobiphenyl (DMAB) for 20 weeks.

Tumor initiation studies in rats with compromised liver function do not clearly show that acetaminophen causes cancer.

In addition to the tumor promotion studies described above, there are two tumor initiation studies of acetaminophen in rats with partial hepatectomies or with choline-induced liver damage: Hasegawa et al., (1988) and Maruyama et al. (1990). These studies have significant limitations for purposes of cancer hazard identification since: (1) the potential for tumor initiation was assessed by measuring the induction of certain liver foci (a potential precursor to tumors), not tumors, (2) the duration of exposure to acetaminophen was only twice a week by gavage for 5 weeks (Hasegawa et al., 1988) or daily in the diet for 25 weeks (Maruyama et al., 1990), and (3) the rats were co-exposed to tumorigenic substances (e.g., phenobarbital, 2-acetylaminofluoriene, carbon tetrachloride). Hasegawa et al. (1988) concluded: "These results indicate that [acetaminophen] possesses no tumor-initiating activity in the rat liver." Similarly, Maruyama et al. (1990) concluded: "Thus, these results indicate that [acetaminophen] does not possess significant carcinogenic activity in damaged rat liver." Both of these studies were reviewed by IARC when it concluded there was "inadequate evidence" of carcinogenicity in animals for acetaminophen.

4.3 Assessment of Exposure Coverage in Animal Carcinogenicity Studies

A Quantitative Systems Pharmacology/ Toxicology software package called DILIsym, that has been validated with acetaminophen for evaluating the population pharmacokinetics and dose response for liver injury, was used to estimate the acetaminophen and metabolite exposures in

humans under therapeutic, supratherapeutic and overdose conditions. The results of these simulations can be found in a separate companion document that has been provided as part of the supplementary materials.

The exposure analysis and simulations support the following conclusions:

- Acetaminophen mg/kg doses in the animal carcinogenicity studies are comparable to or higher than the human therapeutic, supratherapeutic, and overdose scenarios.
- The extent of NAPQI formation is much higher in mice and in a similar range in rats to humans under therapeutic, supratherapeutic, and overdose scenarios
- NAPQI formation in the mouse is comparable to or higher than under acute overdose scenarios.
- Acetaminophen induced minimal hepatotoxicity in the NTP study

Carcinogenicity Studies – Discussion and Conclusions

Overall, the acetaminophen animal carcinogenicity studies are reassuring

When the results of all the animal studies of acetaminophen are considered collectively, the findings are strongly reassuring. The overwhelming weight of the scientific evidence demonstrates that acetaminophen is not carcinogenic in animals at comparable or higher doses and NAPQI levels compared to humans. If acetaminophen posed a carcinogenic hazard to humans, it should have produced a clear and consistent signal of carcinogenicity in animals at a minimally toxic or lower dose, and it did not.

In the NTP cancer bioassay, the evidence of carcinogenicity was limited to "equivocal evidence of carcinogenic activity" in only one sex of one species, and the IARC Working Group specifically discounted that evidence as "not a treatment-related effect" (IARC, 1999). Given the concordance between known human carcinogens and findings in animal studies, the lack of carcinogenicity of acetaminophen in the NTP cancer bioassay and other reasonably well conducted cancer studies provides compelling evidence that acetaminophen is not carcinogenic. IARC determined correctly that there is "inadequate evidence" of carcinogenicity in animals, and there are no new long-term animal carcinogenicity studies or tumor promotion studies since IARC made its determination (IARC, 1999).

5 Genetic Toxicology Studies

The potential genotoxicity of acetaminophen became a topic of discussion in the 1990s, following several publications on the potential of NAPQI to bind to DNA, and results of various *in vitro* and *in vivo* genotoxicity studies. These data were reviewed in detail by a panel of European regulatory genetic toxicology experts (Bergman et al., 1996), who concluded that acetaminophen:

(1) Did <u>not</u> induce gene mutations either in bacteria or in mammalian cells *in vitro* and

(2) Did induce chromosomal damage *in vitro* in mammalian cells at high concentrations, and similar effects could occur *in vivo* at high dosages.

In the case of the latter, chromosome damaging effects were seen only at high concentrations that were cytotoxic to the test system. Since the detailed review of Bergman et al. (1996) several papers related to the potential genotoxicity of acetaminophen have been published. Since the focus of the Bergman paper was the relevance of genotoxicity at therapeutic doses, this assessment builds on the review of Bergman et al. (1996) and considers how the data previously reviewed, or any new data, impact the genotoxic hazard assessment for acetaminophen.

In terms of whether the genotoxicity data are indicative of a genotoxic or cancer hazard, specific focus is placed on:

(1) Relevance of genotoxic endpoints towards assessing potential cancer hazard

(2) Conditions of genotoxic effects and whether the type of damage is stable or persistent

(3) Likely modes of genotoxic action and relevance towards the carcinogenic process, and

(4) Strength of the weight of evidence (WoE) for genotoxic potential indicative of a carcinogenic hazard.

The objectives of this section are to provide an overview of important methodological considerations, a summary and analysis of the available genotoxicity data related to acetaminophen, the Mode of Action (MoA) for its cellular genetic toxicology effects at therapeutic, supratherapeutic and overdose exposures, and provide a WoE assessment on its carcinogenic hazard potential.

5.1 Methodological and Other Important Considerations for Assessment

In terms of WoE, genetic toxicology tests provide information with varying levels of relevance. For example, mutation endpoints are considered more important in determining potential risk than endpoints that are reversible (i.e., DNA breakage) or not associated with a known adverse effect (i.e., sister chromatid exchange). As recommended by Brusick et al. (2016) and Eastmond (2017), in a WoE approach, studies are evaluated based on quality, reproducibility and consistency, significance of the genetic alteration, phylogenetic relevance to humans, type (*in vivo* vs. *in vitro*, cell type, p53 status etc.), and relevance of the route of administration.

Genetic effects identified *in vivo* are generally considered more important than responses from *in vitro* tests and in particular than *in vitro* tests in p53-deficient cell lines that are susceptible to misleading positive results, or in non-mammalian systems (other than the Ames test) for which no recommended testing guidelines are available. As stated in the recent OECD Genetic toxicology Guidance Document (OECD, 2015) "*assays conducted in mammalian cells are preferred because they are considered more relevant*". **Therefore, results in non-mammalian test**

systems such as mussels, insects, plants, yeasts and acellular systems should not be considered as being as relevant (i.e. not be given the same weight) as results from mammalian systems and the Ames test. Also, data from indicator tests such as DNA strand breaks, or from endpoints such as sister chromatid exchanges (SCE), where the biological relevance of the effects is not understood and OECD guidelines have been deleted, should contribute negligible or very low weight.

As stated in the OECD Genetic Toxicology Guidance document (OECD, 2015) "*Indicator tests detect primary DNA damage (i.e. the first in the chain of events leading to a permanent change), but not the consequences of this genetic damage. They are called indicator tests because the measured endpoint does not always lead to mutation, a change that can be passed on to subsequent generations*", and "*When evaluating potential genotoxicants, more weight should be given to the measurement of permanent DNA changes than to DNA damage events that are reversible*". Most regulatory bodies therefore rely on a set of core endpoints that are known or suspected to be directly responsible for neoplastic initiation in somatic cells or alteration of the genetic information in germ cells (EFSA, 2011; ICH, 2011; Kirkland et al., 2011). The endpoints given the greatest weight include chromosomal aberration (CA) or micronucleus (MN) formation *in vivo* and gene mutation *in vitro* in bacteria (Ames) or *in vivo*.

The published studies on the genotoxicity of acetaminophen have therefore been considered in terms of their weighted contribution to an overall indication of genotoxic hazard employing the method for weighting genetic toxicology test methods described in Brusick et al., (2016) (see Table 6 - Table 15). A test's weighted contribution was determined based on information regarding its relevance and reliability, predictivity, the endpoint's reversibility, susceptibility to false responses, and its mechanism's role in initiation of malignancy.

Table 6: Description of Weighted Contribution Categories (adapted from: (Brusick et al., 2016)

Weight Descriptor	Definition
Negligible weight	The endpoint is not linked to any adverse effect relevant to genetic or carcinogenic hazard/ risk (e.g., SCE).
Low weight	The end point is indicative of primary DNA damage, not directly linked to mechanisms of tumorigenicity (e.g., DNA breakage or computer-based SAR results), or the endpoints are evaluated in non-mammalian test systems (other than the Ames test).
Moderate weight	The endpoint may be: (a) only potentially relevant to tumor initiation, (b) subject to secondary effects (cytotoxicity), (c) subject to threshold dependent mechanisms of induction (aneugens) or (d) the test system exhibits a high rate of false responses with respect to carcinogenicity predictivity (e.g., mammalian cell *in vitro* clastogenicity and gene mutation tests, particularly in p53-deficient cells).
High weight	The endpoint is one that has been demonstrated to play a critical role in the process of tumorigenicity (gene mutation in bacteria [Ames test] or *in vivo*, chromosome aberrations or micronuclei *in vivo*).

Bergman et al. (1996) concluded that acetaminophen did induce chromosomal damage (clastogenicity) *in vitro* in mammalian cells at high cytotoxic concentrations, and similar effects could occur *in vivo* at high toxic doses. As will be seen later, more recent publications confirm this. However, to understand the relevance of such genotoxic activity it is important to understand the relationship between clastogenicity and cytotoxicity, and the impact of the different cell types in which clastogenicity has been observed.

As far back as 1981, (Heddle and Salamone) concluded that the types of "*aberrations that involve a rearrangement of gene order rather than a direct loss of a gene are not cell lethal events and, hence, are not contributors to cellular toxicity. In contrast, those aberrations that lead directly to the loss of a section of genetic information are usually cell lethal events and do contribute directly to cellular toxicity*". Thus, chromosome breakage (clastogenicity) and cell death are inextricably linked. It is now accepted that positive chromosomal aberration or micronucleus results at high levels of cytotoxicity could be a misleading indicator of the genotoxic potential of a test substance (Kirkland, 1992). This is prominently discussed in ICH S2(R1) (ICH, 2011):

- As cytotoxicity increases, mechanisms other than direct DNA damage by a compound or its metabolites can lead to 'positive' results that are related to cytotoxicity and not genotoxicity.
- In cytogenetic assays, even weak clastogens that are known to be carcinogens are positive without exceeding a 50% reduction in cell counts. On the other hand, compounds that are not DNA damaging, mutagenic or carcinogenic may induce chromosome breakage but at toxic concentrations.

Hence, OECD guidelines now recommend careful control of cytotoxicity in genotoxicity tests and urge caution in interpreting positive results only observed at levels of cytotoxicity close to or above the recommended maximum.

It is well established that induction of chromosome breaks will lead to cell death. However, chromosome breaks can rejoin leading to stable rearrangements that may be inherited by daughter cells after division, and could pre-dispose to indicate a mutagenic or carcinogenic hazard. Such rearrangements would need to be induced at low levels of cytotoxicity, such that affected cells would survive. Stable chromosome rearrangements are not usually scored in chromosomal aberration tests, because it requires specialized banding techniques, but the induction of unstable rearrangements (complete and incomplete inter- and intra-chromatid exchanges) are an indication of the potential to induce stable rearrangements.

It is also now known that p53-deficient rodent cells are more likely to produce "misleading" positive results (i.e. with substances that are not genotoxic or carcinogenic *in vivo*), particularly for clastogenicity, than p53-competent human cells (Fowler et al., 2012). It is therefore not uncommon to find positive clastogenicity results in p53-deficient Chinese hamster cell lines (CHO, CHL, V79) with substances that are negative in p53-competent human lymphocytes or human

TK6 cells, or to find positive results at lower concentrations in Chinese hamster cells than in human cells. Thus, more weight should be given to results in p53-competent human cells than in p53-deficient hamster cells. Taking these two aspects together, since only chromatid and chromosome breaks (but no unstable rearrangements) were induced by acetaminophen in p53-competent human cells, and only under cytotoxic conditions, the cells will not survive and therefore this type of genotoxic damage does not indicate a clear genotoxic or carcinogenic hazard.

5.2 Assessment of Genotoxicity Studies

The genetic toxicology studies are summarized in Table 7 - Table 15, along with a corresponding weight of evidence assessment of the results.

5.2.1 Mutagenicity Studies

Acetaminophen has been tested for mutagenicity under *in vitro* conditions utilizing both bacterial and mammalian cell systems as well as under *in vivo* conditions. Overall, the evidence indicates that acetaminophen does not have the potential to induce point mutations in bacteria, in *in vitro* mammalian systems or *in vivo* and is therefore not mutagenic.

Table 7: Overview of relevant non-mammalian in vitro mutagenicity studies.

Study	HID Reported Result	WOE Hazard Assessment	Weight
Bacterial mutagenicity (Ames)			
(Burke et al., 1994)	Negative	Supports no hazard	High
(Camus et al., 1982)	Negative	Supports no hazard	High
(Dybing et al., 1984)	Negative	Supports no hazard	High
(Jasiewicz and Richardson, 1987)	Negative	Supports no hazard	High
(Haworth et al., 1983)	Negative	Supports no hazard	High
(Imamura et al., 1983)	Negative	Supports no hazard	High
(King et al., 1979)	Negative	Supports no hazard	High
(Oldham et al., 1986)	Negative	Supports no hazard	High
(NTP, 1993)	Negative	Supports no hazard	High
(Wirth et al., 1980)	Negative	Supports no hazard	High
(Martinez et al., 2000)	Negative	Supports no hazard	High

**Note the "Weight" column reflects the weight that should be given to the study type in the Weight of Evidence Hazard Assessment based on information regarding its relevance and reliability, predictivity, the endpoint's reversibility, susceptibility to false responses, and its mechanism's role in initiation of malignancy.*

Table 8: Overview of in vitro mutagenicity studies in mammalian cells.

Study	Cell Type	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations or Concerns
Mammalian cell gene mutation (Tk locus, 6TG, HPRT)					

Study	Cell Type	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations or Concerns
(Clements, 1992); referred by HID as Hazleton Microtest (1992), cited Muller and Kasper (1995)	Mouse/ lymphoma cells	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Moderate	Positive in the range 3.3-33 mM, exceeding the recommended upper limit, and therefore likely to cause physiological disruption and stress-related damage. Report was unavailable for review; details included were from Bergman et al. (1996). No conclusions could be drawn on the type of damage that paracetamol caused since the size of the mutant colonies was reportedly not analyzed (Bergman et al., 1996). It is possible that small increases in mutation frequencies at high concentrations in this assay can be attributed to chromosomal damage rather than point mutations (Bergman et al. 1996).
(Shimane, 1985)	Chinese hamster lung (V79)	Weakly positive	Multiple factors call into question relevance and use of study for hazard assessment	Low-Moderate	Weakly positive at 100 or 400 µg/mL, depending on exposure period. However, mutant frequencies were low and may have been within historical control range. Also, there was no dose-response with 48 hr exposure. Moreover, V79 cells are p53-deficient and susceptible to misleading positive responses not found in p53-competent, genomically stable (e.g. primary human) cells.
Mammalian cell gene mutation (Oubain resistance)					
(Patierno et al., 1989)	Mouse Fibroblast; C3H/10T1/2 clone 8	Negative	Supports no hazard	Moderate	None
(Sasaki, 1986; Sasaki et al., 1983)	Chinese Hamster Ovary (CHO-K1)	Negative	Supports no hazard	Moderate	None
(Shimane, 1985)	Chinese hamster lung (V79)	Weakly positive	Multiple factors call into question relevance and use of study for hazard assessment	Low-Moderate	Weakly positive at 100 or 200 µg/mL, depending on exposure period. However, mutant frequencies were low and may have been within historical control range. Also, there was no dose-response with 24 hr exposure. Moreover, V79 cells are p53-deficient and susceptible to misleading positive responses not found in p53-competent, genomically stable (e.g. primary human) cells.

**Note the "Weight" column reflects the weight that should be given to the study type in the Weight of Evidence Hazard Assessment based on information regarding its relevance and reliability, predictivity, the endpoint's reversibility, susceptibility to false responses, and its mechanism's role in initiation of malignancy.*

Table 9: Overview of mammalian in vivo mutagenicity studies

Study	Animal Model	Route of Exposure	HID Reported Result	WOE Hazard Assessment	Weight	Considerations or Concerns
Mutagenicity						
(Matsushita et al., 2013)	Rat (F344/NSlc gpt delta transgenic); liver	p.o.	Negative	Supports no hazard	High	None
(Kanki et al., 2005)	Rat (Sprague-Dawley gpt delta transgenic); liver	p.o.	Negative	Supports no hazard	High	None

Pig-a/Pig-ret						
(Suzuki et al., 2016)	Rat (Sprague-Dawley); reticulocytes	p.o.	Negative	Supports no hazard	High	None
(Van der Leede et al., submitted for publication on 10 Oct 2019, Manuscript number EMM-19-0142)	Rat (Sprague-Dawley); total red blood cells and reticulocytes	p.o.	Not Reviewed	Supports no hazard	High	Negative

5.2.2 Clastogenicity Studies

In reliable guideline assays (micronucleus test and chromosomal aberration assay), negative results or irrelevant results were observed within the hazard framework. In well conducted human studies, negative results were observed at the administered therapeutic doses without any cytotoxic effects. Further, negative results were obtained in self-poisoned persons even when cytotoxicity was reported. The weight of evidence suggests that clastogenic effects are not observed unless higher concentrations are reached that affect cellular processes and induce cytotoxicity, which are not expected to lead to viable cells containing stable genetic damage that would be indicative of a clear genotoxic hazard in humans. For in vitro clastogenicity studies, the ICH recommends 1 mM (or 151.16 µg/ml in case of acetaminophen) as the maximum concentration to be tested beyond which cytotoxicity might be expected.

Table 10: Overview of in vitro clastogenicity and SCE studies in mammalian cells

Study	Cell Type	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations or Concerns
Micronucleus (MN) test					
(Ibrulj et al., 2007)	Human Lymphocyte	Weakly positive	Responses not significant. Supports no hazard	Moderate	MN frequencies in acetaminophen-treated cultures (3.5, 5.5, 5.75) were similar to control (5), and there were no significant differences. It is not correct to call this "weakly positive".
(Simkó et al., 1998)	Human Amniotic Fluid (AFC)	Positive	Multiple factors call into question relevance and interpretation of study for hazard assessment	Moderate	MN frequencies in controls were high, and background data for AFC cells not available so it is not clear whether this was normal. Also, slides were not coded so potential bias cannot be excluded. Time and concentration dependence not consistent for cytotoxicity. Inconsistent results reported for MN formation across plots for similar concentrations. Results questionable.
(Muller-Tegethoff et al., 1995)	Primary Rat Hepatocyte	Negative	Supports no hazard	Moderate	Negative
(Dunn et al., 1987)	Rat Kidney Fibroblast (NRK-49F)	Positive	Use of excessive concentrations indicates result is not biologically relevant	Moderate	Positive, but only at very high concentrations (10 & 20 mM), exceeding recommended limit, likely to cause physiological disruption & stress-related damage. Such results should be considered irrelevant and discounted.

Study	Cell Type	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations or Concerns
(Matsushima et al., 1999)	Chinese Hamster Lung (CHL/IU)	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Low-Moderate	CHL cells are p53-deficient and susceptible to misleading positive responses not found in p53-competent, genomically stable (e.g. primary human) cells. Cytotoxicity not assessed. Results considered low-moderate weight.
Chromosomal aberrations (scored as %)					
(NTP, 1993)	Chinese Hamster Ovary (CHO)	Positive/weakly positive	Use of excessive concentrations and p53-deficient cells indicates result is not biologically relevant	Low-Moderate	Positive, particularly after 20-hr treatment -S9, but only tested above 1 mM, exceeding recommended limit, likely to cause physiological disruption & stress-related damage. Such results should be considered irrelevant and discounted. No concurrent measure of cytotoxicity. Also, CHO cells are p53-deficient and susceptible to misleading positive responses not found in p53-competent, genomically stable (e.g. primary human) cells.
(Shimane, 1985)	Chinese Hamster Lung (V79)	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Low-Moderate	Positive at concentrations ranging from 25-200 µg/mL in the absence & presence of metabolic activation. No concurrent measure of cytotoxicity, but probably <50% in this range, based on other data in the paper. Unclear if slides were coded. However, V79 cells are p53-deficient and susceptible to misleading positive responses not found in p53-competent, genomically stable (e.g. primary human) cells.
(Muller et al., 1991)	Chinese Hamster Lung (V79)	Positive	Use of excessive concentrations and p53-deficient cells indicates result is not biologically relevant	Low-Moderate	Positive, particularly after 6 & 12-hr continuous treatments in the absence of metabolism, or when co-cultured with hepatocytes, but mainly at concentrations >1 mM, exceeding recommended limit, likely to cause physiological disruption & stress-related damage. Such results should be considered irrelevant and discounted. No concurrent measure of cytotoxicity. Moreover, V79 cells are p53-deficient and susceptible to misleading positive responses not found in p53-competent, genomically stable (e.g. primary human) cells.
Chromosomal aberrations (scored as #)					
(Ibrulj et al., 2007)	Human Lymphocyte	Positive	Use of excessive concentrations indicates result is not biologically relevant	Moderate	Positive only at 1.3 mM, exceeding recommended limit. Such results should be considered irrelevant and discounted.
(Hongslo et al., 1991)	Human Lymphocyte	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Moderate	If gaps are excluded, weakly positive at 0.75 mM, and positive at 1.5 & 3 mM. However, abnormal chromosome morphology at 3 mM. Slides were coded but no concurrent measure of cytotoxicity, so chromosome breaks could be associated with toxic effects.
(Watanabe, 1982)	Human Lymphocytes	Positive	Use of excessive and cytotoxic concentrations indicates result is not biologically relevant	Moderate	Weakly positive at 200 µg/mL, positive at 400 & 600 µg/mL after 72hrs treatment, but all these concentrations exceed the recommended 1 mM limit, and all induced >50% mitotic inhibition, so aberrations could be due to severe cytotoxicity. Also, gaps were

Study	Cell Type	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations or Concerns
					included, and this is not normal convention. Unclear whether slides were coded, so potential scorer bias cannot be excluded.
(Hongslo et al., 1990)	Mouse Mammary (TA3H)	Weakly positive	Use of excessive and cytotoxic concentrations indicates result is not biologically relevant	Moderate	Positive, but only at concentrations >1 mM, inducing >50% reduction in cell growth, both of which exceed the recommended limits. Such results are therefore irrelevant and should be discounted. Moreover, there are no background data on these mouse mammary tumor cells, and there p53 status and genomic stability are unknown.
(Sasaki et al., 1980)	Chinese Hamster (Don-6)	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Low-Moderate	Positive at 75 & 151 µg/mL after 26-30 hr continuous treatment. No concurrent measure of cytotoxicity. No background data on chromosome damage in Don-6 cells, which are probably also p53-deficient and susceptible to misleading positive responses not found in p53-competent, genomically stable (e.g. primary human) cells.
(Sasaki, 1986; Sasaki et al., 1983)	Chinese Hamster (CHO-K1)	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Low-Moderate	Same data in both papers. If gaps excluded, positive at 70 & 100 µg/mL with 24 hrs treatment. Not stated that slides were coded, so potential scorer bias cannot be excluded. No concurrent measure of cytotoxicity. CHO-K1 cells are p53-deficient and susceptible to misleading positive responses not found in p53-competent, genomically stable (e.g. primary human) cells.
(Ishidate et al., 1978; Ishidate, 1983)	Chinese Hamster Lung Fibroblast	Positive	Publications unavailable for review	Low-Moderate	Publications unavailable for review; unclear whether cytotoxicity was assessed; altered p53 status in CHL cells.
(Ishidate et al., 1988)	Chinese Hamster Lung (V79)	Positive	This is a review of previously published data	Moderate	This is a review of previously published data. The CHO data are from Sasaki et al. 1983, human lymphocyte data from Watanabe 1982 (commented above), and CHL data from Ishidate 1987. There are no V79 data in this paper.
Sister chromatid exchange (SCE)					
(Hongslo et al., 1991)	Human lymphocyte	Positive	Relevance of SCE endpoint is not understood, so not biologically relevant	Negligible	The SCE assay was deleted as a test guideline due to a poor understanding of the mechanisms of action that can be detected by the test and the biological relevance of SCE is not understood. SCE results therefore contribute negligible weight to genotoxic hazard assessment. No concurrent measure of cytotoxicity.
(Hongslo et al., 1990)	Mouse (TA3H)	Positive	Relevance of SCE endpoint is not understood, so not biologically relevant	Negligible	The SCE assay was deleted as a test guideline due to a poor understanding of the mechanisms of action that can be detected by the test and the biological relevance of SCE is not understood. SCE results therefore contribute negligible weight to genotoxic hazard assessment. Positive, but only at concentrations >1 mM, inducing >50% reduction in cell growth, both of which exceed the recommended limits. Such results are

Study	Cell Type	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations or Concerns
					therefore irrelevant and should be discounted. Moreover, there are no background data on these mouse mammary tumor cells, and there p53 status and genomic stability are unknown.
(Holme et al., 1988)	Chinese Hamster Lung (V79)	Positive	SCE endpoint (relevance not understood), and use of p53-deficient cells indicates result is not biologically relevant	Negligible	The SCE assay was deleted as a test guideline due to a poor understanding of the mechanisms of action that can be detected by the test and the biological relevance of SCE is not understood. SCE results therefore contribute negligible weight to genotoxic hazard assessment. V79 cells are p53-deficient and susceptible to misleading positive responses not found in p53-competent, genomically stable (e.g. primary human) cells.
(Hongslo et al., 1988)	Chinese Hamster Lung (V79)	Positive	SCE endpoint (relevance not understood), and use of p53-deficient cells indicates result is not biologically relevant	Negligible	The SCE assay was deleted as a test guideline due to a poor understanding of the mechanisms of action that can be detected by the test and the biological relevance of SCE is not understood. SCE results therefore contribute negligible weight to genotoxic hazard assessment. V79 cells are p53-deficient and susceptible to misleading positive responses not found in p53-competent, genomically stable (e.g. primary human) cells.
(Shimane, 1985)	Chinese Hamster Lung (V79)	Positive	SCE endpoint (relevance not understood), and use of p53-deficient cells indicates result is not biologically relevant	Negligible	The SCE assay was deleted as a test guideline due to a poor understanding of the mechanisms of action that can be detected by the test and the biological relevance of SCE is not understood. SCE results therefore contribute negligible weight to genotoxic hazard assessment. V79 cells are p53-deficient and susceptible to misleading positive responses not found in p53-competent, genomically stable (e.g. primary human) cells.
(NTP, 1993)	Chinese Hamster Ovary	Positive/ weakly positive	SCE endpoint (relevance not understood), and use of p53-deficient cells indicates result is not biologically relevant	Negligible	The SCE assay was deleted as a test guideline due to a poor understanding of the mechanisms of action that can be detected by the test and the biological relevance of SCE is not understood. SCE results therefore contribute negligible weight to genotoxic hazard assessment. CHO cells are p53-deficient and susceptible to misleading positive responses not found in p53-competent, genomically stable (e.g. primary human) cells.
(Sasaki, 1986)	Chinese Hamster Ovary (CHO-K1)	Positive	SCE endpoint (relevance not understood), and use of p53-deficient cells indicates result is not biologically relevant	Negligible	The SCE assay was deleted as a test guideline due to a poor understanding of the mechanisms of action that can be detected by the test and the biological relevance of SCE is not understood. SCE results therefore contribute negligible weight to genotoxic hazard assessment. No concurrent measure of cytotoxicity. CHO cells are p53-deficient and susceptible to misleading positive responses not found in p53-competent, genomically stable (e.g. primary human) cells.

**Note the "Weight" column reflects the weight that should be given to the study type in the Weight of Evidence Hazard Assessment based on information regarding its relevance and reliability, predictivity, the endpoint's reversibility, susceptibility to false responses, and its mechanism's role in initiation of malignancy.*

Table 11: Overview of in vivo clastogenicity, aneugenicity and SCE studies

Study	Animal Model	Route of Exposure	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations or Concerns
Micronucleus test						
(Marshall, 1993); referred by HID as Hazleton Microtest (1993), as cited by Muller and Kasper (1995)	Rat	p.o.	Weakly positive	Multiple factors call into question relevance and use of study for hazard assessment	High	Results confounded by clear bone marrow toxicity due to very high dose. Weak (mainly 2-3-fold but up to 7-fold) positive response in rats 24 and 40 hrs after oral dosing of 3 doses of 900 mg/kg (4-hourly intervals). Report not available. Results taken from Bergman et al. (1996)
(King et al., 1979)(oral)	Mouse	p.o.	Negative	Supports no hazard	High	None
(King et al., 1979) (i.p.)	Mouse	i.p.	Negative	Supports no hazard	High	None
(Sicardi et al., 1991)	Mouse	i.p.	Positive	Multiple factors call into question relevance and use of study for hazard assessment	High	Weak positive (2.4-fold) in mice dosed i.p. with 100 or 150 mg/kg but no response at 200 mg/kg. No dose-response and no cytotoxicity data reported. Results questionable.
(Markovic et al., 2013)	Mouse (dam)	i.p.	Weakly positive	Responses not significant. Supports no hazard	High	Weak positive (3.25-fold) in pregnant BALB/c mice 48 hrs after i.p. dosing at 60 mg/kg on days 12 and 14 of pregnancy. However, the increase was not statistically significant. Slides were not coded so potential scorer bias cannot be excluded.
(Markovic et al., 2013)	Mouse (newborn)	i.p. (dams)	Positive	Multiple factors call into question relevance and use of study for hazard assessment	High	Treatment of pregnant mice as above. Weak positive (2.28-fold) in the blood of the pups that was statistically significant (p<0.0.5), but associated with evidence of oxidative stress and hepatotoxicity. Slides were not coded so potential scorer bias cannot be excluded.
(Van der Leede et al., submitted for publication on 10 Oct 2019, Manuscript number EMM-19-0142)	Rat (Sprague-Dawley); reticulocytes	p.o.	Not Reviewed	Supports no hazard	High	Statistically significant increases in MN in reticulocytes after 1 month of dosing at 500 and 1000 mg/kg/day were attributed to rebound erythropoiesis in response to marked hematotoxicity (severe bone marrow toxicity was seen 4 days after the start of dosing), and therefore the increased MN were concluded to be due to a non-genotoxic mode of action.
Chromosomal aberrations (scored as %)						
(Reddy and Subramanyam, 1985)	Mouse	p.o.	Negative	Supports no hazard	High	None
(Giri et al., 1992)	Mouse	i.p.	Positive	Multiple factors call into question relevance and interpretation of study for hazard assessment	High	Positive (2-4-fold increases) in bone marrow of mice given single i.p. doses of 200 or 400 mg/kg, accompanied by dose-response. Slides were coded and gaps were excluded. Mitotic index showed no bone marrow toxicity. Based on literature (Nayak et al., 2011), these doses would be expected to be near or above hepatotoxicity.

Study	Animal Model	Route of Exposure	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations or Concerns
Chromosomal aberrations (scored as #)						
(Reddy, 1984)	Mouse	p.o.	Positive	Multiple factors call into question relevance and use of study for hazard assessment	High	Same doses and dosing schedule as in Laxminarayana et al., 1980 (same laboratory), but bone marrow sampled at 24, 48 & 72 hrs. If gaps and polyploid cells excluded (which is recommended practice), very small increases in breaks (max. 3/250 cells), but vehicle controls had 0 or only 1 break at all sampling times, which is unusually low. Increases from 0 or 1 in controls to 2 or 3 breaks in treated groups would not be considered biologically relevant. Moreover, no measure of cytotoxicity and slides not coded, so potential scorer bias cannot be excluded. Results highly questionable.
(Severin and Beleuta, 1995) (oral)	Mouse	p.o.	Positive	Multiple factors call into question relevance and use of study for hazard assessment	High	Mice given oral doses of 3 doses of 800 mg/kg at 4-hourly intervals. Clear increase in breaks particularly at 24 hrs after dosing. However, no measure of cytotoxicity. Also, slides not coded, so potential scorer bias cannot be excluded. Hepatotoxicity was likely induced based on previous literature (Uchida et al. 2017).
(Severin and Beleuta, 1995) (i.p.)	Mouse	i.p.	Positive	Multiple factors call into question relevance and use of study for hazard assessment	High	Mice given single i.p. doses of 100, 200 or 400 mg/kg. Increases in breaks at all doses that were 2-9 fold below the positive control and decreased by 72 hours. However, no measure of cytotoxicity. Also, slides not coded, so potential scorer bias cannot be excluded. Based on other studies, 250-300 mg/kg caused severe hepatotoxicity in mice and therefore, the small changes in CA could be due to toxic response (Nayak et al., 2011).

Study	Animal Model	Route of Exposure	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations or Concerns
Aneuploidy						
(Tsuruzaki et al., 1982)	Rat (embryos)	p.o.	Positive	Multiple factors call into question relevance and use of study for hazard assessment	High	Female rats dosed with 500 or 1000 mg/kg from 2 weeks prior to mating until 11.5 days after mating. Rat fetuses showed no increase in structural aberrations. Increases in aneuploidy were reported for both dose groups, but since it is not clear (Japanese paper) how the slides were prepared, it is not known whether the chromosome loss/gain was due to hypotonic treatment of a true effect on the spindle. No measure of cytotoxicity. Not known if slides were coded, so potential scorer bias cannot be excluded. Results questionable.
Sister chromatid exchange						
(Giri et al., 1992)	Mouse	i.p.	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Negligible	The SCE assay was deleted as a test guideline due to a poor understanding of the mechanisms of action that can be detected by the test and the biological relevance of SCE is not understood. SCE results therefore contribute negligible weight to genotoxic hazard assessment. Based on other studies the top dose at least would be expected to be hepatotoxic.

**Note the "Weight" column reflects the weight that should be given to the study type in the Weight of Evidence Hazard Assessment based on information regarding its relevance and reliability, predictivity, the endpoint's reversibility, susceptibility to false responses, and its mechanism's role in initiation of malignancy.*

Table 12: Overview of human clastogenicity and SCE studies

Study	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations or Concerns
Micronucleus test				
(Kocišová and Šrám, 1990)	Weak increased response	Responses not significant. Supports no hazard	High	Same treatment and sampling as in Topinka et al. (1989). The frequencies of MN in lymphocytes at all sampling times were similar to the pre-dose frequency, and not significantly different. Therefore, the result is negative.
(Šrám et al., 1990; Topinka et al., 1989)	Increased response ($p<0.01$)	Multiple factors call into question relevance and use of study for hazard assessment	High	2 studies, probably in the same group of volunteers, one with ascorbic acid, the other without. In both cases MN in buccal cells increased transiently at 72 hrs but not at earlier or later sampling times. The MN frequencies were low, and, based on other publications, probably within the normal range.

Study	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations or Concerns
Chromosomal aberrations				
(Hantson et al., 1996)	No effect	Supports no hazard	High	Negative, even after suicidal doses.
(Hongslo et al., 1991)	Increased response (p<0.1)	Responses excluding gaps small and probably not significant. Supports no hazard	High	When gaps were excluded (as is normal practice) the increase in CA (chromatid breaks) in lymphocytes was small (from 2.16% to 3.43%, not analyzed for statistical significance), and not considered biologically relevant. As many of the six volunteers showed no change or a decrease as those that showed an increase in the levels of aberrant cells.
(Kocišová et al., 1988)	Increased response (p<0.05)	Multiple factors call into question relevance and use of study for hazard assessment	High	2 studies in the same group of volunteers, one with ascorbic acid, the other without. In both cases CA (only chromatid breaks) in lymphocytes increased transiently but at different times, and were normal either before and after, or after the increase. Also, some individuals showed an increase in CA whereas others did not or showed a decrease. Moreover, individuals who had shown a comparatively large increase in chromatid break frequency in the first study showed a small increase or even a decrease in the second study, and vice versa. It is therefore highly likely the increases in CA were due to chance.
(Kirkland et al., 1992)	No effect	Supports no hazard	High	Negative, Double blind and placebo controlled
Sister chromatid exchange				
(Kirkland et al., 1992)	Negative	Did not study SCE	Negligible	SCE were not analyzed.
(Hongslo et al., 1991)	Increased response (p<0.05)	Relevance of SCE endpoint not understood. Result not biologically relevant	Negligible	The biological relevance of SCE is not understood, and the OECD guideline has been deleted. SCE results therefore contribute negligible weight to genotoxic hazard assessment.

**Note the "Weight" column reflects the weight that should be given to the study type in the Weight of Evidence Hazard Assessment based on information regarding its relevance and reliability, predictivity, the endpoint's reversibility, susceptibility to false responses, and its mechanism's role in initiation of malignancy.*

5.2.3 DNA Damage Studies

DNA damage studies are considered "indicator tests" by OECD (OECD, 2015) since they do not measure stable genetic damage. DNA damage may be reversible or may be lethal and not lead to mutations. The weight of evidence demonstrates that acetaminophen does not cause DNA damage in reliable, well-controlled test systems in the absence of cytotoxicity or hepatotoxicity, and such results are consistent with previous studies that genotoxicity resulting from acetaminophen exposure only occurs at high, toxic doses which are not likely to result in viable cells containing stable genetic damage that would be indicative of a clear genotoxic hazard.

Table 13: Overview of in vitro DNA damage/repair studies in mammalian cells

Study	Cell Type	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations or Concerns
DNA damage					
(Andersson et al., 1982; Nordenskjold and Moldeus, 1983)	Cultured skin fibroblasts	Negative	Supports no hazard	Low	Negative; DNA strand breaks are an "indicator test" and contribute only low weight to hazard assessment
(Jetten et al., 2014)	Human liver tissue	Positive	Use of excessive and cytotoxic concentrations indicates result is not biologically relevant	Low	The concentrations of acetaminophen used to treat the liver slices ranged from 2.5-10 mM, and therefore exceeded recommended upper limits for mammalian cell tests. Such results are irrelevant and should be discounted. BMD analysis suggested some donors showed increased comets, but these were only seen at high cytotoxicity (>50%). DNA strand breaks are an "indicator test" and contribute only low weight to hazard assessment.
(Bandi et al., 2014)	Human peripheral blood mononuclear cells	Positive	Use of excessive concentrations indicates result is not biologically relevant	Low	Positive for comets and γH2AX, but data only given for 10 mM, which exceeds the recommended upper limit for mammalian cell tests. Viability was reduced to approximately 50% at this concentration. Such results are irrelevant and should be discounted. DNA strand breaks are an "indicator test" and contribute only low weight to hazard assessment.
(Dybing et al., 1984)	Rat/ Reuber hepatoma cells	Negative	Supports no hazard	Low	Used alkaline elution method for which there is no OECD guideline, so no recommendations for what constitutes an adequate test or how to interpret the results. 10 mM acetaminophen did not induce strand breaks, but NAPQI induced dose-related damage, but 36-100% cytotoxicity was induced across the range. DNA strand breaks are an "indicator test" and contribute only low weight to hazard assessment.
(Sasaki, 1986)	Hamster/ ovary cells (CHO-K1)	Weakly positive	Use of a non-guideline method at excessive concentrations indicates result is not biologically relevant	Low	Used alkaline elution method for which there is no OECD guideline, so no recommendations for what constitutes an adequate test or how to interpret the results. Weak positive response at 5000 µg/mL, exceeding the recommended upper limit for mammalian cell tests. Such results are irrelevant and should be discounted. No details are given, so the extent of cytotoxicity at these high concentrations is not clear. DNA strand breaks are an "indicator test" and contribute only low weight to hazard assessment.
(Hongslo et al., 1988)	Chinese Hamster Lung (V79) cells	Weakly positive	Use of a non-guideline method at excessive concentrations indicates result is not biologically relevant	Low	Used alkaline elution method for which there is no OECD guideline, so no recommendations for what constitutes an adequate test or how to interpret the results. Weak positive response at 3 & 10 mM, exceeding the recommended upper limit for mammalian cell tests. Such results are irrelevant and should be discounted. Cytotoxicity as measured by colony forming ability was only slight. However, V79 cells are p53-deficient and susceptible to misleading

Study	Cell Type	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations or Concerns
					positive responses not found in p53-competent, genomically stable (e.g. primary human) cells. DNA strand breaks are an "indicator test" and contribute only low weight to hazard assessment.
Unscheduled DNA synthesis (UDS)					
(Binkova et al., 1990)	Peripheral blood lymphocytes	Weakly positive	Use of a non-guideline method and non-significant responses indicates not biologically relevant.	Low	UDS was measured by scintillation counting, which is not the recommended method, and can be susceptible to artifacts. Slight increase in UDS over a wide concentration range, but seemingly not statistically significant. UDS is an "indicator test" and considered low weight.
(Dybing et al., 1984)	Mouse hepatocytes	Significantly increased	Use of a non-guideline method and excessive concentrations indicates not biologically relevant	Low	UDS was measured by scintillation counting, which is not the recommended method, and can be susceptible to artifacts. UDS was induced at 5 mM and above, which exceeds the recommended upper limit for mammalian cell tests. Such results are irrelevant and should be discounted. NAPQI did not induce UDS up to 0.25 mM. In any case, UDS is an "indicator test" and considered low weight.
(Holme and Søderlund, 1986)	Mouse hepatocytes	Significantly increased	Use of a non-guideline method and excessive, cytotoxic concentrations indicates not biologically relevant	Low	UDS was measured by scintillation counting, which is not the recommended method, and can be susceptible to artifacts. UDS was induced at 5 mM and above, which exceeds the recommended upper limit for mammalian cell tests. Such results are irrelevant and should be discounted. Also, cytotoxicity was around 50% or higher at these concentrations. UDS is an "indicator test" and considered low weight.
(Holme and Søderlund, 1986)	Rat hepatocytes	Slightly increased	Use of a non-guideline method and excessive concentrations indicates not biologically relevant	Low	UDS was measured by scintillation counting, which is not the recommended method, and can be susceptible to artifacts. UDS was slightly increased at 2.5 mM and above, which exceeds the recommended upper limit for mammalian cell tests. Such results are irrelevant and should be discounted. UDS is an "indicator test" and considered low weight.
(Milam and Byard, 1985)	Rat hepatocytes	No effect	Supports no hazard	Low	UDS was measured by scintillation counting, which is not the recommended method, and can be susceptible to artifacts. No induction of UDS at 3 & 7 mM. UDS is an "indicator test" and considered low weight.
(Sasaki, 1986)	Rat hepatocytes	Significantly reduced	Multiple factors call into question relevance and use of study for hazard assessment	Low	Significant decrease in nuclear granules probably due to toxicity. UDS is an "indicator test" and considered low weight.
(Holme and Søderlund, 1986)	Hamster hepatocytes	Significantly reduced	Multiple factors call into question relevance and use of study for hazard assessment	Low	UDS was measured by scintillation counting, which is not the recommended method, and can be susceptible to artifacts. The decrease in UDS is probably due to cytotoxicity. UDS is an "indicator test" and considered low weight.
(Holme and Søderlund, 1986)	Guinea pig hepatocytes	Significantly reduced	Multiple factors call into question relevance and use of	Low	UDS was measured by scintillation counting, which is not the recommended method, and can be susceptible to artifacts. The decrease in UDS

Study	Cell Type	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations or Concerns
			study for hazard assessment		is probably due to cytotoxicity. UDS is an "indicator test" and considered low weight.
(Hongslo et al., 1988)	Chinese Hamster Lung (V79)	Significantly reduced	Multiple factors call into question relevance and use of study for hazard assessment	Low	UDS was measured by scintillation counting, which is not the recommended method, and can be susceptible to artifacts. The decrease in UDS is probably due to toxicity associated with inhibition of replicative DNA synthesis. In any case, UDS is an "indicator test" and considered low weight.
Impairment of nucleotide excision repair					
(Hongslo et al., 1993)	UV-pretreated Mononuclear blood cells	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Low	Impairment of nucleotide excision repair which occurs at cytotoxic concentrations and exhibits a threshold; however, cytotoxicity not measured.
(Hongslo et al., 1993)	UV-pretreated T lymphocytes	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Low	Impairment of nucleotide excision repair which occurs at cytotoxic concentrations and exhibits a threshold; however, cytotoxicity not measured.
(Hongslo et al., 1993)	UV-pretreated B lymphocytes	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Low	Impairment of nucleotide excision repair which occurs at cytotoxic concentrations and exhibits a threshold; however, cytotoxicity not measured.
(Hongslo et al., 1993)	UV-pretreated Monocytes	Weakly positive	Multiple factors call into question relevance and use of study for hazard assessment	Low	Impairment of nucleotide excision repair which occurs at cytotoxic concentrations and exhibits a threshold; however, cytotoxicity not measured.
(Brunborg et al., 1995)	UV or 3 mM NQO-pretreated mononuclear blood cells	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Low	Impairment of nucleotide excision repair which occurs at cytotoxic concentrations and exhibits a threshold; however, cytotoxicity not measured.
(Brunborg et al., 1995)	UV-pretreated HL-60 cells	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Low	Impairment of nucleotide excision repair which occurs at cytotoxic concentrations and exhibits a threshold; however, cytotoxicity not measured.
(Brunborg et al., 1995)	UV-pretreated fibroblast cells	Weakly positive	Multiple factors call into question relevance and use of study for hazard assessment	Low	Impairment of nucleotide excision repair which occurs at cytotoxic concentrations and exhibits a threshold; however, cytotoxicity not measured.
(Brunborg et al., 1995)	UV-pretreated rat hepatocytes	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Low	Impairment of nucleotide excision repair which occurs at cytotoxic concentrations and exhibits a threshold; however, cytotoxicity not measured.
(Brunborg et al., 1995)	NQO-treated rat/ testicular cells	Positive	Multiple factors call into question relevance and use of	Low	Impairment of nucleotide excision repair which occurs at cytotoxic concentrations and exhibits a threshold; however, cytotoxicity not measured.

Study	Cell Type	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations or Concerns
			study for hazard assessment		
(Hongslo et al., 1988)	UV-pretreated Chinese Hamster lung V79 cells	Positive	Use of excessive concentrations and p53-deficient cells indicate not biologically relevant	Low	Repair inhibited at 3 & 10 mM, which did not significantly reduce colony formation. However, these concentrations exceed the recommended upper limit, and such results are irrelevant and should be discounted. Also, V79 cells are p53-deficient and susceptible to misleading positive responses not found in p53-competent, genomically stable (e.g. primary human) cells.
Impairment of DNA repair					
(Wan et al., 2004)	Rat/ C6 glioma cells	Positive	Use of excessive concentrations indicates not biologically relevant	Low	Inhibition of repair of oxidative damage by OGG1 at 5 mM, associated with ROS production and GSH depletion. This concentration exceeds the recommended upper limit for testing in mammalian cells. Such results are irrelevant and should be discounted.
Oxidation of DNA (8-oxodG)					
(Wan et al., 2004)	Rat/ C6 glioma cells	Positive	Use of excessive concentrations indicates not biologically relevant	Low	Induction of 8-oxoG at 2.5 mM and above, associated with ROS production and GSH depletion. These concentrations exceed the upper limit for testing in mammalian cells. Such results are irrelevant and should be discounted.

**Note the "Weight" column reflects the weight that should be given to the study type in the Weight of Evidence Hazard Assessment based on information regarding its relevance and reliability, predictivity, the endpoint's reversibility, susceptibility to false responses, and its mechanism's role in initiation of malignancy.*

Table 14: Overview of in vivo DNA damage studies

Study	Animal Model	Route of Exposure	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations or Concerns
Comet assay						
(Oshida et al., 2008)	Mouse (liver)	i.p.	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Moderate	Male mice given single i.p. doses up to 300 mg/kg. Liver sampled 4 & 24 hrs after dosing. Clear increase in comets at both sampling times, but only at the top dose, but plasma AST and ALT levels indicated hepatotoxicity. Therefore, the DNA damage may have been secondary to tissue toxicity.
(Oshida et al., 2008)	Mouse (kidney)	i.p.	Negative	Supports no hazard	Moderate	None
(Oshida et al., 2008)	Mouse (bone marrow)	i.p.	Negative	Supports no hazard	Moderate	None
DNA damage						
(Hongslo et al., 1994)	Rat (kidney)	i.p.	Negative	Supports no hazard	Moderate	None
(Hongslo et al., 1994)	Rat (liver)	i.p.	Negative	Supports no hazard	Moderate	None

Study	Animal Model	Route of Exposure	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations or Concerns
(Hongslo et al., 1994)	Mouse (kidney)	i.p.	Negative	Supports no hazard	Moderate	None
(Hongslo et al., 1994)	Mouse (liver)	i.p.	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Moderate	Used alkaline elution method for which there is no OECD guideline, so no recommendations for what constitutes an adequate test or how to interpret the results. Clear increase in DNA strand breaks after 600 mg/kg i.p. dose. Based on other studies this dose would be hepatotoxic, so DNA damage likely secondary to tissue toxicity.
(Van der Leede et al., submitted for publication on 10 Oct 2019, Manuscript number EMM-19-0142)	Rat (Sprague-Dawley); (peripheral blood and liver)	p.o.	Not Reviewed	Supports no hazard	High	Small increases in liver comets were seen in 2 out of 6 male rats dosed at 1000 mg/kg/day for 1 month, but single cell and focal necrosis were observed in the liver of these rats, so it is highly likely that these histopathological changes influenced the DNA damage response
Oxidation of DNA						
(Wang et al., 2015)	Mouse (serum)	p.o.	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Moderate	Mice given single dose of 400mg/kg (presumably oral, but not clear). 8-OH-dG level in liver increased slightly (50%), but AST & ALT levels increased markedly and GSH decreased, indicating hepatotoxicity.
Impairment of nucleotide excision repair						
(Hongslo et al., 1994)	NQO-treated Rat (liver, kidney, spleen)	i.p.	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Moderate	Impairment of nucleotide excision repair which occurs at cytotoxic concentrations and exhibits a threshold. A single i.p. dose of 600 mg/kg would be hepatotoxic.
(Hongslo et al., 1994)	NQO-treated Mouse (liver, kidney, spleen)	i.p.	Positive	Multiple factors call into question relevance and use of study for hazard assessment	Moderate	Impairment of nucleotide excision repair which occurs at cytotoxic concentrations and exhibits a threshold. A single i.p. dose of 600 mg/kg would be hepatotoxic.

**Note the "Weight" column reflects the weight that should be given to the study type in the Weight of Evidence Hazard Assessment based on information regarding its relevance and reliability, predictivity, the endpoint's reversibility, susceptibility to false responses, and its mechanism's role in initiation of malignancy.*

Table 15: Overview of human DNA damage studies

Study	HID Reported Result	WOE Hazard Assessment	Weight*	Considerations
UDS				
(Topinka et al., 1989)	Decreased response	Multiple factors call into question relevance and use of study for hazard assessment	Low	UDS was measured by scintillation counting, which is not the recommended method, and can be susceptible to artefacts. Decrease in UDS is probably due to toxicity. UDS is an "indicator test" and considered low weight.

**Note the "Weight" column reflects the weight that should be given to the study type in the Weight of Evidence Hazard Assessment based on information regarding its relevance and reliability, predictivity, the endpoint's reversibility, susceptibility to false responses, and its mechanism's role in initiation of malignancy.*

A more detailed evaluation of the studies is provided in the two sections that follow.

5.2.4 An Assessment of Studies Evaluated by Bergman, et al. (1996) Demonstrate No Meaningful Evidence of Potential for Acetaminophen to Cause Genetic Toxicity that Could Lead to Cancer

In this section, we review the studies evaluated by Bergman et al. (1996) and in the following section we review the studies since the Bergman et al. (1996) publication.

With regard to the Bergman et al. (1996) conclusions about gene mutations, one study (Clements, 1992) reported induction of *tk* mutations in mouse lymphoma cells in the absence of metabolic activation. However, a biologically relevant response was only observed at 13.2 mM, which far exceeds the current recommended limit (1 mM) and would be considered not biologically relevant by today's standards. Moreover, there was no sizing of mutant colonies, and since acetaminophen does induce chromosomal damage at concentrations also exceeding 1 mM, it is highly likely the mutant colonies were due to chromosome damage and not true gene mutations.

In addition, Bergman et al. (1996) did not review a mammalian cell gene mutation study by Shimane (1985), possibly because it was published in an obscure Japanese journal. In the study by Shimane (1985), V79 cells were treated with acetaminophen at 100, 200 and 400 µg/mL for 24 hrs, or 50, 100 and 200 µg/mL for 48 hrs in the absence of metabolic activation. Solvent control treatments were only included for the 24-hr treatments. After an appropriate expression time, cultures were assessed for mutations to 6-thioguanine (6TG) and ouabain resistance. At 200 µg/mL, cytotoxicity (reduction in colony forming ability) was around 25% for the 24-hr treatment and around 40% for the 48-hr treatment, but at 400 µg/mL cytotoxicity was >50% for both treatment times. 6TG mutant frequencies increased at 200 (>2-fold) and 400 µg/mL (>4-fold) following 24-hr treatment, but there was no statistical analysis, and no historical control data. Moreover, both of these concentrations exceed the current upper limit for testing (1 mM) according to ICH recommendations (ICH, 2011). 6TG mutant frequencies appeared also to increase at all 3 concentrations following 48-hr treatment, but since there was no solvent control

for the 48-hr treatments it is not possible to assess their biological relevance. Ouabain-resistant mutant frequencies increased at 100 and 400 µg/mL, but not at 200 µg/mL following 24-hr treatment, so there was no dose-response. It should be noted that V79 cells are p53-deficient, and highly susceptible to misleading positive results (Fowler et al., 2012), and as such these results would be considered only of moderate weight (see Section 5.1, above). Moreover, these results are in conflict with other studies where *Hprt* and ouabain mutations were not induced (Patierno et al., 1989; Sasaki et al., 1983; Sawada, 1985).

There is therefore no convincing evidence that acetaminophen induces gene mutations in robust, reliable, high weight test systems.

In the case of the conclusions of Bergman et al. (1996) with respect to chromosomal damage, clastogenic effects were seen at high concentrations that were toxic to the test system. Studies of the relationship between genotoxicity and toxic effects in the rat were reported. Bergman et al. (1996) considered the rat to be a suitable model for man, since rat and human hepatocytes display an equal susceptibility to the cytotoxicity of acetaminophen, although this has been challenged by McGill et al. (2012b) who considered that the mitochondrial dysfunction induced by acetaminophen in both mice and humans suggested mice would be a better model. However, it could be argued that since rats were only dosed at 3.3x the dose given to mice (whereas the LD_{50} is 6x higher) that mitochondrial dysfunction in rats would be seen at higher doses, and that the rat may well be an appropriate model for humans.

However, Bergman et al. (1996) described 2 previously unpublished MN studies in rats where slightly increased MN frequencies were seen only at oral doses (3 x 900 mg/kg at 4 hr intervals, or 3 x 500 mg/kg at 4 hr intervals, or 1 x 1500 mg/kg) causing marked liver and bone marrow toxicity. In a more recent study (van der Leede et al., in press) acetaminophen at oral doses up to 2000 mg/kg/day for 3 and 29 days and 1000 mg/kg for 15, including a 1 month recovery phase following the 29 day treatment, did not induce biologically relevant increases in comets in peripheral blood cells, or Pig-a mutations in reticulocytes or erythrocytes, but only slight to minimal hepatotoxicity was seen particularly after extended dosing or recovery. In this study, statistically significant increases in MN in reticulocytes after 1 month of dosing at 500 and 1000 mg/kg/day were attributed to rebound erythropoiesis in response to marked hematotoxicity (severe bone marrow toxicity was seen 4 days after the start of dosing), and therefore the increased MN were concluded to be due to a non-genotoxic mode of action. Also, small increases in liver comets were seen in 2 out of 6 male rats dosed at 1000 mg/kg/day for 1 month, but single cell and focal necrosis were observed in the liver of these rats, so it is highly likely that these histopathological changes influenced the DNA damage response. Therefore, the small increases in DNA damage levels were not considered biologically relevant.

Bergman et al. (1996) also reviewed chromosomal aberration (CA) data from three publications of human studies in which 1 g of acetaminophen was given orally 3 times during an 8-hr period.

They concluded that mixed results were obtained. These 3 studies are discussed briefly as follows:

- Kocisová et al. (1988) reported 2 studies. In the first study, acetaminophen was administered (3 x 1g during 8 hrs) to 11 volunteers (3 males/8 females), and a small but statistically significant ($p<0.05$) increase (from 1.68% pre-dose to 2.77% at 24 hrs after the first dose) in the proportion of cells with CA (excluding gaps) was observed. However, CA frequencies were not significantly different from pre-dose levels at later sampling times (72 or 168 hrs), and had returned to below pre-dose levels by 168 hrs. Thus, the increase in the proportion of cells with CA was transient, which is unusual since in other longitudinal studies CA levels tend to remain increased for periods of weeks or months (Kucerova et al., 1980; Schmid et al., 1985). The transient nature of the response could indicate that the damage was lethal, and that the damaged/dead cells had disappeared by the later sampling times. In the same publication a second study with the same volunteers was performed 1 week later with the same dosing schedule, except that each dose of acetaminophen was given together with 1 g of the anti-oxidant, ascorbic acid. A small but statistically significant ($p<0.05$) increase (from 1.09% pre-dose to 2.22% 72 hrs after the first dose) in the proportion of cells with CA was observed. CA levels were not significantly different from pre-dose at 24 or 168 hrs, so again the increase in the proportion of cells with CA was transient. It is unclear whether the co-administration of ascorbic acid delayed the appearance of CA, or whether the time difference was due to chance. It should be noted that in both studies the increased CA levels were due entirely to chromatid breaks; there were no increases in chromosome breaks or exchanges. It was most interesting that the individual responses of the volunteers in the first and second studies showed that 7 and 6, respectively, of the 11 volunteers showed an *increase* in the number of aberrant cells, whereas 4 and 5 volunteers, respectively, showed *no increase or a decrease* in the numbers of aberrant cells. Since the same volunteers were used in both studies, it was possible to see that no specific sub-group of the volunteers showed a consistent response (i.e. those that showed increased CA levels with acetaminophen alone were not the same as those showing increased CA levels with acetaminophen plus ascorbic acid). On the contrary, it was apparent that those individuals who had shown a comparatively large increase in chromatid break frequency in the first study showed a small increase or even a decrease in the second study, and vice versa. It is therefore highly implausible that the increased CA levels in these 2 studies resulted from the genotoxic effects of acetaminophen, and it is more likely they were due to chance.
- Hongslo et al. (1991) administered acetaminophen (3 x 1g during 8 hrs) to 9 volunteers and reported a small (from 2.38% pre-dose to 5.03% 24 hrs after the first dose) but insignificant ($p<0.1$) increase in the proportion of cells with CA, including gaps. When gaps were excluded (as is normal practice) the increase was much smaller from 2.16% to 3.43% (this was not analyzed for statistical significance). Excluding gaps, the increase was primarily due to a 6-fold increase in chromatid breaks (i.e. similar to the observations of (Kocišová et al., 1988), although no blood samples were taken at later sampling times). As in the Kocisová et al. (1988) study, not all

volunteers showed an increase in the levels of aberrant cells, excluding gaps (7/9 volunteers showed an increase but 2/9 showed a decrease).

- The study of Kirkland et al. (1992) was considered by Bergman et al. (1996) to be the most carefully controlled of these human studies since it was, unlike the other studies, **a double-blind study** (i.e. acetaminophen, 3 x 1 g during 8 hrs, and placebo groups) in 24 volunteers (12 males/12 females). **The study therefore not only compared pre- and post-dose samples from each individual, but also compared acetaminophen-treated with placebo-treated groups (this study was incorrectly characterized in the HID).** Blood samples for the determination of CA frequencies in peripheral lymphocytes were taken 24 hrs prior to dosing and at 24 hrs, 3 days and 7 days after administration of the first dose. Although a larger number of cells than in the two other studies was analyzed no significant increases in % cells with CA (excluding gaps) were found either (a) when CA levels in the acetaminophen-treated individuals (men or/and women) were compared post-dose with pre-dose, or (b) when treated groups at any sampling time were compared with the placebo groups. There was no evidence that any individual responded to acetaminophen or that a group response was masked by non-responders. The study also included determinations of plasma concentrations of paracetamol; C_{max} after the third dose was 0.08 mM in men and 0.11 mM in women.

It is important to note that the HID has multiple statements that are not scientifically accurate or complete in their presentation of the Kirkland et al. (Kirkland et al., 1992), Kocisova, et al. (Kocišová et al., 1988), and Honglso et al (Hongslo et al., 1991), genetic toxicology results. A detailed explanation of the issues associated with the HID assessment can be found in the Section 8.5 of the Appendix. One example of this is that the HID states that:

> "*Acetaminophen induced SCEs in PBL in one study (Hongslo et al. 1991) and had no effect in another study (Kirkland et al. 1992)*" and that "*[i]t is possible that Kirkland et al. (1992) had a reduced ability to detect acetaminophen-related effects on PBL CAs and SCEs due to inter-individual variability between the placebo and acetaminophen- treated groups in "baseline" levels of these markers of clastogenicity." (HID: p. 155).*

This is not correct, since the authors (Kirkland et al., 1992) examined both pre- and post-dose samples as well as acetaminophen and placebo groups and found no increase in CA induction for either comparison group. It is also important to note that the HID incorrectly reported that Kirkland et al. (1992) examined SCEs, which they did not.

Bergman et al. (1996) also noted that a genetic polymorphism with respect to glutathione transferase has been described for Caucasians, a minor proportion of which lack glutathione transferase genes, and this may render them more susceptible to genotoxic compounds. However, as discussed above, the individual data on the volunteers of the Kocisová et al. (1988) studies do not indicate that individual differences affected the increased chromosomal damage that they observed. In the large, double-blind, carefully controlled study of Kirkland et al. (1992), individuals possibly at higher risk were also probably included, yet this study found no indications

of a clastogenic effect at maximum therapeutic dosage. The findings of Kirkland et a. (1992) were confirmed in a study by Hantson et al. (1996), that was published after the Bergman et al. review. This showed that in volunteers who had been administered a single oral dose of 3 g acetaminophen, patients who had received 2 g of acetaminophen by intravenous infusion every 6 hrs for at least 7 days, and in self-poisoned patients who, for suicidal reasons, had ingested more than 15 g acetaminophen, there were no increases in the frequency of structural chromosomal aberrations in the circulating lymphocytes.

Bergman et al. (1996) also reviewed micronucleus (MN) data from two publications of human studies in which 1g of acetaminophen was given orally 3 times during an 8-hr period. Due to lack of methodological detail in these papers they were unable to reach any firm conclusions. These 2 studies are discussed briefly as follows:

- Topinka et al. (1989) administered acetaminophen (3 x 1g during 8 hrs) to 11 volunteers (3 males, 8 females). Another group (or maybe the same group of volunteers, since the design is identical that of (Kocišová et al., 1988), and the average age was the same) were co-administered acetaminophen and ascorbic acid. Buccal cells were sampled at 0 (presumably equivalent to pre-dose), 24, 72 and 168 hrs after the first of the 3 doses. Slides were made, stained with light green, coded (for blinded scoring) and 2000 cells/sample scored for presence of MN. No data on the individual volunteers was presented. A statistically significant 2-fold increase in the group mean frequency of micronucleated buccal cells was seen at 72 hrs but not at 24 or 168 hrs. A slightly smaller, but still statistically significant increase, was seen at 72 hrs in the acetaminophen + ascorbic acid group, but again there were no increases at 24 or 168 hrs. Since no historical data are given, it is unclear whether the raised MN frequencies (0.38% in each case) were within normal control ranges. However, the authors note that "*the statistically significant increase of micronuclei is low in comparison with other groups presented by Stich et al. (1983)*". The pre-dose MN frequencies in this study were 0.19 and 0.23%. However, a survey of multiple publications by Holland et al. (2008) revealed baseline MN frequencies ranging from 0.05-1.15%. Thus, a frequency of 0.38% would be well within the observed normal range and may simply represent "background noise". In the same study the authors observed decreased UDS at all sampling times, but notably at 168 hrs, and so the MN may result from inhibition of ribonucleotide reductase (see later). However, it is therefore curious when increased MN frequencies were only seen at 72 hrs and not also at 168 hrs. The biological relevance of these results is therefore highly questionable.
- Kocisová and Sram (1990) used the same treatment and sampling regimens as described by Topinka et al. (1989) and Kocisová et al. (1988) but with 12 volunteers (3 males/9 females). Blood samples were taken at 0 (presumably equivalent to pre-dose), 24, 72 and 168 hrs after the first of the 3 doses. Lymphocytes were stimulated to divide by phytohemagglutinin, and Cytochalasin B was added to the cultures 44 hrs later. Cultures were harvested at 72 hrs, cells were gently swollen, fixed and stained with Giemsa. It is not stated whether slides were blinded before scoring, but 1000 binucleate cells/sample were scored for MN. The frequencies of MN at all

sampling times were similar to the pre-dose frequency, and not significantly different, whereas the MN frequency in an elderly group of volunteers (included as a "positive control" group since MN frequencies increase with age) was significantly different. **Thus, under the same conditions as this research group found increased CA in blood lymphocytes and reported increased MN in buccal cells (although the biological relevance is debatable), there were no increases in MN frequency in blood lymphocytes.**

5.2.5 An Assessment of Publications Since (Bergman et al., 1996) Demonstrate No Meaningful Evidence of Potential for Acetaminophen to Cause Genetic Toxicity that Could Lead to Cancer

Since the detailed review of Bergman et al. (1996) several papers related to the genotoxicity of acetaminophen have been published. Since the focus of the Bergman paper was the relevance of genotoxicity at therapeutic doses, this assessment builds on the review of Bergman et al. (1996) and considers how the data previously reviewed or any new data impacts the genotoxic hazard assessment for acetaminophen.

In addition to some recent publications discussed in the text above, the following studies have been identified and are considered relevant to a discussion of potential for acetaminophen genotoxicity.

(vi) Gene mutations

Martinez et al. (2000) showed that acetaminophen, when tested up to 1500 µg/plate, was not an oxidative mutagen in the *E. coli* WP2 Mutoxitest. This confirms the lack of gene mutation activity *in vitro* reported by Bergman et al. (1996). Kanki et al. (2005) tested acetaminophen (10000 ppm in diet for 13 weeks, equivalent to 140 mg/rat/day) for induction of gene mutations (6-thioguanine resistance) in female transgenic *gpt* delta rats. The treatment resulted in a statistically significant increase in liver/bodyweight ratio, but there was no increase in GST-P positive liver cell foci, and no increase in *gpt* mutant frequency, even though other substances tested at the same time (IQ and *N*-nitrosopyrrolidine) were positive for both markers. These results confirm *in vivo* the lack of gene mutation activity seen *in vitro*. These negative results were confirmed by Matsushita et al. (2013) in male *gpt* delta rats fed acetaminophen at 6000 ppm in diet for 4 weeks, where there was no increase in *gpt* mutant frequency. The mutation spectra in acetaminophen-treated rats were also similar to those in controls. Acetaminophen inhibited the formation of GST-P positive liver cell foci, as was also shown by Ito et al. (1988).

Suzuki et al. (2016) showed that single oral doses of acetaminophen at 500, 1000 or 2000 mg/kg did not induce *Pig-a* mutations in either erythrocytes or reticulocytes of rats, sampled 1, 2 or 4 weeks after dosing. By contrast, the positive control chemical, *N*-nitroso-*N*-ethylurea, induced a clear time-related response.

In a more recent study in rats (van der Leede et al., in press), acetaminophen at oral doses up to 2000 mg/kg/day for 3 or 29 days, and up to 1000 mg/kg/day for 15 days, including a 1 month recovery phase, did not induce biologically relevant increases in *Pig-a* mutations in reticulocytes or erythrocytes, but only slight to minimal hepatotoxicity was seen particularly after extended dosing or recovery.

(vii) Chromosomal damage

Ibrulj et al. (2007) confirmed the ability of acetaminophen to induce chromosomal aberrations in cultured human lymphocytes, exposed continuously for 72 hrs, at a concentration of 200 µg/mL (1.3 mM), whereas negative results were obtained at 50 and 100 µg/mL. However, no micronuclei were induced at any of the concentrations tested. Although cytotoxicity would be expected at concentrations >1 mM, the effect on nuclear division index was small (in the region of 20% at 200 µg/mL). The chromosomal aberration results are similar to those reported by Honglso et al. (1991) in human lymphocytes exposed to acetaminophen for the last 24 hrs of a 72-hr incubation. It is important to note that almost all induced aberrations in both studies were chromatid breaks which, as discussed earlier, are associated with cell lethality. Acetaminophen had been shown to induce micronuclei in rat kidney fibroblast NRK-49F cells (Dunn et al., 1987), but only at very high concentrations (10 and 20 mM), whilst there were no previous micronucleus data in human lymphocytes.

Matsushima et al. (1999) showed that acetaminophen induced micronuclei in Chinese hamster lung (CHL) cells after extended (24- and 48-hr) treatments in the absence of metabolic activation, but significant effects were seen at lower concentrations (from about 20 µg/mL, 0.13 mM, and above). However, there was no concurrent measure of cytotoxicity reported. As described above, Shimane (1985) also reported induction of chromosomal aberrations in V79 cells in the absence and presence of metabolic activation, at concentrations ranging from 25-200 µg/mL. However, both of these studies used p53-deficient Chinese hamster cell lines, and, as discussed earlier, p53-deficient rodent cells are known to be more sensitive to cytotoxic and genotoxic chemicals than p53-competent human cells (Fowler et al., 2012), particularly in the absence of detoxification processes. Although p53-deficient rodent cells give positive results with chemicals that are not genotoxic or carcinogenic in rodents *in vivo*, and this would be considered to be indicative of absence of genotoxic or carcinogenic effects in humans, a direct comparison with humans has not been possible. The induction of micronuclei at low concentrations in CHL cells is consistent with induction of chromosomal aberrations in the same cells reported by Ishidate et al. (1988), who also reported clastogenic effects in p53-deficient CHO-K1 cells at similar low concentrations, and consistent with the induction of chromosomal aberrations in V79 cells by Shimane (1985), whereas (as expected from comments made earlier) much higher

concentrations (1.3 mM) were required for clastogenic effects in p53-competent human lymphocytes.

Markovic et al. (2013) administered acetaminophen intraperitoneally at 60 mg/kg to pregnant BALB/c mice consecutively on days 12 and 14 of pregnancy. The dose is equivalent to a normal 50 kg human taking 3g of acetaminophen during 1 day. Blood samples were taken from the dams on day 12 of pregnancy and 48 hours after drug administration for *in vivo* micronucleus assays. In each litter, blood samples from 6 animals were analyzed for micronuclei. Anti-oxidant activity (glutathione peroxidase in blood) and an indicator of lipid peroxidation (malondialdehyde in liver) were also measured in the pups. For each of the micronucleus assays, 1000 acridine orange-stained reticulocytes per animal were assessed. This is a much smaller population of cells than is currently recommended in OECD guidelines. Importantly, it is not stated that the slides were "blinded" before scoring, and therefore scorer bias cannot be excluded. Micronucleus frequencies in vehicle control animals were normal (0.86/1000 reticulocytes) and were significantly increased by the positive control chemical (cyclophosphamide). Micronucleus frequencies in the dams treated with acetaminophen were increased slightly (3.25-fold) above vehicle control frequencies at 48 hrs after dosing but were not significantly different. On the other hand, micronucleus frequencies in the blood of the pups showed a smaller increase (2.28-fold) above vehicle controls, but this was statistically significant ($p<0.05$). Glutathione peroxidase activity in the hemolysate of the new-born pups, and malondialdehyde levels in the livers of the pups, were significantly lower than in vehicle control pups. The authors speculate that the reduction of glutathione peroxidase reflected systemic oxidative stress. They state that this reduction is known to occur with acetaminophen treatment, while the reduction of malondialdehyde in the liver can be interpreted as an unspecific reaction to drug treatment that might have cytotoxic, and in particular hepatotoxic, effects associated with oxidative stress and lipid peroxidation. Given that mice are more sensitive than rats or humans to the hepatotoxic effects of acetaminophen, that the increases in micronucleus frequency in the dams were higher than in pups, yet were not statistically significant, and that the slides were probably not "blinded" before scoring, these results should be viewed with caution. The results are probably consistent with the variable *in vivo* micronucleus results in mice summarized in Bergman et al. (1996).

However, as discussed above, Bergman et al. (1996) described 2 previously unpublished MN studies in rats where slightly increased MN frequencies were seen only at oral doses (3 x 900 mg/kg at 4 hr intervals, or 3 x 500 mg/kg at 4 hr intervals, or 1 x 1500 mg/kg) causing marked liver and bone marrow toxicity. In a more recent study in rats (van der Leede et al., in press), acetaminophen at oral doses up to 2000 mg/kg/day for 3 or 29 days, and up to 1000 mg/kg/day for 15 days, including a 1 month recovery phase, statistically significant increases in MN were seen in reticulocytes after 1 month of dosing at 500 and 1000 mg/kg/day but were attributed to erythropoiesis in response to marked hematotoxicity (severe bone marrow toxicity was seen 4

days after the start of dosing), and therefore concluded to be due to a non-genotoxic mode of action.

(viii) DNA damage

Oshida et al. (2008) investigated the induction of DNA strand breaks (comets) in the livers, kidneys and bone marrow of mice given a single intraperitoneal dose of acetaminophen. No DNA damage was induced in kidneys or bone marrow, and comets were only induced in liver at the highest dose (300 mg/kg) where hepatotoxicity was also observed.

In the recent study of van der Leede et al (in press) small increases in liver comets were seen in 2 out of 6 male rats dosed at 1000 mg/kg/day for 1 month, but single cell and focal necrosis were observed in the liver of these rats, so it is highly likely that these histopathological findings influenced the DNA damage response. Therefore, the small increases in DNA damage levels were not considered biologically relevant.

(ix) Oxidative stress

Although, as discussed earlier, NAPQI is likely to induce oxidative stress, and Bisaglia et al. (2002) indicates that acetaminophen also has anti-oxidant properties. Whether acetaminophen exhibits oxidant or anti-oxidant activity may be a question of dose, as has been seen for other substances such as flavonoids and polyphenols [e.g. see (Slezak et al., 2017) where antioxidant activity tends to be manifest at low concentrations, but reactive oxygen species are induced at high concentrations. Thus, for acetaminophen, antioxidant activity could be seen at lower doses/concentrations where NAPQI is effectively bound to glutathione but can induce oxidative stress at higher doses/concentrations where glutathione is depleted.

Powell et al. (2006) dosed male rats with acetaminophen at sub-toxic (150 mg/kg) or overtly toxic (1500 and 2000 mg/kg) doses. Animals were sacrificed 6, 24, or 48 hours later, and liver tissue was used to generate microarray data. Oxidative stress in liver was evaluated by a diverse panel of markers that included assessing expression of base excision repair (BER) genes, quantifying oxidative lesions in genomic DNA, and evaluating protein and lipid oxidation. A sub-toxic dose of acetaminophen produced significant accumulation of nitrotyrosine protein adducts, while both sub-toxic and toxic doses caused a significant increase in 8-hydroxy-deoxyguanosine, markers that are anchored on the mechanism of acetaminophen-induced liver toxicity. Only toxic doses of acetaminophen significantly induced expression levels of BER genes. None of the doses examined resulted in a significant increase in the number of abasic sites, or in the amount of lipid peroxidation.

5.3 Relevance of Metabolites of Acetaminophen for Genetic Toxicology Hazard Assessment

Other possible metabolites of acetaminophen besides NAPQI that the HID reviews, which have only been measured in rodents, are NAPSQI, *p*-benzoquinone, *p*-benzoquinone imine, *p*-aminophenol (PAP), and the *N*-acetyl-*p*-aminophenoxy and *p*-aminophenoxy radicals. Importantly, we could not find any definitive evidence that these have been detected in humans dosed with acetaminophen. These metabolites may be predicted to be DNA-reactive, and some of these have been shown to produce genotoxic effects *in vitro,* and in some cases also *in vivo,* when tested alone. However, as discussed elsewhere in this document, reliable high weight studies such as the Ames test gave consistently negative results with acetaminophen, so if these metabolites were formed when acetaminophen was tested in the presence of metabolic activation, they did not induce gene mutations. Moreover, if these metabolites were formed *in vivo* when animals were dosed with acetaminophen, they did not lead to genotoxic effects at sub-toxic doses, since genotoxic effects in reliable and relevant studies in animals were only seen at hepatotoxic doses. In addition, these metabolites were covered in the animal carcinogenicity studies with acetaminophen, which again did not show any carcinogenic potential (see Section 4).

5.4 Cell Transformation Studies

As discussed in the HID, Patierno et al. 1989 studied in vitro cell transformation of C3H/10T1/2 clone 8 mouse embryo fibroblast (10T1/2) cells exposed to acetaminophen. These cells are considered to be similar to BALB/3T3 and Swiss/3T3 cells, as they are stable in culture and highly sensitive to post-confluence inhibition of cell division (Reznikoff et al., 1973). C3H/10T1/2 cells, together with other immortalized aneuploid mouse cells, represent one of the two major types of systems used for in vitro cell transformation assays, the other type being primary diploid cells, such as Syrian Hamster Embryo cells (Creton et al. 2012).

In this study, Patierno et al. (1989) treated 10T1/2 cells with acetaminophen at concentrations ranging from 0.5 – 2.0 mg/mL (3.3 to 13 mM) for either 24 hours without S-9 or 3 hours with Aroclor 1254-induced hamster liver S-9. In the absence of S-9 acetaminophen induced a small, but dose-dependent increase in the number of type II morphologically transformed foci. A greater number of type II transformed foci were induced by acetaminophen in the presence of S-9. Similar cell transformation results were observed with the carcinogen phenacetin (of which acetaminophen is a major metabolite). Several metabolites of acetaminophen (and phenacetin) were also tested in C3H/10T1/2 cells (NAPQI, PAP, p-benzoquinone), and each were found to be inactive in the cell transformation assay. Patierno et al. (1989) characterized the type II foci induced by acetaminophen and phenacetin as atypical (weak) non-neoplastic morphologically

transformed cells that "did not exhibit any other classical parameters of neoplastic transformation, such as increased saturation density or anchorage independence." (p. 188)

Patierno et al. (1989) indicated that the "results suggest that metabolic intermediates of high concentrations of phenacetin and acetaminophen induce a low frequency of nonneoplastic morphological transformation of 10T½ mouse embryo cells" (Patierno et al., 1989): p. 1038). Further, the authors noted that "[e]ven though the mixed clones reformed weak type II foci when maintained at confluence, they did not exhibit any other classical parameters of neoplastic transformation, such as increased saturation density or anchorage independence" (Patierno et al., 1989): p. 1043). Therefore, the results by Patierno et al. (1989) suggest that acetaminophen does not cause neoplastic transformation in this *in vitro* assay.

5.5 Genetic Toxicology - Discussion and Conclusions

The clastogenic effects of acetaminophen in relevant systems only occur at cytotoxic exposures, such that the cells containing these chromosomal aberrations will not be able to survive to produce stable or persistent genetic damage that could pre-dispose to genetic disease or cancer. Acetaminophen does not induce gene mutations in bacteria or cultured cells *in vitro* (see (Bergman et al., 1996) for details; also, (Martinez et al., 2000), or *in vivo* (Kanki et al., 2005; Matsushita et al., 2013; Suzuki et al., 2016). It can induce genotoxic effects (chromosomal and DNA damage) in cultured cells and animals, but in genomically stable p53-competent cells and in animal species more resistant to the hepatotoxic effects of acetaminophen, this only occurs at extreme and/or toxic exposures.

It is useful to compare the pattern of genotoxicity results observed for acetaminophen with the pattern of results that would be expected for a clear genotoxic carcinogen. This comparison is summarized in Table 16 below, where it can be seen quite clearly that acetaminophen does not present a profile that is typical of a clear genotoxic and carcinogenic hazard:

Table 16: Comparison of test response profiles from acetaminophen to the profile characteristics of confirmed genotoxic carcinogens (adapted from (Brusick et al., 2016); based on (Bolt et al., 2004) and (Petkov et al., 2015).

Characteristic	Carcinogens with a proven genotoxic mode of action	Acetaminophen
Profile of Test Responses in Genetic Assays	Positive effects across multiple key predictive endpoints (i.e. high weight studies such as gene mutation in bacteria or *in vivo*, chromosomal aberrations or micronuclei *in vivo)*	No valid evidence for gene mutation in bacteria, mammalian cells or *in vivo*; no convincing evidence of chromosomal aberrations in humans; chromosomal damage in rodents only at hepatotoxic doses.
Structure Activity Relationships	Positive for structural alerts associated with genetic activity	Not assessed.

Characteristic	Carcinogens with a proven genotoxic mode of action	Acetaminophen
DNA binding	Agent or breakdown product are typically electrophilic and exhibit direct DNA binding	No unequivocal evidence that metabolically activated acetaminophen or NAPQI forms DNA adducts in cells *in vitro* at concentrations that do not also cause cytotoxicity; no reliable evidence of DNA adduct formation in animals or humans *in vivo* at any dose level
Consistency	Positive test results are highly reproducible both *in vitro* and *in vivo*	Conflicting and/or non-reproducible responses in the same test or test category both *in vitro* and *in vivo.*
Response Kinetics	Responses are dose dependent over a wide range of exposure levels	Any positive responses in robust, reliable test systems are generally non-linear, exhibiting a threshold.
Susceptibility to Confounding Factors (e.g. Cytotoxicity)	Responses are typically found at non-toxic exposure levels	Positive responses in robust, reliable test systems typically associated with evidence of overt toxicity.

There is increasing evidence that many substances producing genotoxic responses, particularly *in vitro* in tests detecting chromosomal or DNA damage, exhibit thresholds that are tied to cytotoxicity. Several publications have described modes of action and circumstances that would define such a threshold-mediated genotoxic response (Muller and Kasper, 2000; Scott et al., 1991; Thybaud et al., 2007).

In conclusion, acetaminophen overwhelmingly produces negative results in reliable, robust high weight studies (Brusick et al., 2016), as discussed earlier. Some genotoxic effects (clastogenicity) are seen in moderate weight studies, but in relevant, robust test systems these are only seen at unacceptably high concentrations or under cytotoxic conditions and associated with cell lethality. Therefore, from all of the available data, it is not plausible that acetaminophen induces the stable, genetic damage that would be indicative of a clear genotoxic or carcinogenic hazard in humans.

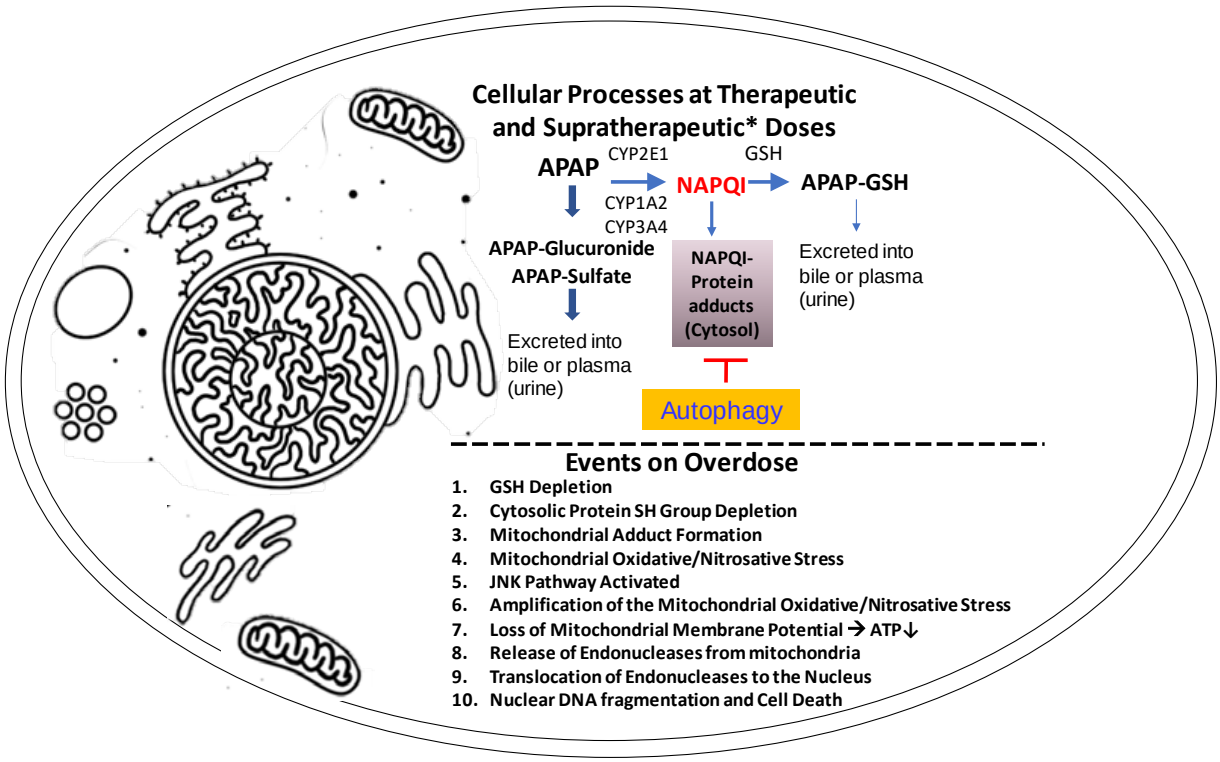
6 Mode of Action Studies – Pathways and Pharmacology Considerations

An understanding of the Mode(s) of Action can provide critical insights and data to support an assessment of the plausibility for a chemical entity to be a carcinogenic hazard (EPA, 2005). There are significant mechanistic data across *in vitro* and *in vivo* test systems and in humans supporting that the mode of action of acetaminophen at therapeutic, supratherapeutic and overdose exposures precludes its ability to cause cancer. There is no meaningful evidence of DNA effects of acetaminophen at therapeutic exposures in animals and humans. There is significant evidence

that has been generated over the past 15-20 years in animals and humans demonstrating that the dominant mechanism for DNA effects in the cell following supratherapeutic exposures to acetaminophen is through inhibition of mitochondrial respiration leading to mitochondrial dysfunction (McGill et al., 2013). Mitochondrial dysfunction occurs when cellular glutathione is depleted, and mitochondrial protein adducts are formed causing mitochondrial burst and substantial cytotoxic oxidative stress and DNA damage as the cell dies (McGill et al., 2013).

Figure 16 shows a high-level schematic diagram of the intracellular modes of action for acetaminophen at therapeutic and supratherapeutic doses and on overdose.

*Figure 16: Mode of action for acetaminophen following therapeutic (4 g/day or less), supratherapeutic (> 4 – 10 g/day) and overdose exposures (> 10-15 g acute). The reactive metabolite of acetaminophen, NAPQI, binds to glutathione and proteins within the cytosol at therapeutic doses and there are no adverse effects. When glutathione is depleted at supratherapeutic doses and on large acute overdose, NAPQI binds to mitochondrial proteins resulting in mitochondrial dysfunction, mitochondrial dependent DNA fragmentation and cell death, which prevents any effects on nuclear DNA that could drive carcinogenesis. Note that CYP1A2 and 3A4 pathways for NAPQI have only been confirmed in animals. *At supratherapeutic doses there can be isolated cells in the centrilobular region of the liver that experience steps 1-8 shown for overdose which may result in some hepatic cell death without any adverse clinical effects.*



In the sections that follow we present data and mechanistic results in more detail supporting that acetaminophen has a non-carcinogenic mode of action.

6.1 Acetaminophen Causes Cellular Toxicity Before it Can Cause Adverse DNA Effects

The underlying mechanisms of acetaminophen-mediated tissue toxicity have been well studied and occur in a dose-dependent manner. Specifically, acetaminophen toxicity depends upon the formation of the reactive metabolite NAPQI. At supratherapeutic doses, excess NAPQI can deplete GSH stores, and protein adducts are formed primarily in hepatocytes because of the higher concentration of CYP2E1 and higher exposures in hepatocytes compared to other cell types. The resulting NAPQI-associated protein adducts can be detected in the cytosol and in mitochondria. However, mitochondrial protein adducts cause mitochondrial dysfunction by increasing the generation of reactive oxygen species (ROS) such as superoxide and peroxynitrite (Ramachandran and Jaeschke, 2018). These processes ultimately result in acetaminophen-metabolite mediated cytotoxicity resulting in organ dysfunction.

The known mechanism by which acetaminophen induces liver damage is particularly pertinent when evaluating the potential carcinogenic hazard of the drug. In a review of the genotoxic mode of action (MOA) of acetaminophen, the same mechanism of action (NAPQI-mediated oxidative stress) was identified for genotoxicity at supratherapeutic doses (Bergman et al., 1996). Notably, all the genotoxic effects of acetaminophen in reliable, robust test systems are related to clastogenic effects under conditions that were toxic to the test system, and notably did not lead to gene mutations and involved dose thresholds for effects (Bergman et al., 1996). Given that the mechanisms by which acetaminophen causes toxicity and genotoxicity *in vitro* and in rodents are threshold-based and are consistent with the molecular mechanisms of acetaminophen-induced cytotoxicity in humans, the absence of carcinogenicity in rodent bioassays would serve to support that it is not a carcinogenic hazard despite causing clastogenic effects at toxic doses.

While acetaminophen has properties that have been associated with the Key Characteristics of Carcinogens (KCC) (e.g. forms reactive metabolite), there is no substantial evidence that these characteristics result in causation of cancer in the case of acetaminophen. For acetaminophen, the proposed DNA effects that could potentially drive a tumorigenic response only occur at doses where there is cell death and no chance for the DNA damage to be propagated to daughter cells. In addition, acetaminophen-induced DNA damage involves a fundamentally different mechanism that is caused by endonuclease-mediated DNA fragmentation that non-reversibly degrades the nucleus of a dying cell (Bajt et al., 2006; Cover et al., 2005b). From a Mode of Action perspective, there is a sequence of intracellular events that occurs following exposure to acetaminophen that may result in cellular toxicity but that prevent acetaminophen from being a carcinogenic hazard at any dose level. The sequence of events and the exposure range at which they occur in rodents and humans are summarized in Table 17.

Table 17: Mechanistic sequence of events that occurs within a cell at therapeutic, supratherapeutic and overdose exposures to acetaminophen across species demonstrating why acetaminophen is not a carcinogenic hazard under any dosing scenario.

Exposure Humans	Exposure Mice (mg/kg)	Exposure Rats (mg/kg)	Events	Ref.
Therapeutic (up to 4 grams/day)	75	15	• 90% of dose is directly conjugated by glucuronidation and sulfation; • <10% of dose is metabolized by Cyps to form NAPQI • Limited amount of NAPQI formed is almost entirely conjugated with GSH; • GSH levels in hepatocytes are very high (5-10 mM); • Temporary reduction in hepatic GSH levels by <5% of baseline; • Very few protein adducts are formed, which are removed by autophagy; • No evidence that any NAPQI reaches the nucleus in the presence of high GSH levels and presence of cytosolic proteins with free SH groups *in vivo*. • No evidence of mitochondrial protein adducts formed, JNK activation, mitochondrial oxidant stress or dysfunction and there is no evidence of DNA damage or cell death.	1-5
Supra-therapeutic (> 4-8 grams/day)	100-150	80	• 90% of dose is directly conjugated by glucuronidation and sulfation; • Still <10% of dose is metabolized by CYPs to form NAPQI • Somewhat higher levels of NAPQI formed (compared to therapeutic doses) • NAPQI almost entirely conjugated with GSH leading to temporary depletion and rapid recovery of hepatic GSH content; GSH levels in hepatocytes are still high (5-10 mM); • Limited protein adducts are formed, which are removed by autophagy; • Higher supratherapeutic doses (in mice) can lead to limited protein adducts in mitochondria leading to temporary JNK activation and even temporary, reversible mitochondria membrane permeability transition pore (MPTP) opening; • The MPTP is reversible and the breakdown of the membrane potential is reversible. • No mitochondrial intermembrane protein release, no nuclear DNA fragmentation and generally no cell death. • No evidence that any NAPQI reaches the nucleus in the presence of high GSH levels and presence of cytosolic proteins with free SH groups *in vivo*	1-7
Overdose (>10-15 g acute exposure)	250	600-1000	**Mice and Humans** • Most of the overdose is still directly conjugated by glucuronidation (predominantly by phase II reactions) and sulfation (saturated); • Still a minority of the overdose is metabolized by CYP2E1 to form NAPQI; • However, much higher amounts of NAPQI are formed after an overdose; NAPQI is conjugated in part with GSH – leading to extensive depletion of GSH; • GSH levels in centrilobular hepatocytes are depleted by >90%; • There is substantial protein adduct formation involving cytosolic and mitochondrial proteins; • The mitochondrial adducts trigger a mild mitochondrial oxidant stress, which is not counteracted by GSH due to its depletion; • The oxidant stress triggers activation of a mitogen activated protein kinase (MAPK) cascade leading to activation of JNK (phosphorylation); • P-JNK translocates to the mitochondria and binds to the anchor protein Sab, which triggers further restriction of the electron flow on the electron transport chain amplifying the oxidant stress and peroxynitrite formation; • The oxidative/nitrosative stress causes mitochondrial DNA (mtDNA) damage but does not affect molecules outside the mitochondria.	1, 7-13

Exposure Humans	Exposure Mice (mg/kg)	Exposure Rats (mg/kg)	Events	Ref.
			• The oxidative/nitrosative stress triggers the mitochondrial membrane permeability transition pore (MPTP) opening, which causes the collapse of the membrane potential and cessation of ATP synthesis; • The MPTP opening causes mitochondrial matrix swelling leading to rupture of the outer membrane and release of the intermembrane proteins endonuclease G and AIF, which translocate to the nucleus and induce DNA fragmentation. • Mitochondrial dysfunction and nuclear DNA fragmentation causes necrotic cell death. • Thus, nuclear DNA fragmentation is completely dependent on mitochondrial dysfunction and represents the point of no-return for cell death. **Rats (**Note: doses that cause toxicity vary across strains**)** • Most of the overdose is still directly conjugated by glucuronidation (predominantly by phase II reactions) and sulfation (saturated); • Still a minority of the overdose is metabolized by Cyps to form NAPQI; • However, much higher amounts of NAPQI are formed after an overdose; is conjugated in part with GSH – leading to extensive depletion of GSH in centrilobular hepatocytes (>90%); • There is substantial protein adduct formation on cytosolic and mitochondrial proteins – although lower levels than found in mice with lower, toxic doses; • The mitochondrial adducts do not trigger any or a relevant initial mitochondrial oxidant stress and peroxynitrite formation that could cause JNK activation or mitochondrial dysfunction and DNA fragmentation. • As a result of the lack of mitochondrial dysfunction and no nuclear DNA damage, there is no relevant cell death. Overall, this further confirms that nuclear DNA damage is dependent on extensive mitochondrial dysfunction.	

References: 1. (McGill and Jaeschke, 2013), 2. (McGill et al., 2013), 3. (Hu et al., 1993), 4. (Heard et al., 2011), 5. (Heard et al., 2016), 6. (Kang et al., in press), 7. (Ni et al., 2016) 8. (Xie et al., 2015a), 9. (Xie et al., 2014), 10. (Cover et al., 2005b), 11. (Bajt et al., 2006) 12. (McGill et al., 2011), 13. (McGill et al., 2012b)

Under therapeutic dosing conditions (Figure 17) there is limited formation of NAPQI, which is bound to glutathione and to cellular proteins to a very limited extent and there is sufficient regeneration of glutathione to bind any NAPQI that is formed. At supratherapeutic doses (>4-8 g in humans), there is some depletion of glutathione and cellular proteins to detoxify NAPQI resulting in mitochondrial adduct formation and potential for disruption of mitochondrial respiration resulting in limited oxidative/nitrosative stress without any DNA damage and potential for isolated hepatocyte necrosis in the centrilobular region of the liver. Note that although CYP1A2 and CYP3A4 were considered based on human in vitro microsomal data (Patten et al., 1993; Raucy et al., 1989; Thummel et al., 1993), both enzymes were found to have negligible contribution in human in vivo studies (Manyike et al., 2000; Sarich et al., 1997).

Figure 17: Schematic diagram showing the molecular cascade within the hepatocyte following therapeutic (3-4 g) and supratherapeutic (> 4 – 8 g/day) of acetaminophen.

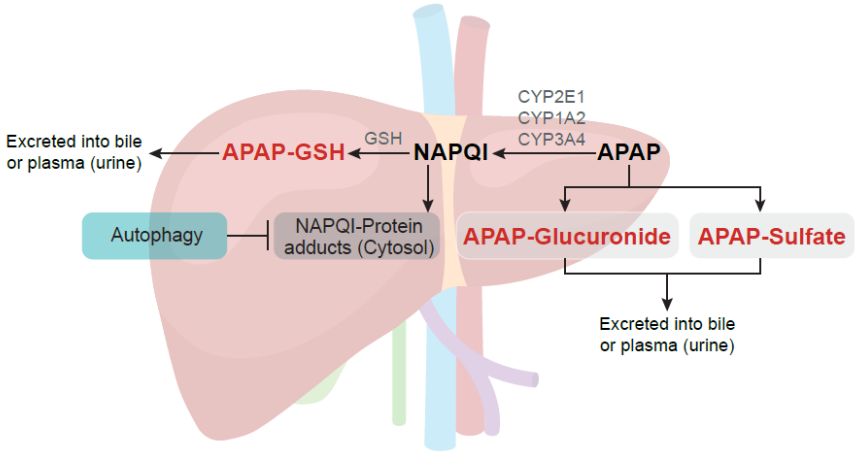
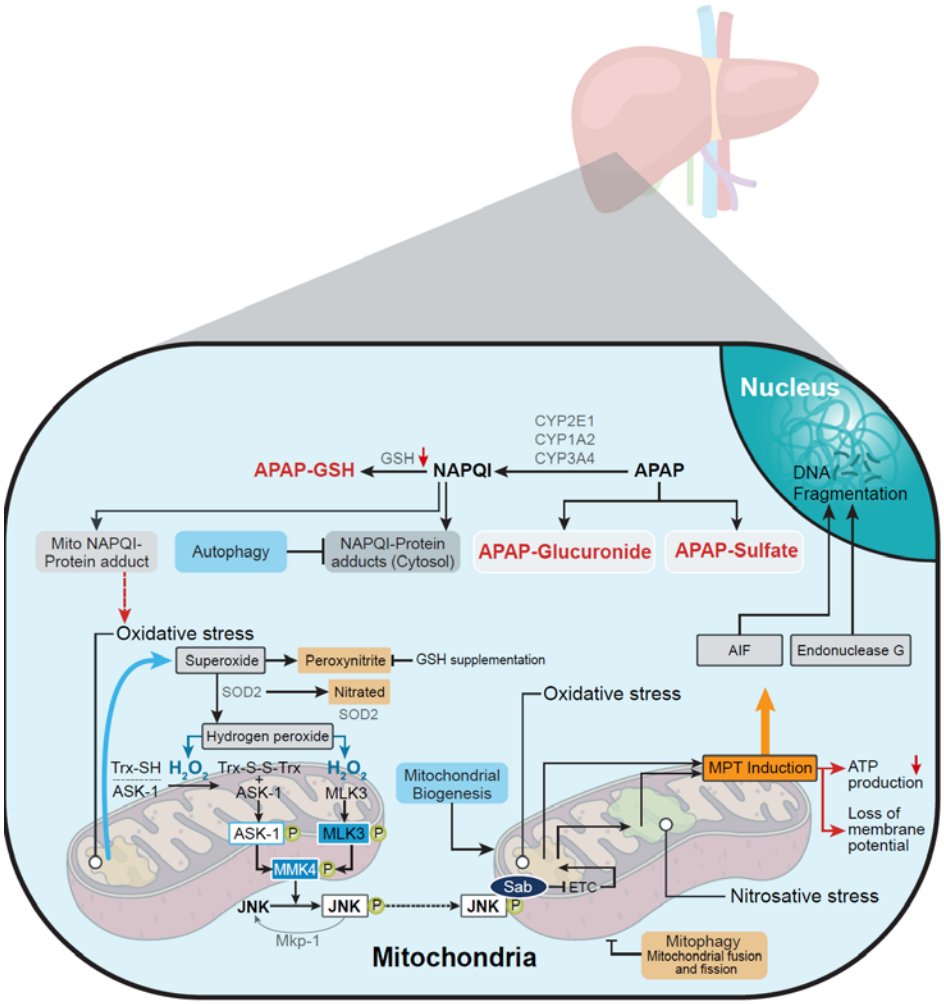


Figure 18: Schematic diagram showing the molecular cascade within the hepatocyte following overdose (> 10-15 g acute dose) of acetaminophen



6.2 Liver Cells Represent a Worst-Case Scenario for NAPQI Formation, Toxicity and Potential Adverse DNA Effects Compared to other Cell Types

Acetaminophen toxicity has been reported to occur very infrequently at supratherapeutic doses and on overdose in organs besides the liver (e.g. kidney) (Hoivik et al., 1995; Kennon-McGill and McGill, 2018). However, since hepatocytes represent a worst-case scenario for acetaminophen exposure and reactive metabolite formation, because of the much higher levels of CYP2E1 and the much higher concentrations of acetaminophen that reach the hepatocytes, in the sections that follow we focus on the data and evidence on adduct formation in hepatocytes across the different exposure conditions that make it exceedingly unlikely for acetaminophen to have any carcinogenic effects.

6.3 Glutathione/Protein Adduct Formation Protect Cells at Therapeutic and Toxic Doses

It is well established that a fraction of any acetaminophen dose is metabolized by cytochrome P450 enzymes leading to formation of the reactive metabolite NAPQI (McGill and Jaeschke, 2013). Although high GSH levels in hepatocytes can effectively detoxify NAPQI by forming a GSH-conjugate at therapeutic doses, very low levels of acetaminophen-cysteine protein adducts are detectable with sensitive mass spectrometric methods in both mice and humans (Heard et al., 2011; McGill et al., 2013). However, this minor adduct formation is pathophysiologically irrelevant as neither mitochondrial dysfunction nor DNA damage is detectable and any temporary loss of GSH is rapidly re-synthesized (McGill et al., 2013). On the other hand, an overdose of acetaminophen results in extensive GSH depletion and a dramatic increase in protein adduct formation. Although a number of these protein adducts have been identified (Cohen et al., 1997; Qiu et al., 1998), no critical protein adduct was identified that could cause cell death. However, any protein adducts of acetaminophen can be readily removed by autophagy ensuring the long-term survival of healthy cells even under chronic acetaminophen use (Ni et al., 2016).

6.4 Mitochondrial Adduct Formation and Toxicity is the Principle Mode of Toxicity and Drives Cell Death and Nuclear DNA Damage Occurs Only as a Consequence of Mitochondrial Dysfunction

In contrast to the glutathione and general protein binding, adducts formed in mitochondria have negative consequences that can lead to cell death. Early studies by Sidney Nelson's group demonstrated that when comparing acetaminophen (APAP) with its regioisomer N-Acetyl-m-aminophenol (AMAP), there was no difference in overall protein binding in mice but only acetaminophen caused protein adducts in mitochondria and induced liver injury (Tirmenstein and Nelson, 1989). These results were confirmed for mouse hepatocytes but not in human hepatocytes where AMAP caused mitochondrial adducts and cell death (Xie et al., 2015b).

Together, these findings suggest that only after a toxic dose of acetaminophen is there substantial protein adduct formation in mitochondria, which is critical for cell death.

A clear consequence of the mitochondrial adducts is a modest oxidant stress, which is insufficient to cause relevant mitochondrial dysfunction, but instead activates redox-sensitive mitogen activated protein kinases such as ASK-1 and MLK3 (Du et al., 2015; Han et al., 2013). The activation of this MAPK cascade results ultimately in the phosphorylation of c-jun N-terminal kinase (JNK) in the cytosol and the translocation of P-JNK to the mitochondria (Hanawa et al., 2008). The binding of P-JNK to the anchor protein Sab on the outer mitochondrial membrane results in a further interruption of the electron transport chain and enhanced electron leakage with formation of superoxide in the mitochondrial matrix (Win et al., 2016). This amplified oxidant stress inside the mitochondria leads to the enhanced formation of peroxynitrite, which is a potent oxidant and nitrating species (Radi, 2004).

6.5 Oxidative Stress is not a Direct Consequence of Acetaminophen Exposure or NAPQI Formation but Only Occurs as a Consequence of Mitochondrial Dysfunction that Leads to Cell Death

Peroxynitrite is the ultimate oxidant responsible for the cell injury after acetaminophen overdose (Knight et al., 2002). Importantly, it is confined to the mitochondria as indicated by selective mitochondrial DNA damage and nitrotyrosine protein adducts selective inside the mitochondria but not in any other compartment of the cell including the nucleus (Cover et al., 2005b). The limitation of the oxidative/nitrosative within the mitochondria is also documented by the selective increase of glutathione disulfide (GSSG) concentrations within the mitochondria (Jaeschke, 1990; Knight et al., 2001) and the use of MitoSox, which is a superoxide indicator that accumulates selectively inside the mitochondria (Yan et al., 2010). The pathophysiological importance of the mitochondrial superoxide formation is also demonstrated by the dramatically enhanced peroxynitrite formation and hepatotoxicity in MnSOD-deficient mice (Ramachandran et al., 2011) and the protective effect of the selective mitochondrial SOD mimetic Mito-Tempo (Du et al., 2017a). The critical role of MnSOD is to prevent the reaction of nitric oxide with superoxide to form peroxynitrite. The enhanced dismutation of superoxide to hydrogen peroxide and oxygen allows the detoxification of hydrogen peroxide by glutathione peroxidase. However, it theoretically enhances the risk of a Fenton reaction and lipid peroxidation. The fact that there is only very limited evidence for lipid peroxidation after acetaminophen overdose and that the lipid-soluble antioxidant vitamin E does not protect (Knight et al., 2003) further supports the hypothesis that peroxynitrite, which is limited to the mitochondrial space, is the critical oxidant in the pathophysiology (Du et al., 2016). Peroxynitrite triggers the opening of the mitochondrial membrane permeability transition pore (MPTP) resulting in the collapse of the membrane potential and cessation of ATP synthesis (Kon et al., 2004). If enough mitochondria are affected,

the cell undergoes necrosis. However, damaged mitochondria can also be removed by mitophagy (Ni et al., 2012) and then replaced by mitochondrial biogenesis (Du et al., 2017b) resulting in the survival of cells especially on the periphery of the necrotic area (Ni et al., 2013).

One of the consequences of the MPTP opening is mitochondrial matrix swelling, which leads to rupture of the outer membrane. In this case, intermembrane proteins such as cytochrome c, Smac/Diablo, endonuclease G and apoptosis-inducing factor (AIF) are being released into the cytosol. This can also be triggered by mitochondrial translocation of bax, which forms heterodimeric pores with other proteins such as bak, bad and bid in the outer membrane (Bajt et al., 2008). Despite the release of cytochrome c from the mitochondria, which could theoretically lead to activation of caspase-9 and trigger apoptosis, there is no evidence for relevant activation of any caspases or morphological characteristics of apoptosis after acetaminophen overdose (Jaeschke et al., 2018). The reason for the lack of apoptosis might be declining ATP levels. In contrast, endonuclease G and AIF translocate to the nucleus and cause DNA fragmentation (Bajt et al., 2006). Thus, this DNA damage is strictly dependent on mitochondrial dysfunction (Cover et al., 2005b) and based on the DNA fragments being produced. Acetaminophen-induced DNA damage is clearly different from caspase-activated DNase-mediated damage during apoptosis (Cover et al., 2005a; Jahr et al., 2001). This means that under conditions when significant DNA fragmentation occurs, the cell passed the point of no-return to necrosis, which makes it impossible that such a cell survives and initiates carcinogenesis.

6.6 DNA Adducts Have not Been Structurally Identified *in Vivo* at any Dose Level

In contrast to protein and mitochondrial adducts, there is no scientifically valid evidence for adduct formation on nuclear DNA after therapeutic or toxic doses of acetaminophen *in vivo*. The limited evidence that acetaminophen can form DNA adducts comes from *in vitro* studies (Dybing et al., 1984; Hongslo et al., 1994; Rogers et al., 1997), and a mouse *in vivo* study (Rogers et al., 1997). These studies show a dose-related increase in the extent of DNA binding of a tritiated label at therapeutic and supratherapeutic concentrations and doses. In addition, comparisons of the relative binding of the tritiated label to the DNA, chromatin, and nucleus demonstrate that almost all of the label was on the chromatin and nucleus, and not on the DNA, which would suggest that the label is binding to histones and proteins rather than the DNA itself. However, the authors only measured radioactivity ***in*** the DNA and assumed this reflected binding of acetaminophen ***to*** DNA; the tritium label can be readily displaced and enter the general cellular pool such that it gets incorporated into normal bases and thence into DNA (metabolic incorporation) and not represent acetaminophen. There was also no clear induction of adducts in liver DNA using the ^{32}P-postlabeling technique. No DNA adducts were identified or characterized, and, as indicated above, the presence of radioactivity in DNA does not prove that adducts have been formed (Phillips et al., 2000). As stated by Bergman et al. (1996) "Definite proof that the covalent binding of radioactivity from ^{3}H-labelled paracetamol to DNA represents

the formation of true DNA adducts would require chemical structural analysis". In conclusion, there is no meaningful evidence of any nuclear DNA adduct formation in humans or animals *in vivo*.

The discussed mechanisms of acetaminophen hepatotoxicity are mainly based on studies in primary mouse hepatocytes or mice *in vivo*, which are the most relevant model for the human pathophysiology (McGill et al., 2012a). However, all critical signaling events in mice have been confirmed in either primary human hepatocytes (Xie et al., 2014), metabolically competent HepaRG cells (McGill et al., 2011) and in acetaminophen overdose patients (Davern et al., 2006; McGill et al., 2012a; McGill et al., 2014). These events include reactive metabolite formation and protein adducts, JNK activation and mitochondrial oxidant stress, mitochondrial dysfunction, DNA fragmentation and cell necrosis. Importantly, these events can be effectively prevented in human hepatocytes by a cytochrome P450 inhibitor (Akakpo et al., 2018) or in overdose patients when treated early with the antidote N-acetylcysteine (McGill et al., 2012a). This indicates that oxidant stress and DNA damage during acetaminophen hepatotoxicity in mice or humans are strictly dependent on the toxic signaling events leading to cell necrosis.

6.7 Potential Effects of Acetaminophen on DNA repair or Genomic Stability in Nonclinical Test Systems are not Relevant to Humans

Several studies show a potential inhibitory effect of acetaminophen on reparative and replicative DNA synthesis *in vitro* and *in vivo* using a thymidine uptake assay. It has been proposed that, by analogy with hydroxyurea, this may be a result of the inhibition of ribonucleotide reductase and may explain genotoxic effects seen at high doses (Bergman et al., 1996; Thybaud et al., 2007).

The following factors suggest that there is insufficient evidence to support that the results showing potential effects on ribonucleotide reductase *in vitro* have any relevance to the carcinogenicity hazard potential of acetaminophen based on the following:

- There are no studies showing direct binding of acetaminophen to ribonucleotide reductase (Hinson et al., 2004)
- No data was identified demonstrating that acetaminophen inhibits ribonucleotide reductase or disrupts the ribonucleotide pool *in vivo*
- Studies claiming that there is inhibition of ribonucleotide reductase have been conducted in in vitro model test systems that have highly questionable relevance to humans or animals (e.g. mouse mammary immortalized tumor cell line with mutations introduced (Hongslo et al., 1990)); in multiple *in vitro studies* the conditions tested are implausible in humans (i.e. high concentrations for 48 hours in a static system).
- The reduced thymidine uptake is transient, reversing in vivo within 2 to 4 hours (Hongslo et al., 1994; Lister and McLean, 1997).

- There is no evidence that the effects are sustained with multiple dosing at therapeutic or non-toxic supratherapeutic doses and lead to sustained DNA effects at non-toxic concentrations.
- There are other potential mechanisms, besides direct inhibition of ribonucleotide reductase, that could cause the effects seen in these *in vitro* model systems by Hongslo et al. (Hongslo et al., 1990). One potential alternative mechanism for the effects observed on DNA repair could be acetaminophen induced mitochondrial permeability transition *in vitro* that occurs in two phases (glutathione depletion/covalent binding followed by mitochondrial dysfunction). Mitochondrial dysfunction can drive toxicity, inhibit ribonucleotide reductase function in the cytosol (Desler et al., 2010; Desler et al., 2007) and the *in vitro* effects on DNA that were observed.

When viewed in the context of the negative carcinogenicity studies and other genetic toxicology studies, the data support that this mechanism does not represent a genotoxic or carcinogenic hazard to humans.

Another potential mode of action for genomic instability that is observed with acetaminophen involves elevation of intracellular Ca^{2+}. High cytotoxic concentrations/doses of acetaminophen induce a marked increase in intranuclear Ca^{2+}, resulting in endonuclease activation and DNA fragmentation, such that any genotoxic effects may be a consequence of cytotoxic events, and, as discussed earlier, affected cells would not survive. Thus, since increased Ca^{2+} levels are only associated with high cytotoxicity, any resultant genotoxicity will exhibit a threshold. Human plasma concentrations under normal acetaminophen usage are much lower than cytotoxic concentrations, so that under normal usage acetaminophen would not induce genotoxicity associated with increased Ca^{2+} levels. Under conditions of overdose, the high cytotoxicity will mean the cells containing genotoxic damage will not survive.

6.8 Receptor Pharmacology and High Throughput Screening (HTS) Data Show no Evidence of Carcinogenic Potential

Acetaminophen was tested in ToxCast/Tox21 for activity in 309 *in vitro* assays that are relevant to one or more of the KCCs (Table HTS). Acetaminophen was inactive in 306 of the assays, 289 of which were tested in human models (94%, cell lines or cell-free). The three active assay endpoints were all tested in human cell models, and were related to epigenetic alterations, progesterone receptor binding, and androgen receptor antagonism. However, these active assays were all flagged for data quality issues, and the activity in these assays was inconsistent with other assays that test for similar signals. Thus, acetaminophen was generally considered inactive in HTS assay endpoints related to the KCCs at concentrations up to 200 µM.

In conclusion, when viewed in context of the preclinical findings, which would account for many of the limitations in interpretation of *in vitro* assays, as well as account for activity associated with metabolites (even following chronic exposure to very high doses), the activity observed in the HTS data are without biological significance. Numerous preclinical assays demonstrate a lack

of adversity associated with the molecular or cellular signals obtained in the ToxCast/Tox21 assays.

6.9 Mechanisms of Pharmacological Action May be Protective Against Carcinogenicity

Current evidence regarding acetaminophen's analgesic mechanism of action has been proposed to involve 1) the inhibition of cellular prostaglandin production (Anderson, 2008; Graham and Scott, 2005), 2) increased cannabinoid receptor activity (Anderson, 2008; Hogestatt et al., 2005; Sharma and Mehta, 2014), 3) the inhibition of nitric oxide production (Sharma and Mehta, 2014) and 4) anti-oxidant/peroxynitrite scavenging properties (Dou et al., 2017; Schildknecht et al., 2008).

Studies conducted *in vitro* have shown that acetaminophen at pharmacologically relevant concentrations acts as a cellular peroxynitrite scavenger (Dou et al., 2017; Schildknecht et al., 2008), suggesting it may have a protective effect against oxidative stress, and therefore even could be protective against potential carcinogenesis. In tissue and *in vivo* animal studies acetaminophen has also been shown to reduce ROS/RNS in multiple tissue types (Blough and Wu, 2011). Acetaminophen has been shown to have antioxidant effects in the rat liver (DuBois et al., 1983) and acetaminophen (20 mg/kg) has also been shown to decrease liver mitochondrial H2O2 formation in both control and HF diet fed mice (Shertzer et al., 2008). Acetaminophen has also been shown to have protective effects at low doses on renal injury in a Zucker rat obesity model for renal injury; the effects appear to be mediated, at least in part, through attenuation of ER stress (Wang et al., 2014).

There are also several reports of anti-proliferative and anti-tumor effects of acetaminophen in different nonclinical models. Bush et al reported that acetaminophen "exhibited antiproliferative activity against all tested ovarian cancer cell lines" *in vitro* and describe potential pathways driving its antiproliferative effects (Bush et al., 2016). Takehara, et al. (2011) demonstrated that a breast cancer stem cell line treated with acetaminophen *in vitro* resulted in the loss of their tumorigenic ability in nude mice. Furthermore, administration of acetaminophen inhibited the growth of tumor xenografts of MDA-MB-231 cells in both the presence and absence of simultaneous administration of doxorubicin, a typical anti-tumor drug for breast cancer.

6.10 Clarification of Acetaminophen Metabolism in Humans versus Rodents

Acetaminophen pharmacokinetics and metabolism have been extensively studied over the past 60 years, generating hundreds of publications. OEHHA reviewed over 400 studies, using several published reviews as an initial guide for their selection. Sections in the HID on the absorption, distribution, and excretion of acetaminophen appropriately summarize the extensive human data available. However, in the metabolism section, "*data from animal studies are included when human data are unavailable or incomplete"*, a statement which assumes these data apply to

humans, and information from human and animal studies are often intermingled. It is not clear which information provided is relevant to humans, especially as it pertains to enzymes involved in acetaminophen metabolism and purported reactive metabolites. Therefore, Figure 5 from the HID (page 142) is reproduced below (Figure 19) with overlay marks to distinguish evidence-based, confirmed metabolic pathways and metabolites in humans versus other proposed or documented metabolites in rodents and *in vitro* tests. Some of the reactive metabolites have only been identified in rodents at hepatotoxic doses.

Although humans and animals share some, but not all, of the reported pathways and metabolites of acetaminophen, results from animal studies and *in vitro* tests should not be indiscriminately extrapolated to humans (Caparrotta et al., 2018; Prescott, 2000; Rumack, 2004). Studies in various animal models and *in vitro* tests are hypothesis generating, necessitating confirmation and elucidation in subsequent *in vivo* human studies. Given important species differences in acetaminophen metabolism and by dose, some extrapolated suppositions regarding metabolic outcomes in humans have been proven incorrect or not clinically significant through human studies (Prescott, 2000; Rumack, 2004).

Some misconceptions and incorrect interpretations of human acetaminophen metabolism based on animal studies and *in vitro* tests remain today. They continue to be cited in scientific and medical reviews, and online medicinal product forums. A high-level summary of human metabolism and comparison of species differences are presented in this section to help clarify to the Committee which acetaminophen metabolites are relevant to humans.

Figure 19: Figure 5 from the HID (page 142) is reproduced below with overlay marks to distinguish evidence-based, confirmed metabolic pathways and metabolites in humans versus other proposed or documented metabolites in rodents and in vitro tests

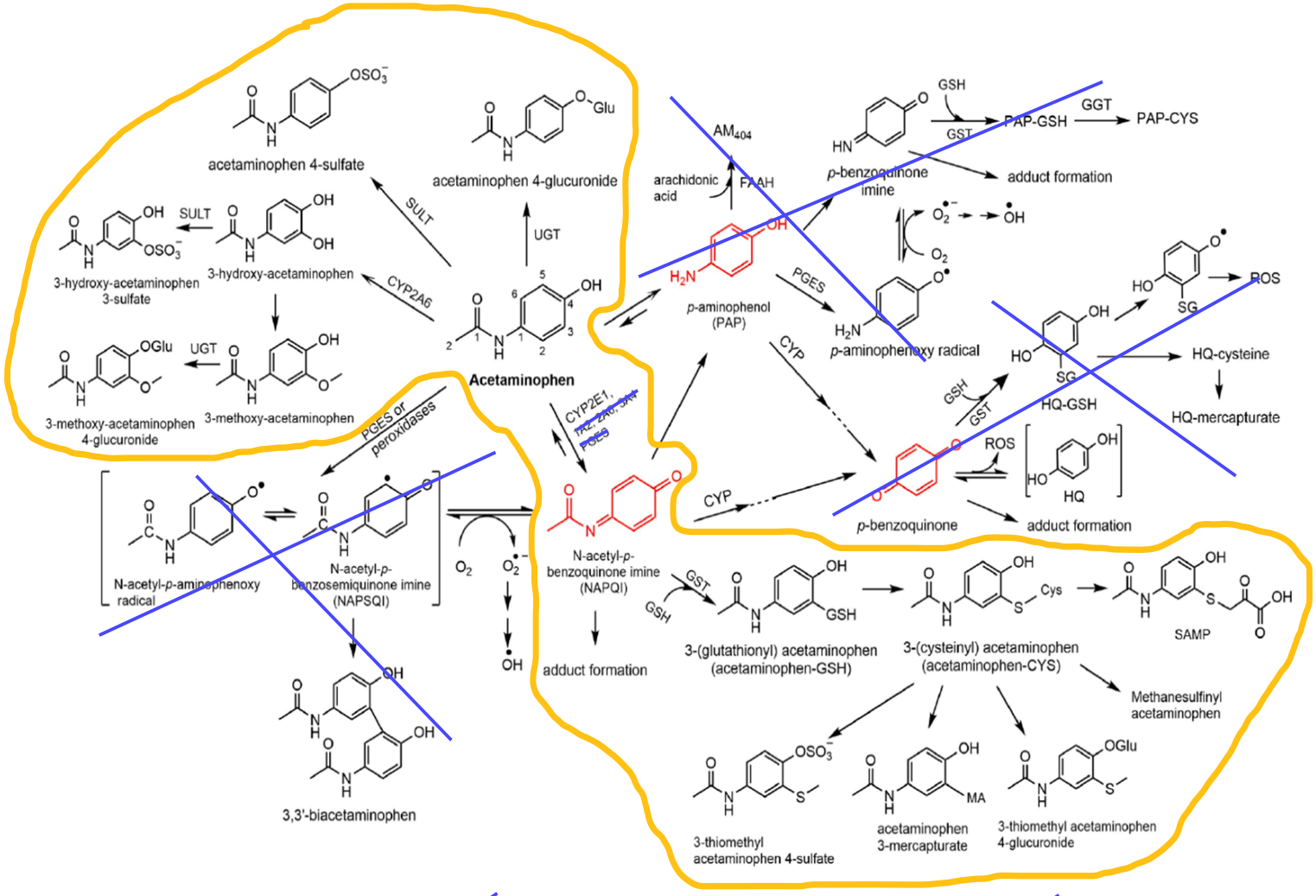


Figure 5. The proposed metabolism of acetaminophen in humans and animals

6.10.1 High-level Summary of Human Acetaminophen Metabolism

Acetaminophen undergoes mixed-competitive and sequential biotransformation, primarily in the liver. Three main pathways are involved: conjugation with glucuronide, conjugation with sulfate, and oxidation via cytochrome P450 (CYP450) enzymes, which are included within the yellow boundary of Figure 5 from the HID. Glucuronidation is the main metabolic pathway in adults, whereas the sulfate conjugate predominates in premature infants, newborns, and young infants because hepatic glucuronidation is relatively immature at birth (Gow et al., 2001; Miller et al., 1976).

Acetaminophen is conjugated with glucuronic acid by UDP-glucuronosyltransferases (UGT), specifically the isoforms UGT1A6, UGT1A9, and UGT2B15 (Baker et al., 2005; Court et al., 2001; Miners et al., 2011; Mutlib et al., 2006). Acetaminophen is a substrate for three sulfotransferases, SULT1A1, SULT1A3, and SULT1C4 (Pacifici, 2004). Sulfation of acetaminophen is partly governed by the availability of inorganic sulfate, which is rate limiting in the formation of the cofactor of sulfation, 3'-phosphoadenosine-5'phosphosulfate (PAPS). The other rate-limiting reaction is sulfotransferase activity. With repeated therapeutic and supratherapeutic dosing of acetaminophen, sulfotransferase activity decreases forming less sulfate conjugate; whereas acetaminophen induces UGT enzymes forming more glucuronide conjugate (Brown et al., 2008; Gelotte et al., 2007; Hindmarsh et al., 1991). The fraction of acetaminophen dose oxidized by the CYP450 pathway remains relatively the same as measured by urine excretion.

The main oxidative pathway forms the highly reactive intermediate, *N*-acetyl-*p*-benzoquinone imine (NAPQI), which is conjugated with glutathione (GSH) to form cysteine, mercapturate, methylthio-, and methanesulfyinyl-APAP metabolites (Mitchell et al., 1974). These inert thiol metabolites circulate either free in plasma or conjugated with glucuronide or sulfate. The principal CYP450 isoenzyme involved *in vivo* is hepatic CYP2E1. Although CYP1A2 and CYP3A4 were considered other pathways for NAPQI based on human *in vitro* microsomal data (Patten et al., 1993; Raucy et al., 1989; Thummel et al., 1993), both enzymes were found to have negligible contribution in human *in vivo* studies (Manyike et al., 2000; Sarich et al., 1997).

After large acetaminophen overdoses when GSH stores are reduced or depleted, excess NAPQI forms protein adducts through binding to cysteine groups, primarily on mitochondrial proteins, leading to hepatic cell death (Mazaleuskaya et al., 2015). More recently, low concentrations of APAP-cysteine protein adducts were detected in adults after consuming the maximum daily dose (4 g/d) for 2 days and reaching a plateau at 7 days (Heard et al., 2016; Heard et al., 2011). A small amount of an acetaminophen dose is oxidized by CYP2A6 to form 3-hydroxyacetaminophen and 3-methoxyacetaminophen (Andrews et al., 1976; Slattery et al., 1989).These catechol metabolites are further conjugated with glucuronide or sulfate (Chen et al., 1998).

6.10.2 Differences in Acetaminophen Metabolism Among Species

Table 18 highlights important quantitative differences involving both metabolic activation and parallel nontoxic conjugation of acetaminophen among humans, rats, and mice (Prescott, 1996). The metabolite pattern varies among species and by differences in dose, routes of administration, and experimental conditions and indicate that the results from *in vitro* test systems and animal models with acetaminophen need to be viewed with caution. Glucuronidation of acetaminophen is the predominate conjugation pathway in humans and mice, whereas sulfation is the predominate conjugation pathway in rats. Mice form the largest amount of acetaminophen cysteine via the reactive intermediate, NAPQI, making them more sensitive to hepatotoxic doses.

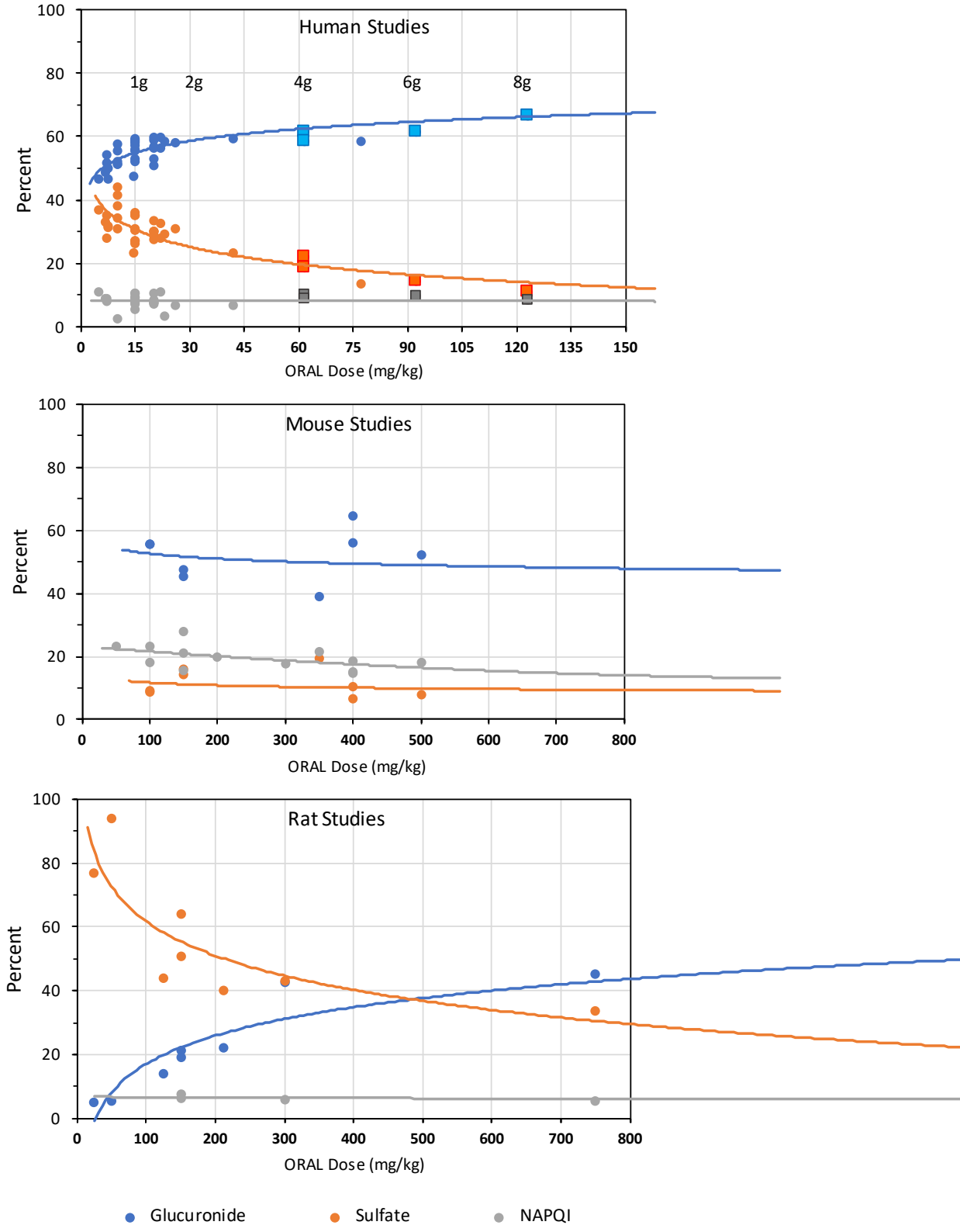
Table 18. Urine Metabolite Pattern[a] of Acetaminophen Varies Among Species

Metabolite	**Human**	**Rat**	**Mouse**
Parent, free Acetaminophen (APAP)	2 - 4		
APAP-Glucuronide	45 - 65	10 - 20	50 - 60
APAP-Sulfate	25 - 35	40 - 80	10 - 20
Thiols via NAPQI			
APAP-Cysteine (and conjugates)	2 - 6	1 - 5	15- 25
APAP-Mercapturate (and conjugates)	3 - 6	2 - 6	1 - 3
Methylthio-APAP (and conjugates)	2 - 3[b]		1.6[c]
Methanesulfinyl-APAP (and conjugates)			trace[c]
Catechols		1.8	
Hydroxy-APAP (and conjugate)	3 - 8		*3.7*[c]
Methoxy-APAP (and conjugate)	3 - 6		*2.7*[c]

a: Adapted from Tables 6.1-6.3 in Prescott 1996; b: Sum of methylthio- and methanesulfinyl-APAP (Gelotte et al., 2007); c: Percent of 250 mg/kg dose (hepatotoxic in mice) excreted, Rashed et al., 1990;

Not only does the urine metabolite pattern vary by species, it further varies by administered dose. Figure 20 illustrates the dose dependence of acetaminophen glucuronidation, sulfation, and thiol formation (via NAPQI) by species. Metabolite data are presented as percent excreted in urine relative to either the administered dose or the total amount of metabolites recovered. Each panel includes a scatter plot of mean values from published studies of single oral doses (circles) in humans, mice, and rats. Only the human panel includes mean values from multiple daily doses identified as squares (Gelotte et al., 2007). Trendlines are overlaid to highlight the discordant shifts in metabolite patterns with increasing doses among these three species.

Figure 20. Acetaminophen Metabolite Differences by Species and Dose, Expressed as Percent Excreted in Urine Relative to Either Acetaminophen Dose or Total Amount of Metabolites.



For reference, therapeutic single doses in humans range from about 7.5 to 20 mg/kg or 500 to 1300 mg. Systemic exposures (area under the plasma concentration-time curve, AUC) of acetaminophen for these human doses correspond to approximately 92 to 246 mg/kg and 47 to 124 mg/kg doses in mice and rats (scaled based on human equivalent dose based on body surface area with a conversion factor of 12.3 for mice and 6.2 for rats (Nair and Jacob, 2016).

6.10.3 Species Differences in the Metabolic Activation of Acetaminophen

Several pages in the HID address the metabolic activation of acetaminophen by cytochrome P450, but most data and information are summarized from *in vivo* studies and *in vitro* tests in rodents. Regarding the various isoenzymes associated with CYP450, Prescott said, *"Because of dose dependence and species differences in the expression, activity and inducibility of these isoenzymes, it is not justifiable to extrapolate the results of animal studies to clinical conditions in man*"(Prescott, 2000). It is well accepted and confirmed in humans that about 5 to 12% of an acetaminophen dose is oxidized to NAPQI via by CYP2E1 and conjugated with GSH, and undergoes further transformation to thiol metabolites. A small fraction of acetaminophen is oxidized by CYP2A6 to catechol metabolites (Andrews et al., 1976; Slattery et al., 1989).

An overview of the differences in acetaminophen oxidation by CYP450 isoenzymes and deacetylation among humans and rodents is presented in Table 17, referencing several studies cited in the HID. It becomes clear that there is no meaningful evidence of the formation of additional reactive metabolites (or their transformed species) beyond NAPQI and its thiol metabolites in humans.

In rodents, acetaminophen is believed to be deacetylated to form p-aminophenol (PAP) at hepatotoxic doses in mice and hamsters or shown *in vitro* in rats (Gemborys and Mudge, 1981; Mugford and Tarloff, 1995; Newton et al., 1982; Rashed et al., 1990). PAP can become a reactive intermediate after undergoing enzymatic or nonenzymatic oxidation and cause cellular damage resulting in nephrotoxicity. However, during 60 years of clinical investigations, PAP has not been identified as a metabolite of acetaminophen in humans in any prospective, well-controlled metabolism study. Two studies in the HID were cited as evidence for PAP being an acetaminophen metabolite in humans. One study apparently detected PAP in urine from three patients after large acetaminophen overdoses of 40, 50, and 75 g (Clark et al., 1986), but proof of identity was not rigorous and was based on a nonspecific color reaction and thin layer chromatography using only one solvent system (Prescott, 1996).

The second, uncontrolled study was designed to quantify acetaminophen and PAP in urine from male partners of couples planning for pregnancy and to search for associations of each compound with semen quality (Smarr et al., 2017). Acetaminophen and PAP were detected in urine from 93% and 100% of the study population, respectively. However, the investigators claimed that PAP was a metabolite of acetaminophen with no credible evidence. In fact, they

noted a critical limitation of their study: "*In LIFE, information on sources of potential paracetamol exposure (for example, self-report of medication use, occupational or environmental exposures) was not collected (Smarr et al., 2017)*." Also, it is well known that occupational and environmental continuous exposure to aniline is ubiquitous, and that aniline is readily metabolized to acetaminophen and PAP via separate pathways and excreted in urine (Dierkes et al., 2014; Holm et al., 2015).

In a comprehensive metabolomics analysis of serum and urine from adult volunteers who were administered oral daily doses of 0.5, 2 and 4 g acetaminophen, 22 metabolites, including conjugates, were identified using a combination of nuclear magnetic resonance (NMR), liquid chromatography–mass spectrometry (LC-MS) and/or gas chromatography–mass spectrometry (GC–MS) (Jetten et al., 2012). Neither PAP nor its conjugates were identified in serum or urine at any of the doses using these highly sensitive assay techniques.

Table 19. Acetaminophen Oxidation by Cytochrome P450 Enzymes and Deacetylation in Various Test Systems

Pathway / Metabolite	***In Vivo* Human Studies**	***In Vitro* Human Tests**	**Human References**	***In Vivo* Rodent Studies**	***In Vitro* Rodent Tests**	**Rodent References**
NAPQI Intermediate to Thiols	CYP2E1 5-12% of dose	CYP2E1, CYP1A2, CYP2A6, CYP3A4, CYP2D6	***In Vivo*:** (Manyike et al., 2000; Sarich et al., 1997), ***In Vitro*:** (Bloom et al., 2019; Laine et al., 2009; Thummel et al., 1993)	CYP2E1, CYP1A2 (mice)	CYP3A1 (rats)	***In Vivo*:** (Snawder et al., 1994) **In Vitro:** (Prasad et al., 1990)
Catechols	CYP2A6 3-8% of dose	CYP2A6	***In Vivo*:** (Gelotte et al., 2007; Slattery et al., 1989) ***In Vitro*:** (Chen et al., 1998)	6.4% of dose[b] in mouse urine; 1.8% fractional recovery in rat urine		***In Vivo*:** (Forte et al., 1984; Rashed et al., 1990; Thummel et al., 1988) ***In Vitro*:** (Chen et al., 1998)
p-aminophenol	No evidence[a]	No evidence	***In Vivo*:** (Clark et al., 1986; Smarr et al., 2017)	1.7% of dose[b] in mouse urine; 0.1-0.9% of dose[c] in hamster urine; 1.5-3.6% of dose[d] in rat urine	Incubations: perfused rat kidneys and microbes from rat cecum	***In Vivo*:** (Gemborys and Mudge, 1981; Rashed et al., 1990); (Newton et al., 1983) ***In Vitro*:** (Mugford and Tarloff, 1995; Newton et al., 1982)
p-benzoquinone	No evidence	No evidence	None	Indirect evidence in mice: posited as metabolites	Purified CYPs from pheno-barbital treated rats	***In Vivo*:**(Pascoe et al., 1988) ***In Vitro*:** (Dahlin et al., 1984; Eastmond, 1993)
NAPSQI Intermediate	No evidence	No evidence	None	No evidence	Horseradish peroxidase, purified CYPs	***In Vitro*:** (Potter and Hinson, 1987a, b, 1989)
4-aminophenoxyl free radicals	No evidence	No evidence	None	No evidence	*In vitro* tests	***In Vitro*:** (Fischer et al., 1985; Josephy et al., 1983; Potter and Hinson, 1987a, b; West et al., 1984)

a: p-Aminophenol has not been confirmed as a metabolite of acetaminophen in humans in any prospective, well-controlled metabolism study. See text for discussion.
b: Percent of 250 mg/kg dose (LD50) in mice
c: Percent of metabolites excreted at 50 to 300 mg/kg in hamsters; LD50 is 350 mg/kg
d: Percent of recovered dose from 250 to 750 mg/kg in rats
Key: Catechols – 3-hydroxyacetaminophen and 3-methoxyacetaminophen; NAPQI – N-acetyl-p-benzoquinone imine; NAPSQI – N-acetyl-p-benzosemiquinone imine

6.11 Relevance of Metabolites and Data on Structural Analogs

Several investigators have raised the potential for formation of reactive metabolites of acetaminophen besides NAPQI, including p-benzoquinone and p-aminophenol (McGill and Jaeschke, 2013). We are not aware of any data showing that these metabolites are formed in humans. In addition, given that they have only been detected in rodents, the negative NTP carcinogenicity studies demonstrate that if they are formed, they do not cause cancer in rodents at the levels that they were formed in these bioassays. Therefore, the carcinogenicity and genotoxicity data for p-benzoquinone and p-aminophenol should not be considered in the hazard assessment of the carcinogenicity of acetaminophen.

The HID presents the carcinogenicity of phenacetin as a potential source of concern for acetaminophen because acetaminophen is a metabolite of phenacetin. There is no scientific basis for this concern. Phenacetin has been reported to be a carcinogen in man, rats and mice, inducing urothelial tumors of the renal pelvis and tumors in the nasal cavity (Angervall et al., 1969; Bengtsson et al., 1968; Isaka et al., 1979; Johansson et al., 1974; Nakanishi et al., 1982; Taylor, 1972). On the other hand, acetaminophen, which is the major metabolite of phenacetin (Brodie and Axelrod, 1949; Nery, 1971b), induced no urothelial tumors nor tumors in the nasal cavity in rats nor mice. These findings suggest that hydroxylated metabolites of phenacetin, not acetaminophen, are likely the proximal mutagens and carcinogens, as concluded by multiple studies (Calder et al., 1976; Nery, 1971a; Shudo et al., 1978).

6.12 Implications of Mode of Action to Hazard Potential in Sub-Populations

6.12.1 Patient Variability in Metabolism

Metabolism of acetaminophen varies among individuals as a result of genetic polymorphisms and nongenetic factors (Court et al., 2017; Critchley et al., 1986; van der Marel et al., 2003). Yet, given the dominant pathways of glucuronide and sulfate conjugation (~85-90%), small changes in oxidation to NAPQI, if they occur, are not clinically significant and often fall within the expected range for therapeutic doses (de Morais et al., 1992; Forrest et al., 1979; van Rongen et al., 2016; Zapater et al., 2004).

6.12.2 Patients with Purported Susceptibility to Liver Injury

A recent critical review of the literature concluded that no patient group is unequivocally at elevated risk of acetaminophen -induced liver toxicity (Caparrotta et al., 2018). This review included clinical studies addressing genetic and nongenetic factors that may alter acetaminophen metabolism, such as enzyme polymorphism, race/ethnicity, Gilbert's syndrome, liver disease, age, obesity, nutritional state, alcohol use, and potential drug interactions. It excluded animal studies, given important species differences in metabolism making extrapolation to humans

inappropriate. Other reviews of clinical data addressed acetaminophen use by liver-impaired patients (Hayward et al., 2016) and by populations in which low glutathione has been observed (Lauterburg, 2002), concluding no evidence for greater risk. Another review highlighted common misconceptions of purported drug, alcohol, and fasting interactions with acetaminophen that were based on data from animal studies, *in vitro* tests, and case reports (Rumack, 2004).

Simulations have been performed to evaluate the potential for acetaminophen to be a hazard in patient sub-populations and in overdose patients using a Quantitative Systems Toxicology Platform called DILIsym that has been developed and validated using acetaminophen. These simulations support that there is also not a carcinogenicity hazard in patients with susceptibility for liver injury. The methodology and results of these simulations can be found in a separate supplementary document that has been made available to the CIC.

6.12.3 Overdose Patients

Clinical Evidence Supports Complete Recovery on Overdose and No Carcinogenic Hazard

The histology of liver injury due to significant overdoses of acetaminophen in rodents and man is well described. In man, key histologic features vary from limited centrilobular necrosis to confluent necrosis in more serious cases. In those subjects who recover from this injury, complete recovery characterized by normalization of liver function and restoration of hepatic architecture is the typical pattern. A single case report (Baeg et al., 1988) and case series from the 1970s have evaluated histology of overdose subjects both in the acute phase and generally after 3 months post overdose (Clark et al., 1973; Hamlyn et al., 1977; Lesna et al., 1976; Portmann et al., 1974). In the case series, in most patients at follow up biopsy necrotic zones were found to have been completely reconstituted with restoration of hepatic architecture. In a very small fraction of biopsied patients, minor abnormalities and fibrosis were seen. Fibrosis, if it occurred, was generally mild, and was seen only in very severe cases of injury. In many patients, serial biopsies demonstrated resolution of fibrosis over several months. Regarding acetaminophen as the potential causative factor of the fibrosis, this is not possible due to absence of pertinent medical information to rule out other potential etiologic factors (ETOH, viral, NASH etc.). In conclusion, the clinical data on liver injury from acetaminophen overdose when it does not require a liver transplant demonstrates that the injury resolves fully with no evidence of the type of chronic liver disease that would carry an increased risk for cancer.

7 Conclusions

This document provides a comprehensive weight of evidence assessment of the available animal carcinogenicity, genotoxicity, mode of action and epidemiology data. The human and animal studies are numerous and reassuring that acetaminophen is not a carcinogenic hazard at any dose level. The genetic toxicology and mode of action data help explain why we do not see a

signal of carcinogenicity in humans and laboratory animals. In conclusion, based on the weight of evidence, acetaminophen has not been clearly shown to cause cancer.

8 Appendices

8.1 Assessment of Epidemiologic Evidence by Cancer Site: Summaries of Publications

8.1.1 Urinary Tract System: Kidney, Renal Pelvis and Ureter, Bladder

(i) Urinary Tract Cancers (combined or not specified)

Cohort Studies

Two cohort studies conducted on urinary tract cancers and acetaminophen use both reported no association (Friis et al., 2002; Walter et al., 2011a). Fries et al 2002 used data from the Prescription Database of North Jutland County and the Danish Cancer Registry to compare cancer incidence among individuals ever prescribed with acetaminophen versus the expected incidence based on the North Jutland population who did not receive acetaminophen prescriptions. After a 9-year follow-up period, the standardized incidence ratio (SIR) for urinary tract cancer among those prescribed with acetaminophen, but not with aspirin or NSAIDs (N=13,482) was 1.0 (95% CI 0.7-1.4).

Walter et al 2011a assessed self-reported acetaminophen use over the previous 10 years in the VITamins And Lifestyle (VITAL) cohort study (N=62,841) (Walter et al., 2011a). After a mean follow-up period of 6.5 years, the reported adjusted hazards ratio (aHR) for urinary tract cancer among low acetaminophen users (<4 days/week or <4 years) was 1.1 (95% CI 0.76-1.59) compared to non-users. The aHR among high acetaminophen users (≥4 day/week and ≥4 years) was 1.05 (95% CI 0.6-1.83) compared to non-users.

Case Control Studies

Six case-control studies were conducted on urinary tract cancers and acetaminophen use (Ross et al 1989 is not shown in the forest plot since the 95% CI was not provided in the study). One study reported increased risk associated with self-reported acetaminophen use (Steineck et al., 1995). Steineck et al 1995 assessed self-reported acetaminophen use in a population-based case-control study in Sweden (N=325 cases and 393 controls). Compared to non-users, the adjusted odds ratio (aOR) for squamous or transitional cell carcinoma among ever users of acetaminophen was 1.6 (95% CI 1.1-2.3).

The other 5 case-control studies did not report increased risk for urinary tract cancers associated with ever/regular acetaminophen use (Linet et al., 1995; McCredie and Stewart, 1988; Pommer et al., 1999; Rosenberg et al., 1998; Ross et al., 1989). McCredie and Stewart 1988 assessed self-reported acetaminophen use in a population-based case-control study in Australia (N=55 cases and 688 controls). Compared to those unexposed to acetaminophen, the aOR for ureter cancer among those with ≥ 0.1 kg lifetime consumption of acetaminophen was 2.0 (95% CI 0.8-4.5). A

no increased risk was observed among those with lower level of lifetime consumption of acetaminophen (≥0.1 kg).

Ross et al 1989 assessed self-reported acetaminophen use in a population-based case-control study in the US (N=187 cases and 187 controls). Compared to those without regular acetaminophen use, the unadjusted RR for renal pelvis and ureter cancer among regular users (>30 day/year) was 1.3. The unadjusted RR among those with >30 consecutive days acetaminophen use was 2.0. No confidence intervals were calculated, but the findings were not significant as reported by the authors.

Linet et al 1995 assessed self-reported acetaminophen use in a population-based case-control study in the US (N=502 cases and 496 controls). Compared to those without regular acetaminophen use, the aOR for renal pelvis and ureter cancer among regular users (≥2 doses/week for ≥1 month) was 1.0 (95% CI 0.6-1.8).

Rosenberg et al 1998 assessed self-reported acetaminophen use in a hospital-based case-control study in the US (N=498 cases and 8,149 controls). Compared to never users of acetaminophen, the aRR for transitional cell cancer among those with regular use for a duration of ≥5 years (≥2 days/week for ≥1 month) and began at least 1 year prior hospitalization was 1.1 (95% CI 0.5-2.6).

Pommer et al 1999 assessed self-reported acetaminophen use in a population-based case-control study in Germany (N=647 cases and 647 controls). Compared to non- or rare users, the aOR for renal pelvis and ureter cancer among those with ≥1 kg cumulative lifetime use was 2.25 (95% CI 0.28-17.96).

Assessment of Evidence

Given the limitations of the studies above and that both cohort studies and 5 of the 6 case control studies did not report an increase in RR, it cannot be concluded that acetaminophen use is clearly shown to cause increased risk for urinary tract cancer. Across 8 studies (including Ross et al 1989 which is not included in the Forest plot/Figure 4), there are certain methodological limitations that should be considered when interpreting their results:

- Five relied on self-reported acetaminophen use, which could introduce recall bias (McCredie and Stewart, 1988; Rosenberg et al., 1998; Ross et al., 1989; Steineck et al., 1995; Walter et al., 2011a).
- Two studies did not analyze the effect of cumulative dose (**Friis** et al 2002, and **Rosenberg** et al 1999).
- Two studies did not analyze the effect of duration of acetaminophen use (**McCredie and Stewart** 1988, and **Pommer** et al 1999).

- Seven studies did not analyze the effect of latency between start of exposure and cancer diagnoses or onset of symptoms (**Walter** et al 2011, **McCredie and Stewart** 1988, **Ross** et al 1989, **Linet** et al 1995, **Steineck** et al 1995, **Rosenberg** et al 1999, and **Pommer** et al 1999).
- Five studies did not account for protopathic bias (**Walter** et al 2011, **McCredie and Stewart** 1988, **Ross** et al 1989, **Steineck** et al 1995, and **Pommer** et al 1999).
- Seven studies were not able to account for channeling bias (**Friis** et al 2002, **McCredie and Stewart** 1988, **Ross** et al 1989, **Linet** et al 1995, **Steineck** et al 1995, **Rosenberg** et al 1999, and **Pommer** et al 1999).
- Four studies did not account for confounding by indication (**McCredie and Stewart** 1988, **Linet** et al 1995, **Steineck** et al 1995, and **Pommer** et al 1999).

All studies were able to confirm cancer cases either through histopathologic results, medical records, or through cancer registries.

(ii) Renal Cancer

Cohort Studies

One of 4 cohort studies reported an association between regular acetaminophen use and increased risk for RRC (Karami et al., 2016). Karami et al 2016 assessed self-reported acetaminophen use in the US Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO) (N=98,807). Compared to irregular users, the aOR for RCC among regular users (≥1 times/week) was 1.68 (95% CI 1.19-2.39). An increased aOR was also observed among regular use with <10 years duration compared to irregular users (OR=2.09 95% CI 1.39-3.14). However, no association was observed in longer duration of use (≥10 years) compared to irregular users (OR=1.08 95% CI 0.55-2.1). This may be the result of remembering and reporting more relatively recent use than distant use, a type of exposure misclassification.

Three of 4 cohort studies conducted on renal cell carcinoma and acetaminophen use reported no positive associations (Cho et al., 2011; Friis et al., 2002; Walter et al., 2011a). In the study by Friis et al 2002, the SIR for renal parenchyma cancer which included RCC among those prescribed acetaminophen but not with aspirin or NSAIDs (N=13,482) was 1.0 (95% CI 0.4-2.1).

Walter et al 2011a (N=62,841) reported the aHR for renal cancer among high acetaminophen users (≥4 days/week and ≥4 years) to be 0.96 (95% CI 0.46-1.98) compared to non-users.

Cho et al 2011 assessed self-reported acetaminophen use in the Nurses' Health Study and Health Professionals Follow-up Study (N=126,928). Compared to irregular users, the aRR for RCC among regular users (≥2 times/week) was 1.32 (95% CI 0.96-1.84).

Case-Control Studies

Four of the 12 case-control studies conducted on RCC and acetaminophen use reported increased risk associated with ever/regular acetaminophen use or use within the highest level of exposure (if ever/regular use were not reported) (Derby and Jick, 1996; Gago-Dominguez et al., 1999; Kaye et al., 2001; McCredie et al., 1993). McCredie et al 1993 assessed self-reported acetaminophen use in a population-based case-control study in Australia (N=503 cases and 523 controls). Compared to irregular users, the aRR for RCC among regular users (≥20 times during lifetime) of acetaminophen in any form was 1.5 (95% CI 1.0-2.3). Also, the aRR among regular users of acetaminophen who never took phenacetin or aspirin was 1.6 (95% CI 1.0-2.8) compared to irregular users. Analysis by duration of use showed an aRR of 2.3 (95% CI 1.0-5.4) among those with >7 years of acetaminophen use who never took phenacetin or aspirin.

Derby and Jick 1996 used data from the Group Health Cooperative (GHC) of Puget Sound in a nested-case control study (N=222 cases and 885 controls). Acetaminophen use was determined using data from the GHC pharmacy which included OTC and prescription drug use, although acetaminophen obtained from a local pharmacy or grocery store would have been missed. Compared to non-users, the unadjusted RR for RCC among those with ≥1.0 kg lifetime consumption of acetaminophen was 2.6 (95% CI 1.1-6.0).

Gago-Dominguez et al 1999 assessed self-reported acetaminophen use in a population-based case-control study in USA (N=1276 cases and 1204 controls). Compared to irregular users of analgesics, the aOR for RCC among regular users (≥2 times/week for ≥1 month) of acetaminophen was 1.7 (95% CI 1.3-2.1). The aOR among exclusive users of acetaminophen was 1.6 (95% CI 1.1-2.4) compared to irregular users of analgesics. The study also reported an increased association with increasing maximum weekly dose of acetaminophen use. Note, however, that the study reported positive associations with aspirin, nonaspirin NSAIDs and phenacetin, suggesting that it was the indication, not the drug itself, that was the possible source of association with the cancer.

Kaye et al 2001 used data from the General Practice Research Database (GPRD) in a nested-case control study (N=20 cases and 434 controls). Acetaminophen use was determined using a prescription data base. Compared to non-users, the aOR for RCC among those with any acetaminophen use 1 to 5 years prior the index date was 1.6 (95% CI 1.0-2.6). Also, analysis by number of prescriptions showed an aOR of 2.3 (95% CI 1.0-5.3) among those with ≥20 prescriptions compared to non-users. OTC use of acetaminophen is missing and observed use is in sicker patients, who obtained the prescription for this OTC product while seeing their GP.

Eight of the 12 case-control studies conducted on RCC and acetaminophen use did not show increased risk associated with ever/regular acetaminophen use or use within the highest level of exposure (if ever/regular use were not reported) (Chow et al., 1994; Karami et al., 2016; Kreiger et al., 1993; McCredie et al., 1988; McCredie et al., 1995; McLaughlin et al., 1985; Mellemgaard

et al., 1994; Rosenberg et al., 1998). McLaughlin et al 1985 assessed self-reported acetaminophen use in a population-based case-control study in the US (N=495 cases and 697 controls). Compared to never users, the aOR for RCC among female ever users of acetaminophen was 1.2 (95% CI 0.8-1.9). The aOR among male ever users was 0.7 (95% CI 0.5-1.0).

McCredie et al 1988 assessed self-reported acetaminophen use in a population-based case-control study in Australia (N=360 cases and 985 controls). Compared to non-users, the aRR for RCC among regular users (≥0.1 kg consumption) of acetaminophen was 1.2 (95% CI 0.8-1.8).

The study by Kreiger et al 1993 assessed self-reported acetaminophen use in a population-based case-control study in Canada (N=518 cases and 1381 controls). Compared to never or irregular users, the aOR for RCC among female ever users of acetaminophen was 0.6 (95% CI 0.4-1.6). The aOR among male ever users was 0.9 (95% CI 0.4-1.8).

Chow et al 1994 assessed acetaminophen use through self-reports and pharmaceutical records in a population-based case-control study in the US (N=690 cases and 707 controls). Compared to never users, the aOR for RCC among female regular users (≥2 times/week for ≥1 month) of acetaminophen was 2.1 (95% CI 0.6-6.9). The aOR among male regular users was 1.2 (95% CI 0.5-3.2).

Mellemgaard et al 1994 assessed self-reported acetaminophen use in a population-based case-control study in Denmark (N=368 cases and 396 controls). Compared to never users, the aOR for RCC among female ever users (≥2 times/week for ≥1 month) of acetaminophen was 1.0 (95% CI 0.4-2.5). The aOR among male ever users was 1.1 (95% CI 0.5-3.0).

McCredie et al 1995 conducted a pooled analysis of the studies by McCredie et al 1993, Chow et al 1994, and Mellemgaard et al 1994. Results showed that compared to never or irregular users, the aRR for RCC among regular users (≥0.1 kg consumption) of acetaminophen was 1.1 (95% CI 0.9-1.5).

Rosenberg et al 1998 assessed self-reported acetaminophen use in a hospital-based case-control study in the US (N=383 cases and 8,149 controls). Compared to never users of acetaminophen, the aRR for RCC among those with regular use for a duration of ≥5 years (≥2 days/week for ≥1 month) and begun at least 1 year prior to hospitalization was 1.1 (95% CI 0.5-2.6).

In the case-control study by Karami et al 2016, the authors assessed self-reported acetaminophen use in the US Kidney Cancer Study (N=1,217 cases and 1,235 controls). Compared to irregular users, the aOR for RCC among regular users (≥1 times/week for ≥3 months) was 1.09 (95% CI 0.87-1.37). However, a significant association was observed among regular use with ≥10 years duration compared to irregular users (OR=1.54 95% CI 1.09-2.16). Also, a significant association was observed in regular over-the-counter users (OR=1.35 95% CI 1.01-1.83) but not in prescription users (OR=0.96 95% CI 0.74-1.24). The discrepant findings between OTC and

prescription use does not support an etiologic hypothesis and no explanation was provided by the authors.

Five case-control studies also assessed the association between acetaminophen use and renal pelvis cancer. All 5 studies did not show significantly increased risk associated with ever/regular acetaminophen use or use within the highest level of exposure (if ever/regular use were not reported) (Kaye et al., 2001; McCredie and Stewart, 1988; McCredie et al., 1993; McLaughlin et al., 1985; Pommer et al., 1999). McLaughlin et al 1985 reported the aOR for renal pelvis cancer among female ever users of acetaminophen to be 2.2 (95% CI 0.8-5.8) compared to never users. The aOR among male ever users was 1.2 (95% CI 0.6-2.5). McCredie and Stewart 1988 reported the aOR for renal pelvis cancer among those with ≥ 1.0 kg lifetime consumption of acetaminophen to be 0.8 (95% CI 0.4-1.7). McCredie et al 1993 reported the aRR for renal pelvis among regular users (≥20 times during lifetime) of acetaminophen in any form to be 1.3 (95% CI 0.7-2.4) compared to irregular users. Pommer et al 1999 reported the aOR for renal pelvis cancer among those with ≥ 1 kg cumulative lifetime use to be 3.27 (95% CI 0.25-43.02) compared to non or rare users. Kaye et al 2001 reported the unadjusted OR for renal pelvis cancer among those with any acetaminophen use 1 to 5 years prior the index date to be 1.2 (95% CI 0.4-3.1).

(iii) Bladder Cancer

Cohort Studies

All 3 cohort studies conducted on bladder cancer and acetaminophen use reported no associations (Friis et al., 2002; Genkinger et al., 2007; Walter et al., 2011a). Friis et al 2002, the SIR for bladder cancer among those prescribed with acetaminophen, but not with aspirin or NSAIDs (N=13,482) was 1.0 (95% CI 0.7-1.5).

The study by Genkinger et al 2007 assessed self-reported acetaminophen use in the Health Professionals Follow-Up Study (N=49,448). The reported aRR for bladder cancer among regular (≥1 time/week for ≥3 months for ≥2 years prior to interview) users in 1986 and 1988 was 0.9 (95% CI 0.49-1.65) compared to non-users.

Walter et al 2011a (N=62,841) reported the aHR for bladder cancer among high acetaminophen users (≥4 days/week and ≥4 years) to be 1.5 (95% CI 0.57-3.89) compared to non-users.

Case Control Studies

Eight of the 9 case-control studies on bladder cancer and acetaminophen use reported no significant associations (Castelao et al., 2000; Derby and Jick, 1996; Fortuny et al., 2006; Fortuny et al., 2007; McCredie et al., 1988; Piper et al., 1985; Pommer et al., 1999). Piper et al 1985 assessed self-reported acetaminophen use in a population-based case-control study in the US (N=173 cases and 173 controls). Compared to irregular users, the unadjusted OR for bladder cancer among regular users (≥30 days/year) of acetaminophen only was 1.5 (95% CI 0.4-7.2).

McCredie and Stewart 1988 (N=162 cases and 381 controls) reported the aOR for bladder cancer among those with ≥ 1.0 kg lifetime consumption of acetaminophen to be 0.7 (95% CI 0.4-1.3) compared to those unexposed to acetaminophen.

Derby and Jick 1996 (N=504 cases and 885 controls) reported the unadjusted RR for bladder cancer among those with ≥1.0 kg lifetime consumption of acetaminophen to be 1.3 (95% CI 0.6-2.8) compared to non-users.

Pommer et al 1999 (N=571 cases and 647 controls) reported the aOR for bladder cancer among those with ≥ 1 kg cumulative lifetime use to be 0.83 (95% CI 0.33-2.07) compared to non or rare users.

Castelao et al 2000 assessed self-reported acetaminophen use in a population-based case-control study in the US (N=1,514 cases and 1,514 controls). Compared to non or irregular users of acetaminophen, the aOR for bladder cancer among those with regular use (≥2 times/week for ≥1 month) was 0.85 (95% CI 0.6-1.19).

The study by Kaye et al 2001 (N=189 cases and 744 controls) reported the aOR for bladder cancer among those with any acetaminophen use 1 to 5 years prior the index date to be 0.9 (95% CI 0.6-1.3).

Fortuny et al 2006 assessed self-reported acetaminophen use in a hospital-based case-control study in Spain (N=958 cases and 1,029 controls). Compared to non-users of acetaminophen, the aOR for bladder cancer among ever users was 0.8 (95% CI 0.6-1.0).

Fortuny et al 2007 assessed self-reported acetaminophen use in a population-based case-control study in the US (N=376 cases and 463 controls). Compared to non-users of acetaminophen, the aOR for bladder cancer among ever users was 0.8 (95% CI 0.5-1.6).

One of the 9 case-control studies reported a positive association between acetaminophen use and bladder cancer. Baris et al 2013 assessed self-reported acetaminophen use in a population-based case-control study in the US (N=1,171 cases and 1,418 controls) (Baris et al., 2013). Compared to never users of acetaminophen, the aOR for bladder cancer among those with regular use (≥2 times/week for ≥1 month) was 1.3 (95% CI 1.1-1.7). An association was also observed among those with regular use for <5 years but none was observed for higher categories of duration of use.

8.1.2 Lymphohematopoietic Neoplasms

(i) Lymphoma, Non-Hodgkin Lymphoma (NHL), NOS and its subtypes

One case-control study assessed the association between acetaminophen use and lymphoma (not otherwise specified). Becker et al 2009 assessed self-reported acetaminophen use in a hospital-based case-control study in Europe (N=2,362 cases and 2,458 controls) (Becker et al., 2009). Compared to non-users of acetaminophen, the aOR for lymphoma among any users was

2.29 (95% CI 1.49-3.51). However, this study relied on self-reported exposure and was subject to recall bias, it did assess latency or dose, and was a hospital-based case-control study, so at risk of selection bias.

One of 2 cohort studies reported significant association between acetaminophen use and lymphoma but no increase in RR for SLL or CLL. Walter et al 2011b reported the aHR for NHL among high acetaminophen users (≥4 days/week and ≥4 years) to be 1.81 (95% CI 1.12-2.93) compared to non-users. One the other hand, Walter also assessed the RR for SLL/CLL which was not increased. The aHR for SLL/CLL was 0.84 (0.31-2.28). In the study by Friis et al 2002, the association between acetaminophen use and NHL was not significant. The SIR among those prescribed with acetaminophen but not with aspirin or NSAIDs (N=13,482) was 1.2 (95% CI 0.7-2.0).

One of 2 case-control studies reported a significant association between acetaminophen use and lymphoma. Baker et al 2005 assessed self-reported acetaminophen use in a hospital-based case-control study in the US (N=625 cases and 2,512 controls) (Baker et al., 2005). Compared to irregular users of acetaminophen, the aOR for NHL among female regular (≥once/week for 6 months) users was 1.71 (95% CI 1.18-2.5). Also, the aOR for SLL among female regular users was 2.41 (95% CI 1.08–5.41). However, it is important to note that Baker found no significant association among males and all analyses by duration, frequency of use, and cumulative acetaminophen use showed non-significant results. Inconsistency between sexes suggests it is not biologic. And note issues with autoimmune diseases and lymphoma especially mentioned above.

In the study by Kato et al 2002, the association between acetaminophen use and NHL was not significant (Kato et al., 2002). Self-reported acetaminophen use was assessed in a population-based case-control study in the US (N=376 cases and 463 controls). Compared to non-users of acetaminophen, the aOR for NHL among regular (≥ 1 time/month for ≥ 6 months) users for a duration of >10 years was 1.39 (95% CI 0.45-4.26).

(ii) Hodgkin Lymphoma (HL)

One cohort study reported no significant association between acetaminophen use and Hodgkin Lymphoma. In the study of Friis et al 2002, the SIR for HL among those prescribed with acetaminophen but not with aspirin or NSAIDs (N=13,482) was 1.4 (95% CI 0.5-8.0).

One case-control study reported significant association between acetaminophen use and Hodgkin Lymphoma. Chang et al 2004 assessed self-reported acetaminophen use in a population-based case-control study in the US (N=565 cases and 679 controls) (Chang et al., 2004). Compared to irregular users of acetaminophen, the aOR for HL among regular (≥2 times/week) users was 1.71 (95% CI 1.29-2.31). When compared to never users, the aOR among regular users was 2.17 (95% CI 1.58-2.98).

(iii) Multiple Myeloma (MM)

One cohort study reported no significant association between acetaminophen use and Multiple Myeloma. In the study of Friis et al 2002, the SIR for MM among those prescribed with acetaminophen but not with aspirin or NSAIDs (N=13,482) was 1.6 (95% CI 0.6-3.2).

One case-control study reported significant association between acetaminophen use and MM. Moysich et al 2007 assessed self-reported acetaminophen use in a hospital-based case-control study in the US (N=117 cases and 483 controls). Compared to irregular users of acetaminophen, the aOR for MM among regular (≥1/week for ≥6 months) users was 2.95 (95% CI 1.72-5.08). Significant findings were also observed among those who used acetaminophen >7 times/week (aOR=4.36 95%CI 1.7-11.2) and with >10 years duration of use (aOR=3.26 95%CI 1.52-7.02).

(iv) Leukemia (adult)

One of 2 cohort studies reported significant association between acetaminophen use and lymphoma. Walter et al 2011b reported an aHR for myeloid leukemia among high acetaminophen users (≥4 days/week and ≥4 years) to be 2.26 (95% CI 1.24-4.12) compared to non-users. In the study by Friis et al 2002, the association between acetaminophen use and leukemia was not significant. The SIR among those prescribed with acetaminophen but not with aspirin or NSAIDs (N=13,482) was 0.9 (95% CI 0.5-1.6).

Two of the 3 case-control studies that determined the association between acetaminophen use and leukemia reported significant results (Ross et al., 2011; Weiss et al., 2006), however, there was no significant increase in either of these studies in the sub-types evaluated. Weiss et al 2006 assessed self-reported acetaminophen use in a hospital-based case-control study in the US (N=169 cases and 676 controls). Compared to never users of acetaminophen, the aOR for leukemia among ever users was 1.53 (95% CI 1.03-2.26). However, analyses by sub-type showed no significant findings for acute lymphoblastic/lymphocytic leukemia (ALL, aOR=1.73, 95% CI 0.79-3.78) and acute myeloid leukemia (AML, aOR=1.5, 95% CI 0.98-2.3).

Ross et al 2011 assessed self-reported acetaminophen use in a population-based case-control study in the US (N=670 cases and 701 controls). Compared to never users of acetaminophen, the aOR for myeloid leukemia among female ever users was 1.60 (95% CI 1.04-2.47). However, analyses by sub-type showed no significant findings for AML (aOR=1.46, 95% CI 0.87-2.44) and chronic myeloid leukemia (CML, aOR=1.24, 95% CI 0.64-2.42). No significant associations were seen for myeloid leukemia and subtypes among males.

One of the 3 case-control studies that evaluated the association between acetaminophen use and leukemia reported non-significant results. Friedman et al 1982 conducted a case-control study in the US that assessed self-reported acetaminophen use among leukemia cases, hospital controls, and members of the Kaiser-Permanente Medical Care Program (N=409 cases and 818 controls). Compared to non-users of acetaminophen, the aOR for leukemia among any users was

0.44 (95% CI 0.2-1.0) using hospital controls and 1.13 (95% CI 0.43-2.91) using member controls. Also, the aOR for myeloid leukemia among any users was 0.67 (95% CI 0.19-2.34) using hospital controls and 1.67 (95% CI 0.4-6.87) using member controls.

(v) Leukemia (childhood)

Both of the case-control studies that assessed the association between acetaminophen use and childhood leukemia reported no significant results. Ognjanovic et al 2015 assessed self-reported acetaminophen use among mothers in a population-based case-control study in the US and Canada (N=441 cases and 323 controls) (Ognjanovic et al., 2011). Compared to irregular users of acetaminophen, the aOR for ALL in children of mothers with regular (≥5 times) acetaminophen use prior to knowing pregnancy was 1.16 (95%CI 0.80-1.68). The aOR for AML was 0.66 (95% CI 0.43-1.01). All other analyses by timing of acetaminophen use showed non-significant results.

Couto et al 2015 assessed self-reported acetaminophen use among mothers in a hospital-based case-control study in Brazil (N=231 cases and 411 controls) (Couto et al., 2015). Compared to no report of acetaminophen use, the aOR for ALL in children of mothers with reported acetaminophen use was 0.56 (95%CI 0.28-1.10). The aOR for AML was 0.48 (95% CI 0.15-1.48). All other analyses by age of children showed non-significant results.

8.1.3 Liver Cancer

Cohort Studies

One of 2 cohort studies reported significant association between acetaminophen use and liver cancer. In Lipworth et al 2003 (N=49,890), the SMR for liver cancer among those prescribed with acetaminophen was 2.2 (95% CI 1.6-2.9) (Lipworth et al., 2003). Significantly increased SMRs were also observed among those prescribed with latencies of <1 year (SMR=3.8, 95%CI 2.3-5.9) and ≥5 years (SMR=2.6, 95%CI 1.1-5.2). While significantly increased SMRs were observed among those given with 1 (SMR=2.7, 95%CI 1.6-4.1) and 2 to 4 prescriptions (SMR=2.1, 95% CI 1.1-3.6), no significant increase was observed in higher number of prescriptions. This inverted latency and dose response suggests the association was not biological. Due to the limitations of Lipworth et al study design (i.e., mortality study, no confounders other than age and sex used), we do not consider this valid evidence for an association.

In the study by Friis et al 2002, the association between acetaminophen use and liver cancer was not significant. The SIR among those prescribed with acetaminophen but not with aspirin or NSAIDs (N=13,482) was 1.8 (95% CI 0.7-3.6).

Case Control Studies

Two case-control studies reported significant association between acetaminophen use and liver cancer (Yang et al., 2016).

McGlynn et al 2015 used data from the UK CPRD in a nested-case control study (N=1,195 cases and 4,640 controls). They determined acetaminophen use through prescription records. The unadjusted OR for liver cancer among those who ever used acetaminophen was 1.52 (95% CI 1.31-1.75). No adjusted analysis was conducted.

Yang et al 2016 used similar data as that of McGlynn et al 2015 and noted very low increases. Compared to those with <2 prescriptions of acetaminophen, the aOR for liver cancer among ever users (≥2 prescriptions) was 1.18 (95% CI 1.00-1.39). A significant association was also observed among ever users (aOR=1.2, 95%CI 1.02-1.42) after excluding exposure 2 years prior to the case diagnosis. Also, significant associations were observed after excluding those with liver disease among those with ≥2 (aOR=1.24, 95%CI 1.05-1.47) and ≥40 prescriptions (aOR=1.61, 95%CI 1.22-2.12).

8.2 Early Studies Including Assessments of Phenacetin Without Explicitly Accounting for Phenacetin as a Source of Confounding

Baris D, Karagas MR, Koutros S, et al.. Nonsteroidal anti-inflammatory drugs and other analgesic use and bladder cancer in northern New England. Int J Cancer. 2013;132:162-173.

Castelao JE, Yuan JM, Gago-Dominguez M, et al. Non-steroidal anti-inflammatory drugs and bladder cancer prevention. Brit J Cancer. 2000;82(7):1364-1369.

Cho E, Curhan G, Hankinson SE, et al. Prospective evaluation of analgesic use and risk of renal cell cancer. Arch Intern Med. 2011;171(16):1487-93.

Chow WH, McLaughlin JK, Linet MS, et al. Use of analgesics and risk of renal cell cancer. Int J Cancer. 1994;59(4):467-70.

Cramer DW, Harlow BL, Titus-Ernstoff L, et al. Over-the-counter analgesics and risk of ovarian cancer. Lancet. 1998;351(9096):104-7.

Derby LE, Jick H. Acetaminophen and renal and bladder cancer. Epidemiology. 1996:358-62.

Fortuny J, Kogevinas M, Garcia-Closas M, et al. Use of analgesics and nonsteroidal anti-inflammatory drugs, genetic predisposition, and bladder cancer risk in Spain. Cancer Epidemiol Biomarkers Prev. 2006;15(9):1696-1702.

Fortuny J, Kogevinas M, Zens MS, et al. Analgesic and anti-inflammatory drug use and risk of bladder cancer: a population based case control study. BMC Urology. 2007;7(1):13.

Gago-Dominguez M, Yuan J-M, Castelao JE, et al. Regular use of analgesics is a risk factor for renal cell carcinoma. Brit J Cancer. 1999;81(3):542-548.

Kreiger N, Marrett LD, Dodds L, et al. Risk factors for renal cell carcinoma: results of a population-based case-control study. Cancer Causes Control. 1993;4(2):101-10.

Linet MS, Chow WH, McLaughlin JK, et al. Analgesics and cancers of the renal pelvis and ureter. Int J Cancer. 1995;62(1):15-8.

McCredie M, Ford JM, Stewart JH. Risk factors for cancer of the renal parenchyma. Int J Cancer. 1988;42(1):13-6.

McCredie M, Pommer W, McLaughlin JK, et al. International renal-cell cancer study. II. Analgesics. Int J Cancer. 1995;60(3):345-9.

McCredie M, Stewart JH, Day NE. Different roles for phenacetin and paracetamol in cancer of the kidney and renal pelvis. Int J Cancer. 1993 Jan 21;53(2):245-9.

McCredie M, Stewart JH. Does paracetamol cause urothelial cancer or renal papillary necrosis?. Nephron. 1988;49(4):296-300.

McLaughlin JK, Blot WJ, Mehl ES, Fraumeni JF, J. Relation of analgesic use to renal cancer: Population-based findings. Natl Cancer Inst Monogr. 1985;69:217-22.

Mellemgaard A, Niwa S, Mehl ES, et al. Risk factors for renal cell carcinoma in Denmark: role of medication and medical history. Int J Epidemiol. 1994;23(5):923-30.

Piper JM, Tonascia J, Matanoski GM. Heavy phenacetin use and bladder cancer in women aged 20 to 49 years. N Engl J Med. 1985 ;313(5):292-5.

Pommer W, Molzahn M. Urothelial cancer at different tumour sites: role of smoking and habitual intake of analgesics and laxatives. Results of the Berlin Urothelial Cancer Study. Nephrol Dial Transplant. 1999;14(12):2892&2897.

Rosenberg L, Rao RS, Palmer JR, et al. Transitional cell cancer of the urinary tract and renal cell cancer in relation to acetaminophen use (United States). Cancer Causes Control. 1998;9(1):83-8.

Ross RK, Paganini-Hill A, Landolph J, et al. Analgesics, cigarette smoking, and other risk factors for cancer of the renal pelvis and ureter. Cancer Res. 1989;49(4):1045-8.

Steineck G, Wiholm BE, De Verdier MG. Acetaminophen, some other drugs, some diseases and the risk of transitional cell carcinoma: A population-based case-control study. Acta Oncol. 1995;34(6):741-8.

8.3 Forest Plots by Cancer Type

The forest plots include one-point estimate and confidence interval for each cancer type within each study. The estimate in most cases is the RR of any acetaminophen use versus no acetaminophen use or nonuse of acetaminophen. For those studies that did not provide a RR forever versus never use, then either regular use or the highest exposure category use was used.

Figure 21. Forest plot: urinary tract cancers

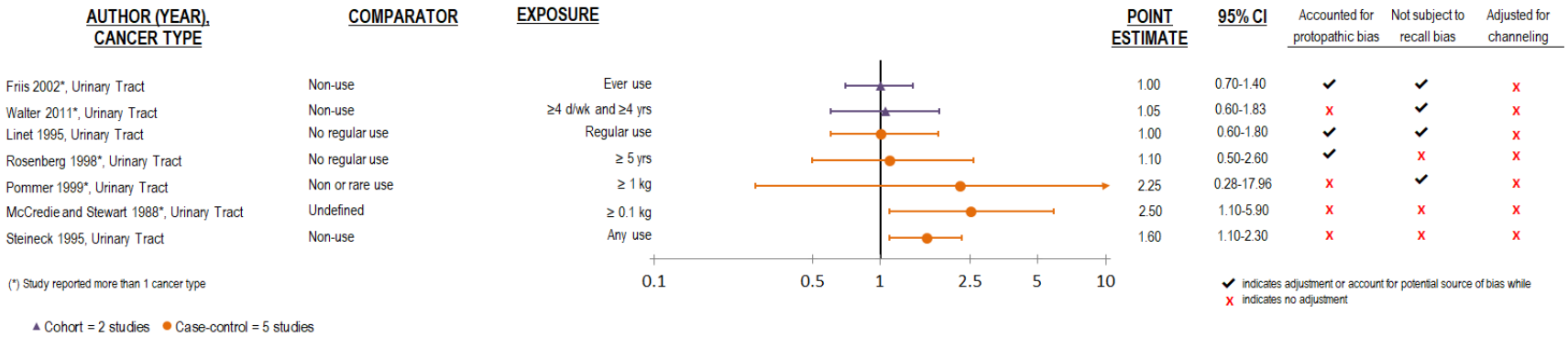


Figure 22. Forest plot: renal cancer

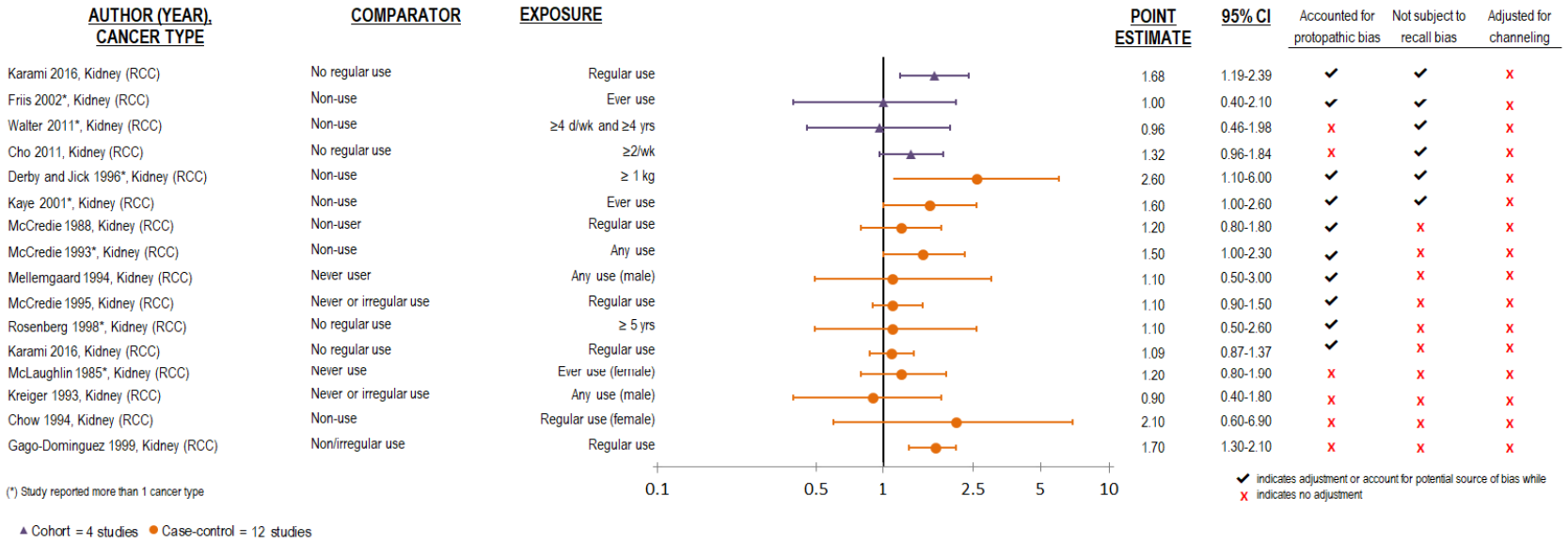


Figure 23. Forest plot: renal pelvis cancer

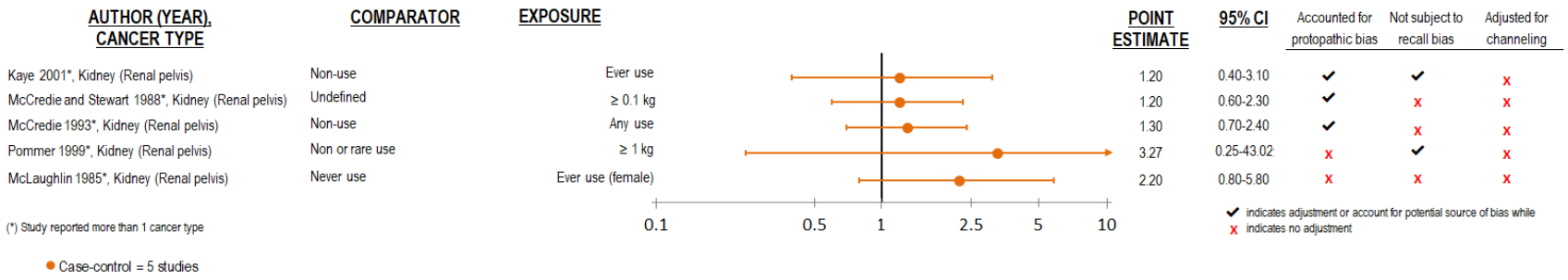


Figure 24. Forest plot: bladder cancer

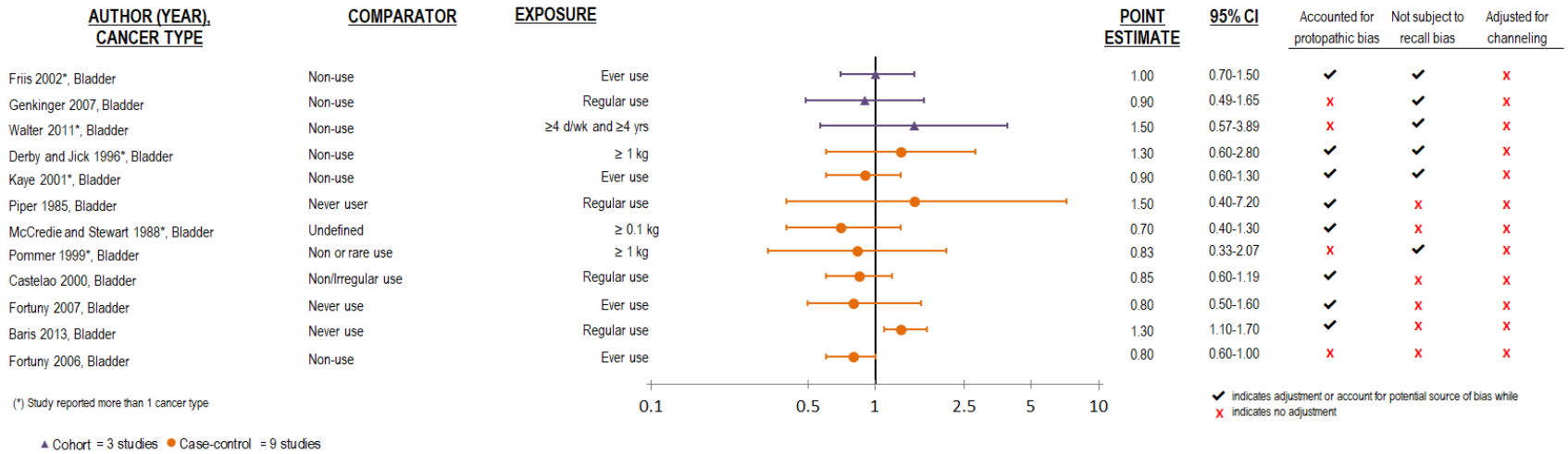
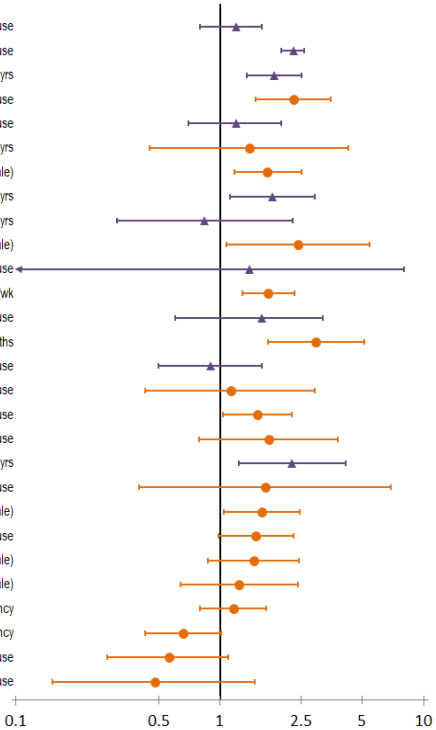


Figure 25. Forest plot: lymphohematopoietic neoplasms

AUTHOR (YEAR), CANCER TYPE	COMPARATOR	EXPOSURE	POINT ESTIMATE	95% CI	No recall bias	Analyzed cumulative dose	Duration analyzed	Latency analyzed	Account for protopathic bias	Account for channeling	Outcome confirmed
Friis 2002*, Hematologic cancer	Non-use	Ever use	1.20 [a]	0.80-1.60	✓		✓	✓	✓	✓	✓
Lipworth 2003*, Hematologic cancer	Non-use	Any use	2.30 [a]	2.00-2.60	✓			✓			
Walter 2011b*, Hematologic cancer	Non-use	≥4 days/wk and ≥4 yrs	1.84	1.35-2.50	✓				✓	✓	✓
Becker 2009, Lymphoma	Non-use	Any use	2.29 [a]	1.49-3.51							✓
Friis 2002*, Lymphoma (NHL)	Non-use	Ever use	1.20 [a]	0.70-2.00	✓		✓	✓	✓	✓	✓
Kato 2002, Lymphoma (NHL)	Never use	>10 yrs	1.39	0.45-4.26			✓		✓	✓	✓
Baker 2005*, Lymphoma (NHL)	No regular use	Regular use (female)	1.71	1.18-2.50		✓	✓				✓
Walter 2011b*, Lymphoma (NHL)	Non-use	≥4 days/wk and ≥4 yrs	1.81	1.12-2.93	✓				✓	✓	✓
Walter 2011b*, Lymphoma (NHL, SLL/CLL)	Non-use	≥4 days/wk and ≥4 yrs	0.84	0.31-2.28	✓				✓	✓	✓
Baker 2005*, Lymphoma (NHL, SLL)	No regular use	Regular use (female)	2.41	1.08-5.41		✓	✓				✓
Friis 2002*, Hodgkin's Lymphoma	Non-use	Ever use	1.40 [a]	0.00-8.00	✓		✓	✓	✓	✓	✓
Chang 2004, Hodgkin's Lymphoma	No regular use	≥2 times/wk	1.72	1.29-2.31							✓
Friis 2002*, Multiple Myeloma	Non-use	Ever use	1.60	0.60-3.20	✓		✓	✓	✓	✓	✓
Moysich 2007, Multiple myeloma	No regular use	≥1 tab/week for ≥ 6 months	2.95	1.72-5.08			✓			✓	✓
Friis 2002*, Leukemia (combined)	Non-use	Ever use	0.90 [a]	0.50-1.60	✓		✓	✓	✓	✓	✓
Friedman 1982*, Leukemia (combined)	Non-use	Any use	1.13 [a]	0.43-2.91							✓
Weiss 2006*, Leukemia (combined)	Never	Ever use	1.53 [a]	1.03-2.26		✓	✓		✓		✓
Weiss 2006*, Leukemia (ALL)	Never	Ever use	1.73 [a]	0.79-3.78		✓	✓		✓		✓
Walter 2011b*, Leukemia (myeloid)	Non-use	≥4 days/wk and ≥4 yrs	2.26	1.24-4.12	✓				✓	✓	✓
Friedman 1982*, Leukemia (myeloid)	Non-use	Any use	1.67 [a]	0.40-6.87							✓
Ross 2011*, Leukemia (myeloid)	Non-use	Any use (female)	1.60	1.04-2.47			✓		✓	✓	✓
Weiss 2006*, Leukemia (AML)	Never	Ever use	1.50 [a]	0.98-2.30		✓	✓		✓		✓
Ross 2011*, Leukemia (AML)	Non-use	Any use (female)	1.46	0.87-2.44			✓		✓	✓	✓
Ross 2011*, Leukemia (CML)	Non-use	Any use (female)	1.24	0.64-2.42			✓		✓	✓	✓
Ognjanovic 2011*, Pediatric leukemia (ALL)	Non-use	Any use, prior pregnancy	1.16	0.80-1.68							✓
Ognjanovic 2011*, Pediatric leukemia (AML)	Non-use	Any use, prior pregnancy	0.66	0.43-1.01							✓
Couto 2015*, Pediatric leukemia (ALL)	Non-use	Any use	0.56	0.28-1.10							✓
Couto 2015*, Pediatric leukemia (AML)	Non-use	Any use	0.48	0.15-1.48							✓



(*) Study reported more than 1 cancer type

▲ Cohort = 3 studies ● Case-control = 10 studies

a- no control for confounders other than age and/or sex

Figure 26. Forest plot: liver cancer

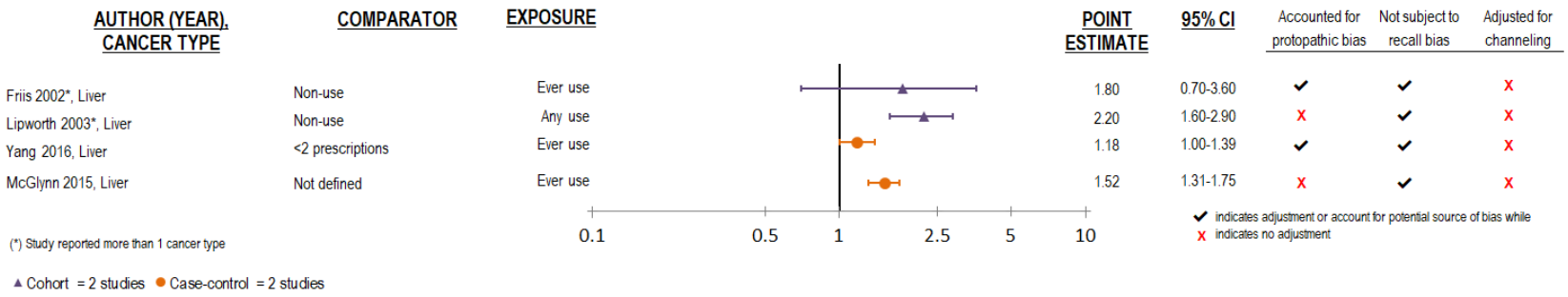
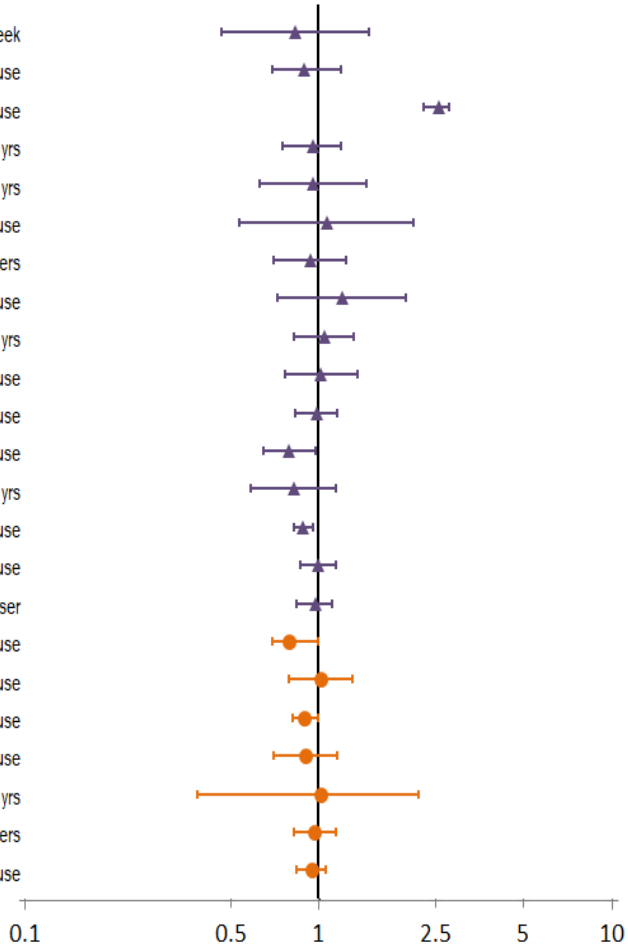


Figure 27. Forest plot: breast cancer

AUTHOR (YEAR), CANCER TYPE	COMPARATOR	EXPOSURE	POINT ESTIMATE	95% CI
Harris 1999, Breast	Not defined	≥ 4 pills/week	0.84	0.47-1.50
Friis 2002*, Breast	Non-use	Ever use	0.90 [a]	0.70-1.20
Lipworth 2003*, Breast	Non-use	Any use	2.60 [a]	2.30-2.80
Harris 2003, Breast	<5 yrs use	≥5 yrs	0.96 [a]	0.76-1.20
Marshall 2005, Breast	Non-use	daily use for ≥5 yrs	0.96	0.63-1.47
Gallicchio 2006, Breast	Non-use	Regular use	1.07	0.54-2.12
Gallicchio 2007, Breast	Non-use	Users	0.94 [a]	0.71-1.25
Kwan 2007, Breast	Non-use	Ever use	1.21	0.73-2.00
Gill 2007, Breast	Non-use	≥6 yrs	1.05	0.83-1.33
Friis 2008, Breast	Non-use	Ever use	1.02	0.77-1.36
Eliassen 2009, Breast	Non-use	Current regular use	0.99	0.84-1.16
Bosco 2011, Breast	Non-use	Current use	0.80	0.65-0.98
Walter 2011*, Breast	Non-use	≥4 d/wk and ≥4 yrs	0.83	0.59-1.15
Zhang 2012, Breast	Non-use	Current use	0.89	0.83-0.96
Clarke 2017, Breast	Non-use	Current use	1.00	0.87-1.15
Kehm 2019, Breast	Non regular user	Regular user	0.98	0.85-1.12
Meier 2002*, Breast	Non-use	Regular use	0.80	0.70-1.00
Terry 2004, Breast	Non-use	Ever use	1.02	0.80-1.31
Garcia Rodriguez 2004a, Breast	Non-use	Current use	0.90	0.82-1.00
Rahme 2005, Breast	No regular use	≥90 days of use	0.91	0.71-1.16
Harris 2006, Breast	Non-use	≥2 x/wk for ≥2 yrs	1.02	0.39-2.20
Brasky 2010, Breast	Non-user	Users	0.97	0.83-1.15
Ashok 2011, Breast	Non-use	Any use	0.95 [a]	0.85-1.06



(*) Study reported more than 1 cancer type

▲ Cohort = 16 studies ● Case-control = 7 studies

a- no control for confounders other than age and/or sex

Figure 28. Forest plot: ovarian cancer

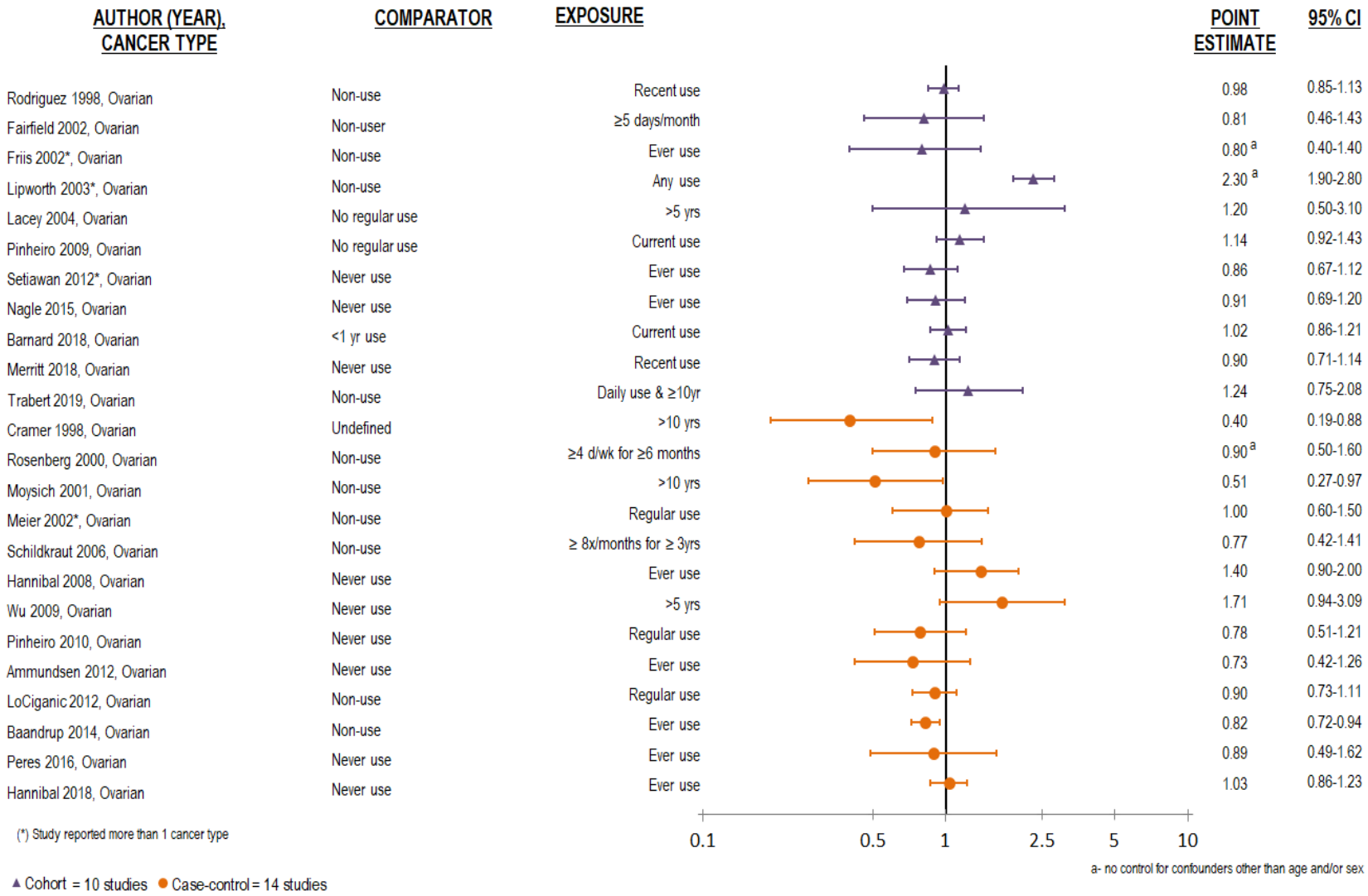


Figure 29. Forest plot: cervical cancer

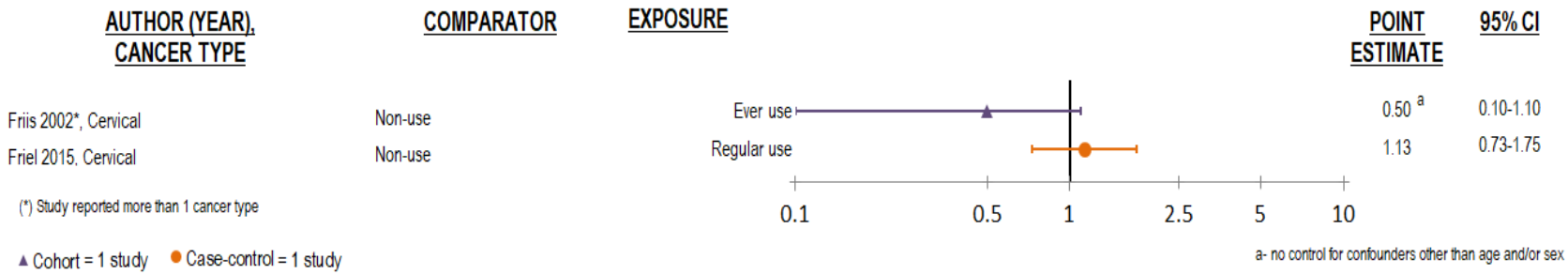


Figure 30. Forest plot: uterine/endometrial cancer

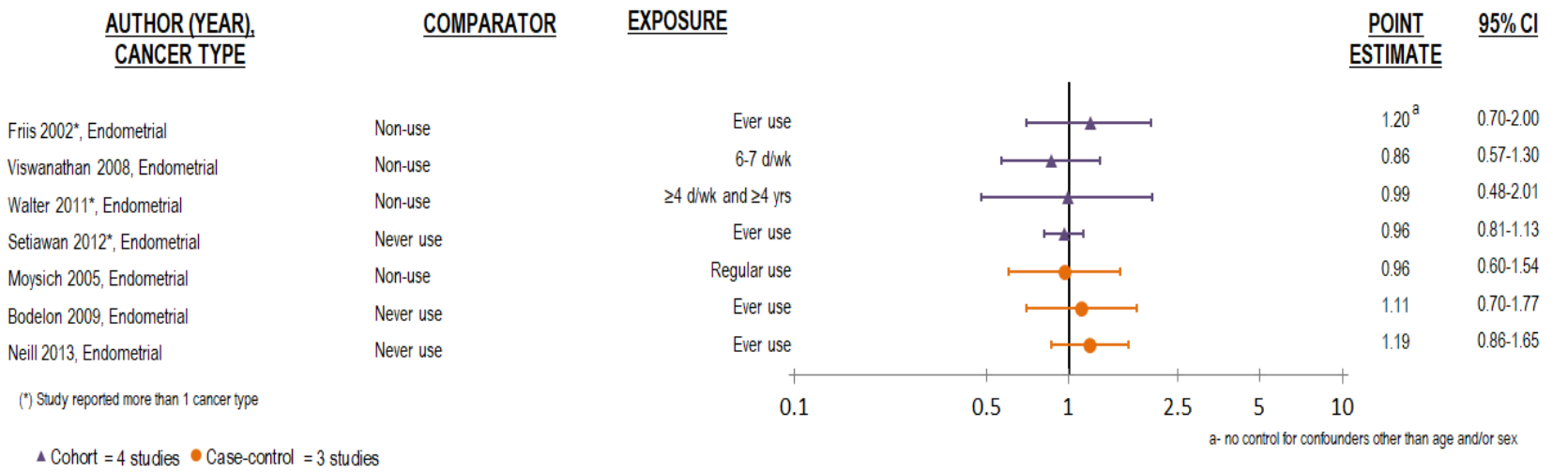


Figure 31. Forest plot: prostate cancer

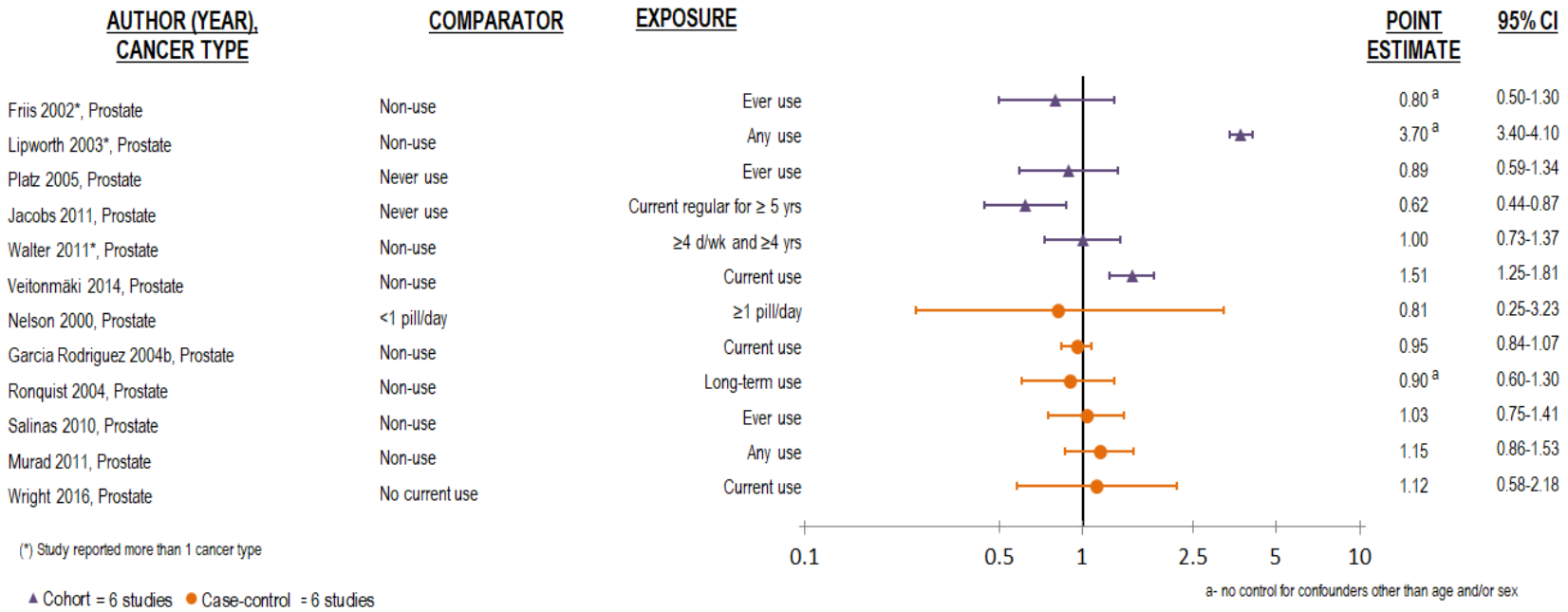


Figure 32. Forest plot: skin cancer

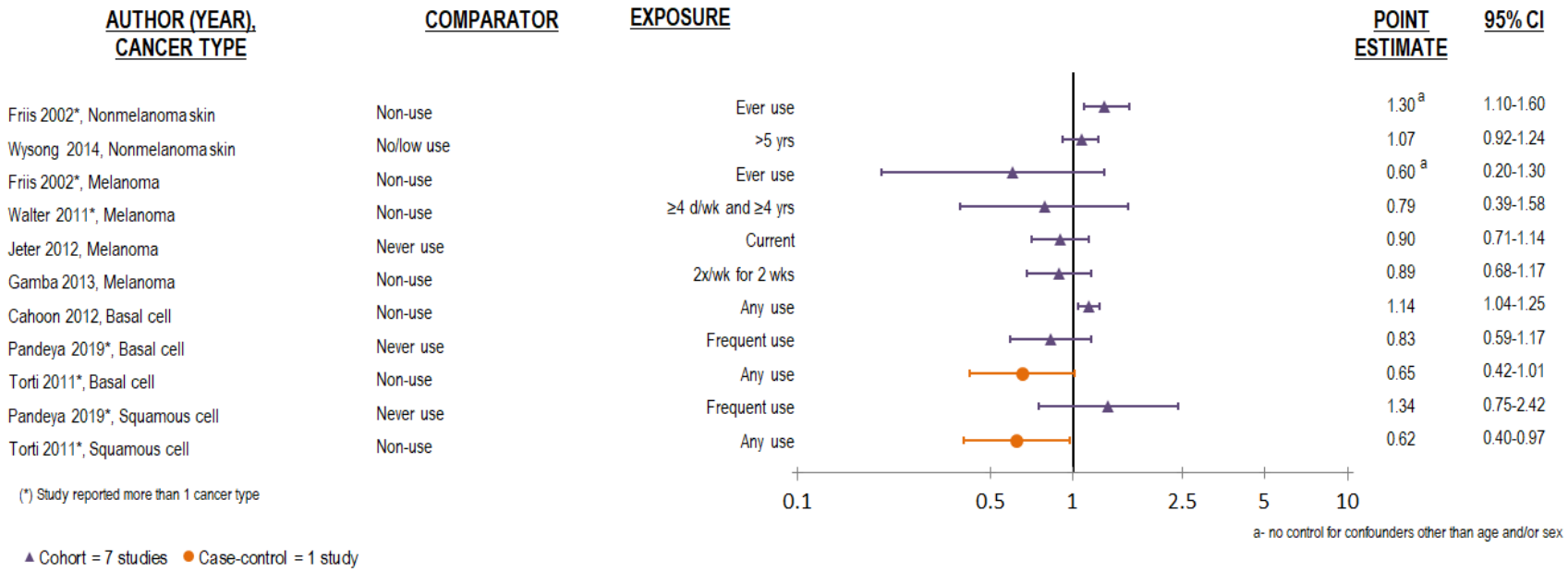


Figure 33. Forest plot: colorectal cancer

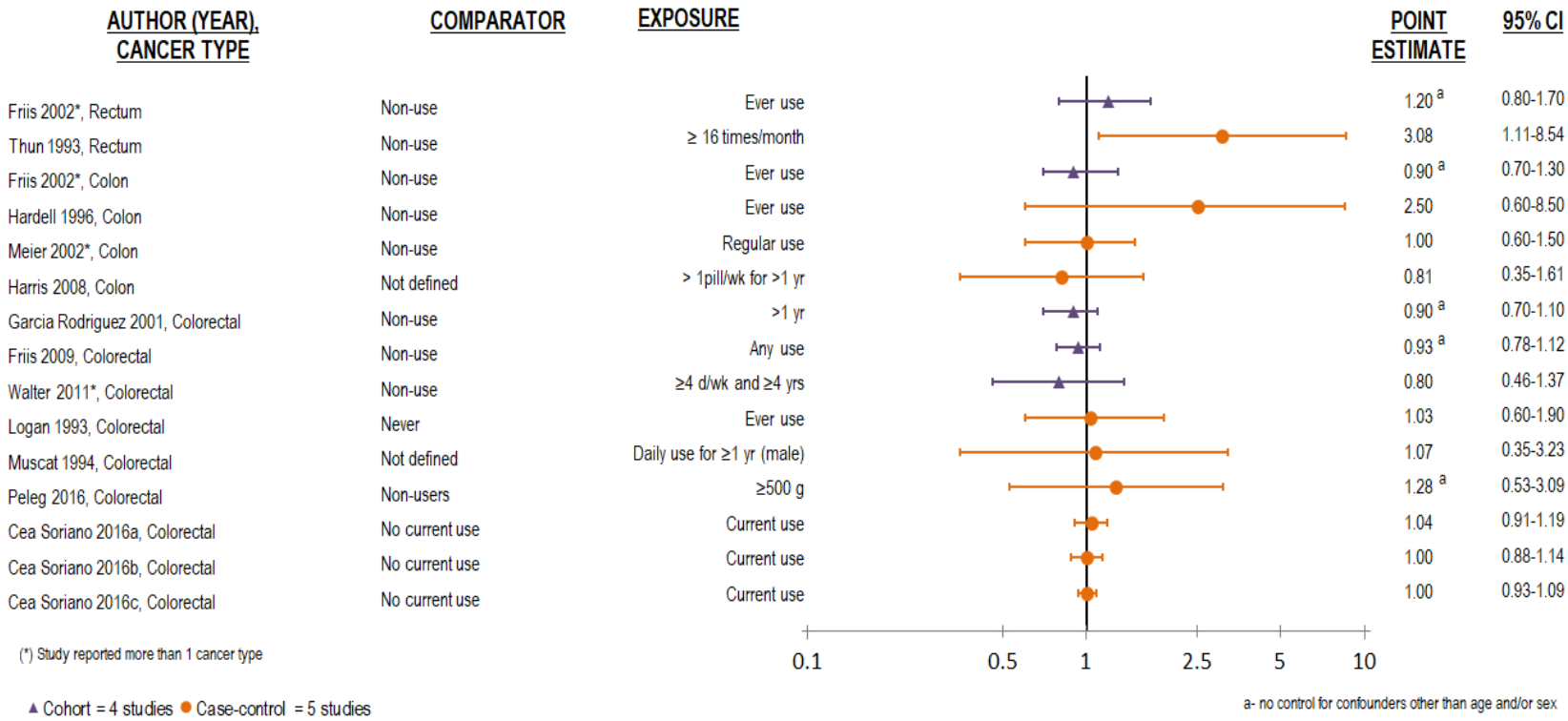


Figure 34. Forest plot: brain cancer

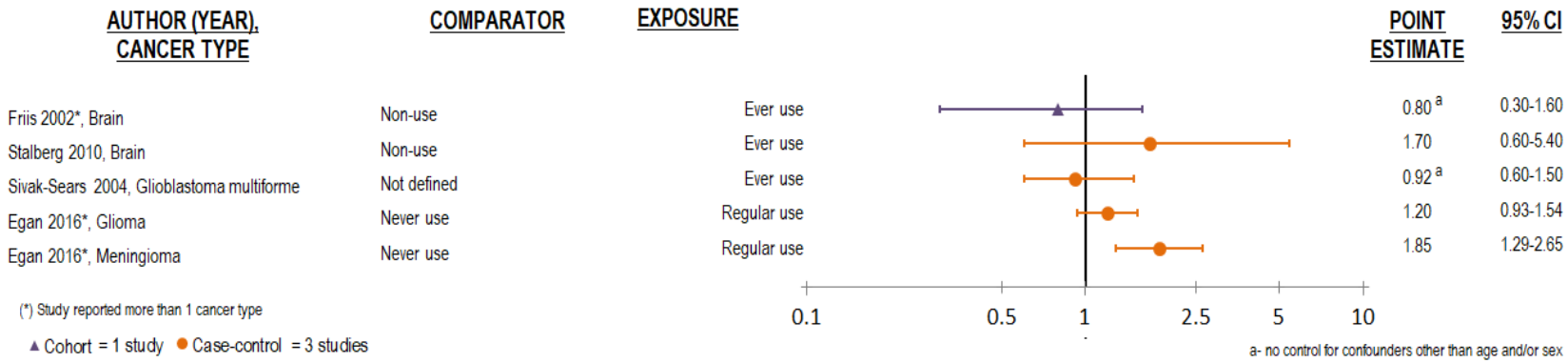


Figure 35. Forest plot: lung cancer

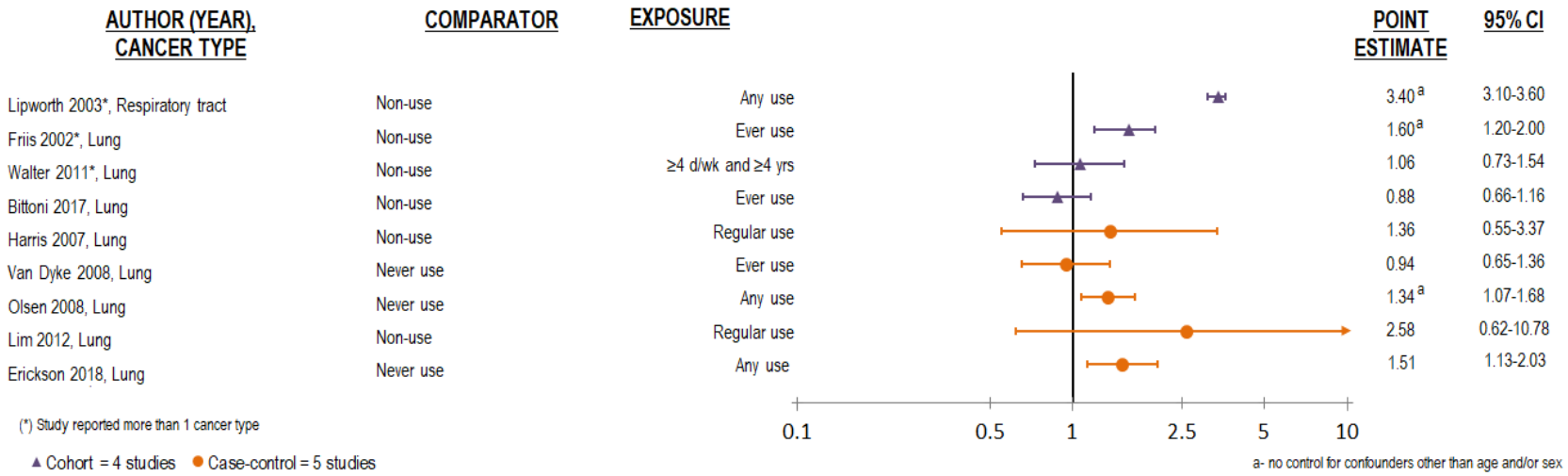


Figure 36. Forest plot: gastrointestinal cancer

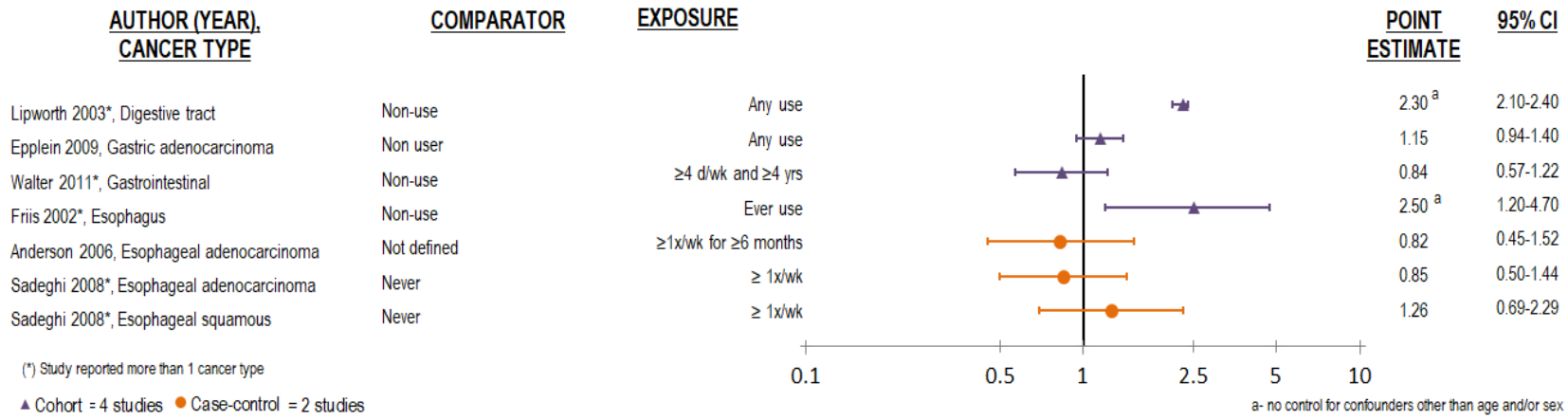


Figure 37. Forest plot: pancreatic cancer

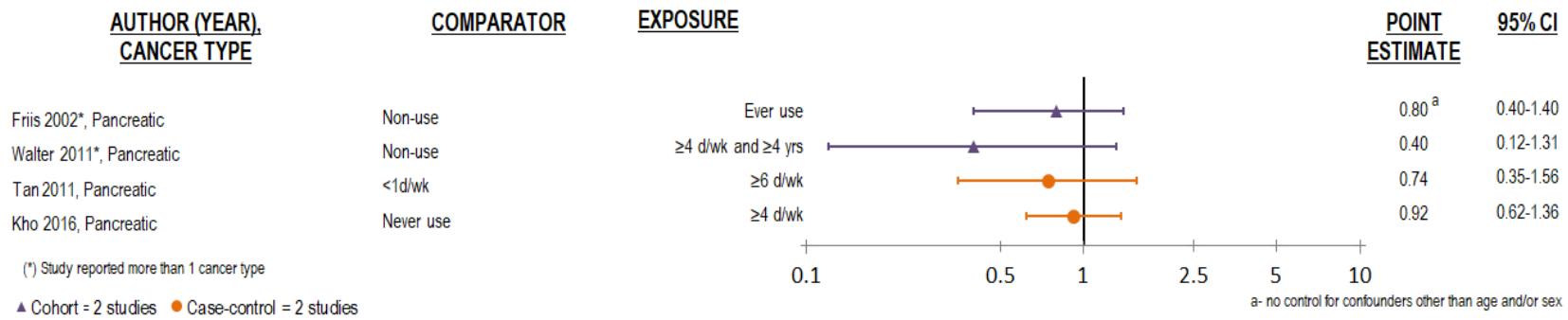
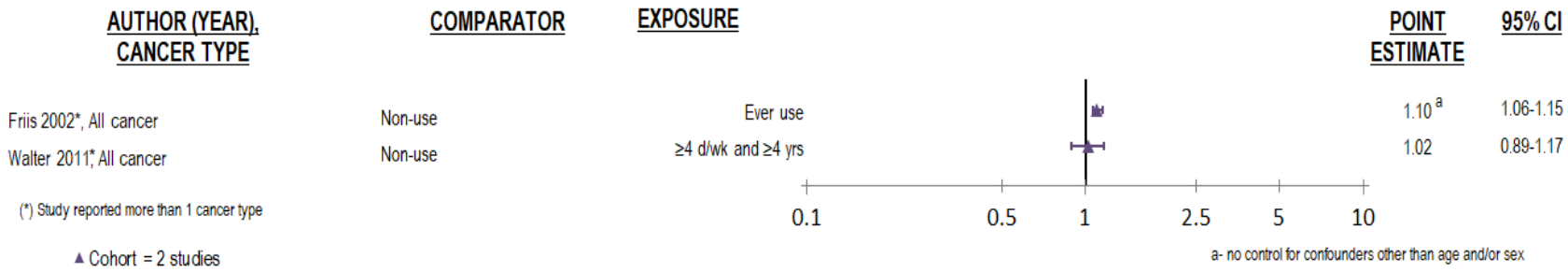


Figure 38. Forest plot: all cancers combined



8.4 Quantifying Bias in Epidemiologic Studies on the Association Between Acetaminophen and Cancer

8.4.1 Background and study design

Over 130 epidemiologic studies have been conducted to examine whether use of acetaminophen predisposes to the occurrence of one or more forms of cancer. There are many limitations to many of these studies as noted earlier, including vulnerability to channeling, protopathic bias, and uncontrolled confounding. However, the magnitude of the bias resulting from these limitation remains unknown, hampering the interpretability of the results of these studies.

Recent methodological developments have focused on using large sets of negative controls – exposure-outcome pairs where no causal effect is believed to exist – to measure the operating characteristics of study designs by observing to what extent these designs produce effect size estimates in line with the truth (that there is no effect for the negative controls). Previously, this approach has been used to show substantial bias in a comparative cohort study comparing acetaminophen to ibuprofen, even after adjustment using propensity scores (Weinstein et al., 2017).

Similarly, we set out to quantify bias in study designs used in observational research on the relationship between acetaminophen and cancer. The protocol for this study has been posted on-line at:

https://github.com/OHDSI/StudyProtocols/tree/master/QuantifyingBiasInApapStudies

We mimic the design choices made in prior studies as best we can in 10 different design variants and apply these designs to the well-known CPRD database. Each design is used to estimate the association between acetaminophen and a set of 37 negative control outcomes, outcomes a priori selected because they are known not to be caused by acetaminophen, as well as 4 cancer outcomes. This allowed us to see how far off the results for the negative controls are from the truth (that there is no effect), as well as how far away the results for the cancer outcomes are from the negative controls.

Of the 10 designs we evaluated, 8 are variants of the case-control design, where we systematically varied the mechanism by which controls were selected, how exposure was defined, and which covariates were used to adjust for potential confounding. The other 2 designs are variants of the cohort design, where we used the study by Walter et al. (2011) as a prime example of such studies (Walter et al., 2011b).

Data source:

CPRD (Clinical Practice Research Datalink) which was used in Kaye (2001), Yang et al (2016), McGlynn et al (2015) and other studies.

Methods:

- **Population:**
  - Case-control: restricted age to 30 years and older.
  - Cohort study: restricted to ages 50-76 at baseline, excluding people with prior history of cancer other than nonmelanoma skin cancer reported at baseline
- **Cancers:** We include 4 types of cancer which have been associated with acetaminophen use in prior studies:
  - Renal cell carcinoma
  - Primary liver cancer[5]
  - Lymphoma
  - Multiple myeloma

Negative controls:

Negative control outcomes are those determined a priori to have no association with the exposure of interest. We used the same set of negative control outcomes as an earlier study (Weinstein et al., 2017). Briefly, the negative outcomes must meet the following requirements to be considered as negative controls:

(1) that there is no Medline abstract where the MeSH terms suggest a negative association between the drug and the condition (Winnenburg et al., 2015).
(2) that there is no mention of the drug-condition pair on a US Product Label in the "Adverse Drug Reactions" or "Postmarketing" section.
(3) there are no US spontaneous reports suggesting that the pair is in an adverse event relationship.

Other steps are taken to ensure the controls are well-identified within the disease code vocabulary (see the protocol in Section 8.4 for further details). Once potential negative control candidates were selected, manual clinical review to exclude any pairs that may still be in a causal relationship or similar to the study outcome was performed to select the top 50 or so concepts by patient exposure.

The 37 negative control outcomes we used from the prior study are as follows:

1. Achilles tendinitis

[5] Although hepatocellular carcinoma specifically might be of more clinical interest than the broader 'Primary Liver Cancer' selected here, the data do not support a finer distinction. Most primary liver cancers are coded as 'Primary malignant neoplasm of liver' (READ code B150.00). For this reason we define our outcome of interest as 'Primary liver cancer', similar to other studies performed in CPRD.

2. Atrophic vaginitis
3. Breath smells unpleasant
4. Bronchiectasis
5. Disorders of initiating and maintaining sleep
6. Ear problem
7. Falls
8. Foot-drop
9. Ganglion and cyst of synovium, tendon and bursa
10. Hemangioma
11. Hydrocele
12. Hyperthyroidism
13. Impaired glucose tolerance
14. Impingement syndrome of shoulder region
15. Impotence
16. Incontinence of feces
17. Interpersonal relationship finding
18. Irregular periods
19. Irritability and anger
20. Joint stiffness
21. Loss of sense of smell
22. Mixed hyperlipidemia
23. Osteitis deformans
24. Panic attack
25. Perforation of tympanic membrane
26. Pes planus
27. Premature menopause
28. Prolapse of female genital organs
29. Pure hypercholesterolemia
30. Respiratory symptom
31. Restless legs
32. Restlessness and agitation
33. Rosacea
34. Simple goiter
35. Skin sensation disturbance
36. Snapping thumb syndrome
37. Urinary symptoms

Each design variant described below was used to estimate effect sizes for the negative controls as well as the outcomes of interest.

Case-control studies

Selection of controls: The case-control studies selected controls in 2 ways:

1. Sampling index dates from the distribution observed for cases, and randomly applying these to viable controls (i.e. non-cases that were observed at the index date).
2. Randomly selecting up to four matched controls per case. The matching variables were age, sex index date, time observed prior to index date, practice

Exposure status: Some studies implemented a 1-year lag assuming exposures within the year prior to index were not believed biologically plausible for any effect to occur within a shorter time frame. We evaluated 2 definitions of exposure:

1. All time prior: exposed on or any time prior to the index date, where the index date is the date of the outcome (for cases).
2. One-year delay: exposed on or any time prior to the index date, where the index date is one year before the date of the outcome (for cases).

Statistical model: After controls had been selected, exposure status was ascertained, and covariates were constructed, we fit a logistic regression to estimate the effect size (odds ratio) and 95% confidence interval. For those analyses where controls were matched to cases this regression was conditioned on the matched sets.

Table 20. Case-control design analysis variants.

ANALYSIS ID	CONTROL SELECTION	EXPOSURE STATUS	COVARIATE ADJUSTMENT
1	Sampling	All time prior	Age, sex, index year
2	Sampling	All time prior	Age, sex, index year, BMI, alcohol, smoking, diabetes
3	Sampling	One-year delay	Age, sex, index year
4	Sampling	One-year delay	Age, sex, index year, BMI, alcohol, smoking, diabetes
5	Matching	All time prior	None
6	Matching	All time prior	BMI, alcohol, smoking, diabetes
7	Matching	One-year delay	None
8	Matching	One-year delay	BMI, alcohol, smoking, diabetes

These 8 analyses were used to estimate odds ratios for all 37 negative controls and 4 outcomes of interest, resulting in 8 x (37 + 4) = 328 odds ratios and confidence intervals.

Cohort studies

Design variables: sex, smoking, Charlson Index (instead of self-rated health), history of RA, history of arthritis or chronic neck/back/joint pain, history of fatigue or lack of energy, and history of migraines or frequent headaches. The following variables in the Walter et al. study could not be included because they are not available in CPRD: race/ethnicity, education, number of first-degree relatives with a history of leukemia or lymphoma.

Exposure status: Similar to Walter et al, we focused on 'high use', defined in the original study as >= 4 days/week for >= 4 years. In our analysis, we classified subjects as 'exposed' if they were continuously exposed in the 4 years prior to the index date, allowing for gaps representing use of acetaminophen only 4 out of 7 days, with a minimum allowed gap of 30 days.

- Subjects were classified as 'unexposed' if they were not prescribed any acetaminophen in the 4 years prior to the index date.
- Similar to Walter et al.(Walter et al., 2011b) a separate analysis was performed excluding those who experienced the outcome in the 2 years following the index date.

Table 21. Cohort design analysis variants.

ANALYSIS ID	EXCLUDE SUBJECTS WITH THE OUTCOME IN THE 2 YEARS FOLLOWING THE INDEX DATE
9	No
10	Yes

These 2 analyses were used to estimate hazard ratios for all 37 negative controls and 4 outcomes of interest, resulting in 2 x (37 + 4) = 82 hazard ratios and confidence intervals.

Patient characteristics for the cohort study

Descriptive analyses were based on covariate balance of the variables described in the protocol. These include demographics and parameterizations of all conditions, drug exposures, procedures, the Charlson Index as well as other characteristics.

Additionally, the propensity score was estimated for each patient using a large-scale propensity score approach. Below, we provide the propensity score distribution plot for exposed and unexposed to assess comparability.

An explicit head-to-head comparison between 2 cohorts of baseline covariates, using standardized difference as a measure comparing individual factors, was conducted. Covariates with standardized difference > 10% were highlighted as potential imbalanced confounding factors.

Quantification of bias

We plot the estimated odds ratios/hazard ratios and standard errors (linearly related to the width of the confidence interval).

Study designs that adequately control for confounding factors should produce odds ratio estimates in line with the known true effect size (i.e., a odds ratio/hazard ratio of 1.0) for the negative control outcomes.

We compute the percentage of negative controls having a p-value below 0.05, with the expectation that for an unbiased study design this percentage should be 5%.

Results:

Case-control studies (see forest plots below)

- The outcomes of interest were within the range of systematic error of the negative controls, therefore could not be distinguished.
    - This can be seen in the forest plots. The negative controls (blue points) vary widely and the RRs of the outcomes of interest (yellow diamonds) fall within the range of variation in the negative controls.
- Designs which exclude exposure in the year prior to index have less bias than those which do not.
- Regardless of study design, the RR for the outcomes of interest were within the range of systematic error of the negative controls.

Table 22. Count and fraction of negative controls (for which there was enough data to compute an estimate) having a (two-sided) p < 0.05.

Analysis ID	Description	Controls with estimate	Controls significant	Fraction significant (p < 0.05)
1	Sampling, all time prior, adj. for age, sex & year	21	21	100.0%
2	Sampling, all time prior, adj. for age, sex, year, BMI, alcohol, smoking & diabetes	22	21	95.5%
3	Sampling, year delay, adj. for age, sex & year	21	18	85.7%
4	Sampling, year delay, adj. for age, sex, year, BMI, alcohol, smoking & diabetes	21	20	95.2%
5	Matching, all time prior	35	35	100.0%

6	Matching, all time prior, adj. for BMI, alcohol, smoking & diabetes	35	35	100.0%
7	Matching, year delay	35	35	100.0%
8	Matching, year delay, adj. for BMI, alcohol, smoking & diabetes	35	35	100.0%

Case-control study results

Analysis 1: Sampling, all time prior, adj. for age, sex & year

Figure 39:Forest plot showing point estimate and 95% confidence intervals for all negative controls and outcomes of interest.

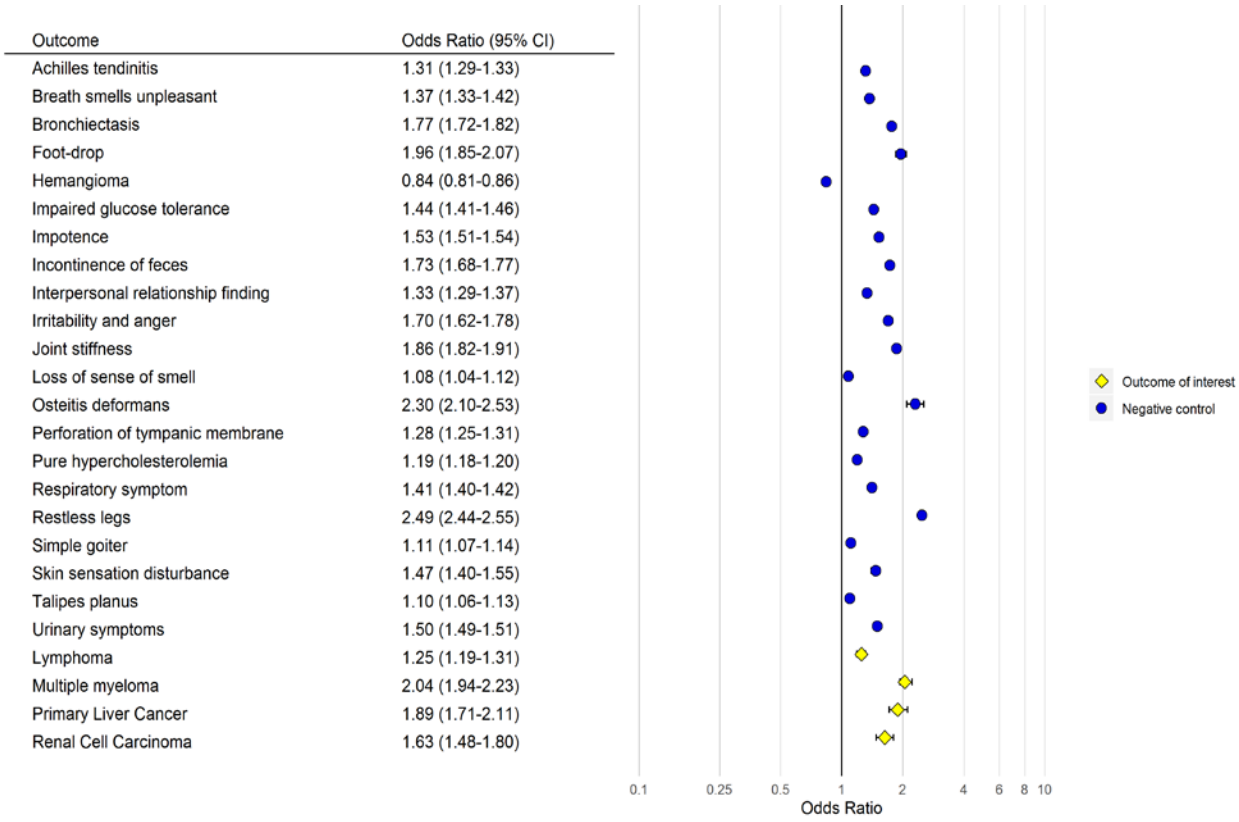
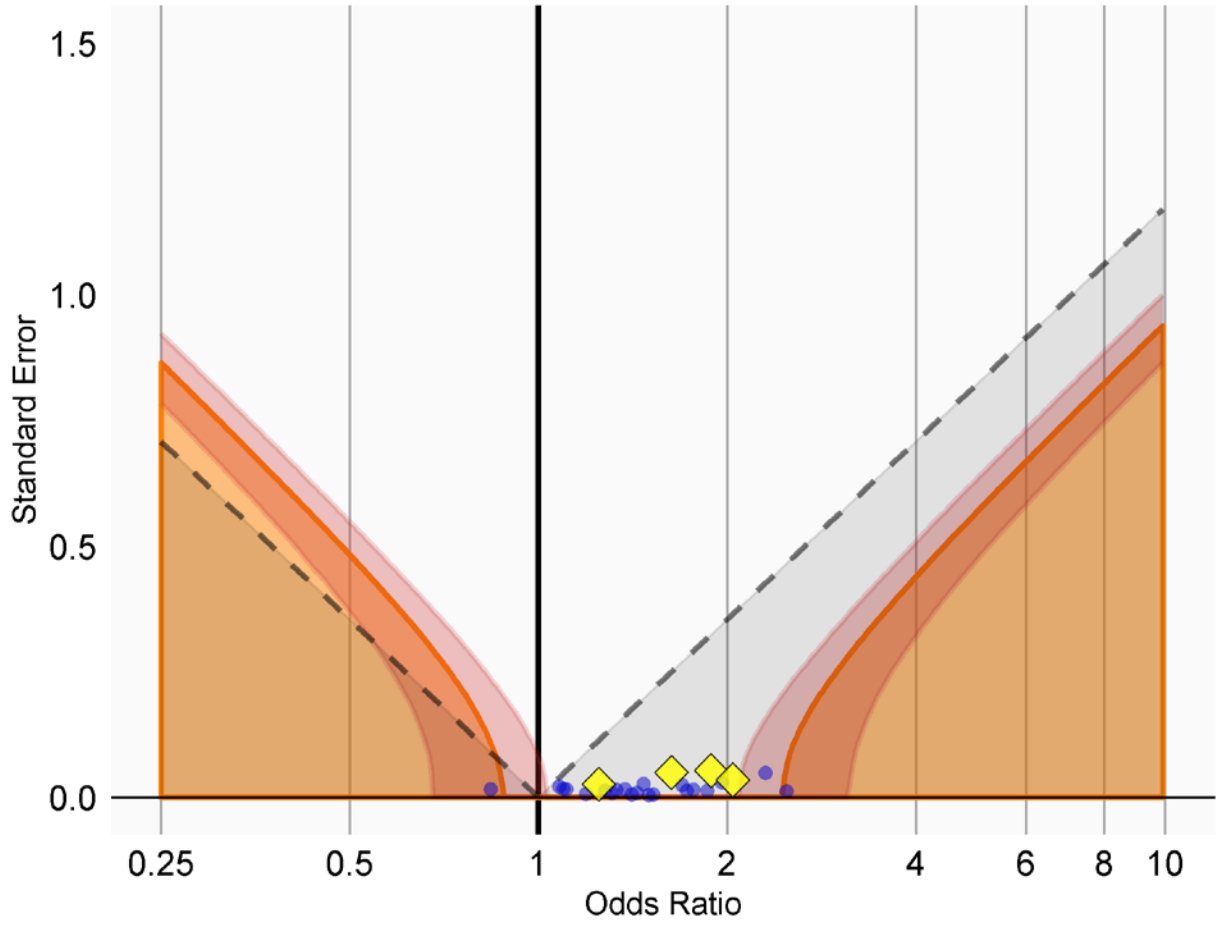


Figure 40: *Bias plot, showing effect size on the x-axis, and standard error (related to the width of the confidence interval) on the y-axis. Blue dots indicate negative controls, yellow diamonds indicate outcomes of interest. Estimates below the dashed lines have p < 0.05 using traditional p-value calculation. Estimates in the orange area have calibrated p < 0.05. The pink area denotes the 95% credible interval around the boundary of the orange area.*



Analysis 2: Sampling, all time prior, adj. for age, sex, year, BMI, alcohol, smoking & diabetes

Figure 41: Forest plot showing point estimate and 95% confidence intervals for all negative controls and outcomes of interest.

Outcome	Odds Ratio (95% CI)
Achilles tendinitis	1.21 (1.19-1.23)
Breath smells unpleasant	1.34 (1.30-1.38)
Bronchiectasis	1.83 (1.78-1.88)
Foot-drop	1.88 (1.77-1.99)
Hemangioma	0.82 (0.80-0.85)
Impaired glucose tolerance	1.18 (1.16-1.20)
Impotence	1.30 (1.30-1.32)
Incontinence of feces	1.68 (1.64-1.73)
Interpersonal relationship finding	1.22 (1.19-1.26)
Irritability and anger	1.50 (1.43-1.58)
Joint stiffness	1.74 (1.70-1.78)
Loss of sense of smell	1.06 (1.01-1.10)
Osteitis deformans	2.32 (2.11-2.56)
Perforation of tympanic membrane	1.22 (1.19-1.25)
Premature menopause	1.00 (0.92-1.09)
Pure hypercholesterolemia	1.05 (1.04-1.06)
Respiratory symptom	1.29 (1.28-1.30)
Restless legs	2.30 (2.25-2.35)
Simple goiter	1.08 (1.05-1.12)
Skin sensation disturbance	1.39 (1.32-1.46)
Talipes planus	1.05 (1.02-1.09)
Urinary symptoms	1.42 (1.41-1.43)
Lymphoma	1.27 (1.21-1.34)
Multiple myeloma	2.20 (2.09-2.31)
Primary Liver Cancer	1.70 (1.60-1.82)
Renal Cell Carcinoma	1.56 (1.42-1.65)

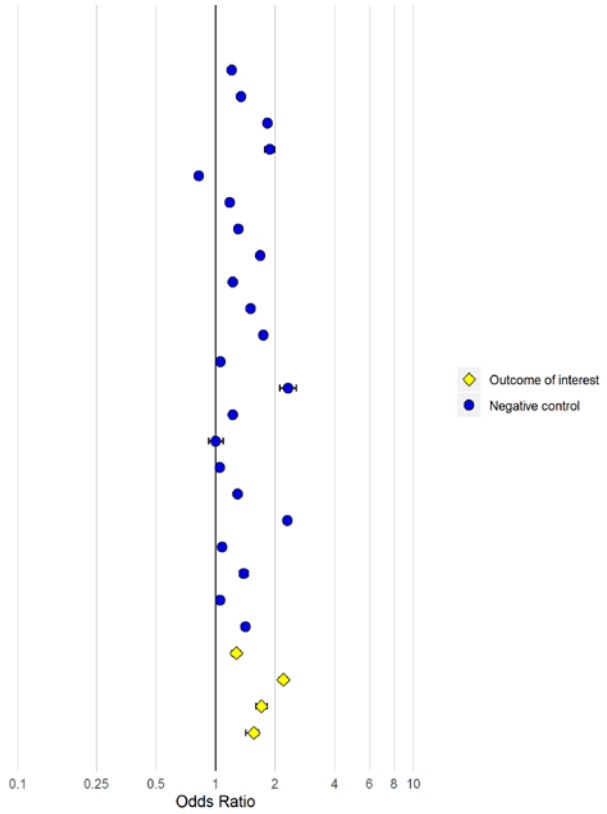
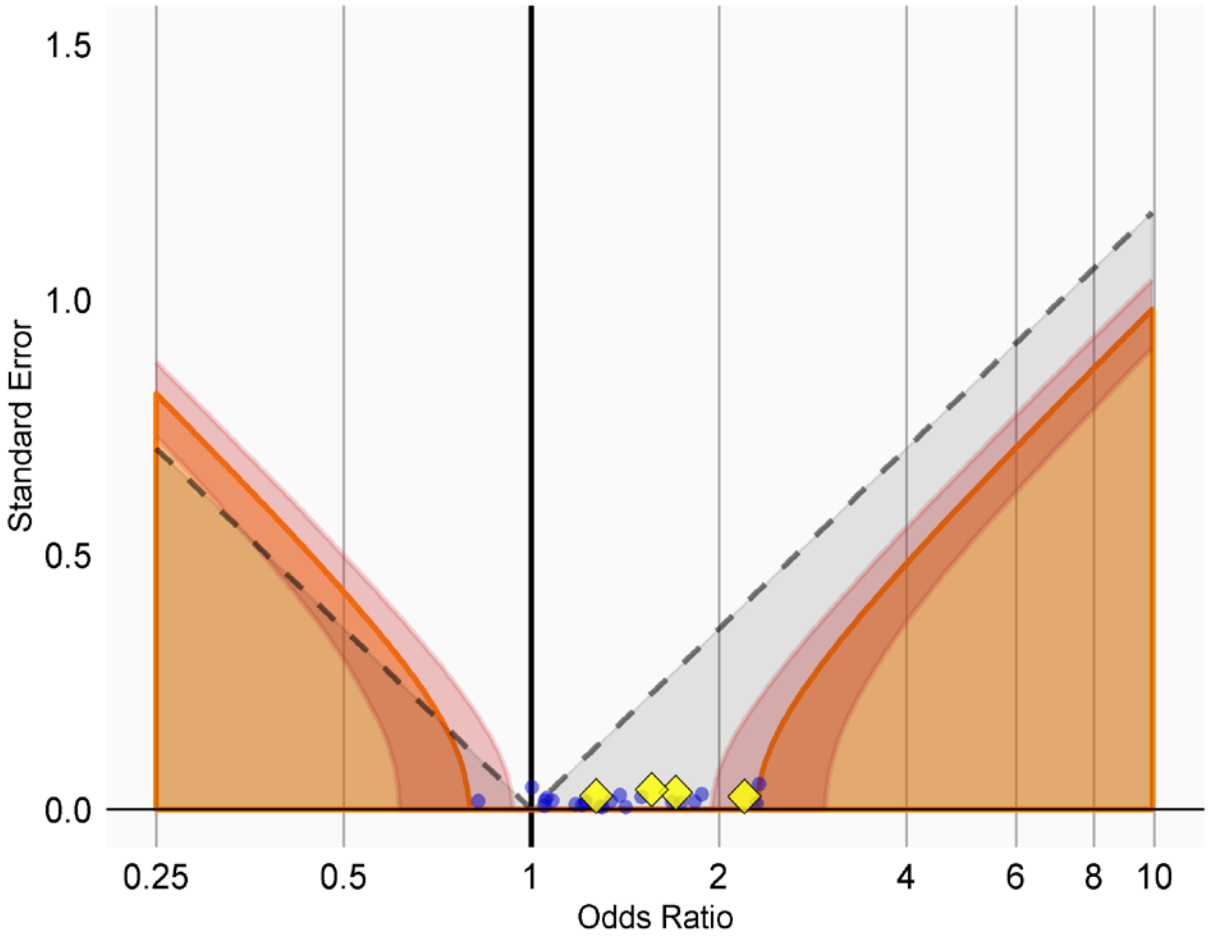


Figure 42: Bias plot, showing effect size on the x-axis, and standard error (related to the width of the confidence interval) on the y-axis. Blue dots indicate negative controls, yellow diamonds indicate outcomes of interest. Estimates below the dashed lines have p < 0.05 using traditional p-value calculation. Estimates in the orange area have calibrated p < 0.05. The pink area denotes the 95% credible interval around the boundary of the orange area.



Analysis 3: Sampling, year delay, adj. for age, sex & year

Figure 43: Forest plot showing point estimate and 95% confidence intervals for all negative controls and outcomes of interest.

Outcome	Odds Ratio (95% CI)
Achilles tendinitis	1.18 (1.17-1.20)
Breath smells unpleasant	1.18 (1.14-1.22)
Bronchiectasis	1.55 (1.50-1.59)
Foot-drop	1.38 (1.30-1.45)
Hemangioma	0.75 (0.75-0.77)
Impaired glucose tolerance	1.30 (1.28-1.33)
Impotence	1.37 (1.35-1.38)
Incontinence of feces	1.38 (1.34-1.41)
Interpersonal relationship finding	1.17 (1.14-1.21)
Irritability and anger	1.45 (1.38-1.52)
Joint stiffness	1.44 (1.41-1.48)
Loss of sense of smell	1.00 (0.96-1.04)
Osteitis deformans	1.47 (1.35-1.61)
Perforation of tympanic membrane	1.10 (1.08-1.13)
Pure hypercholesterolemia	1.07 (1.06-1.09)
Respiratory symptom	1.25 (1.24-1.26)
Restless legs	2.02 (1.97-2.06)
Simple goiter	0.97 (0.94-1.01)
Skin sensation disturbance	1.29 (1.22-1.35)
Talipes planus	0.98 (0.94-1.01)
Urinary symptoms	1.32 (1.31-1.33)
Lymphoma	0.89 (0.84-0.93)
Multiple myeloma	0.95 (0.87-1.03)
Primary Liver Cancer	1.21 (1.12-1.35)
Renal Cell Carcinoma	1.13 (1.02-1.24)

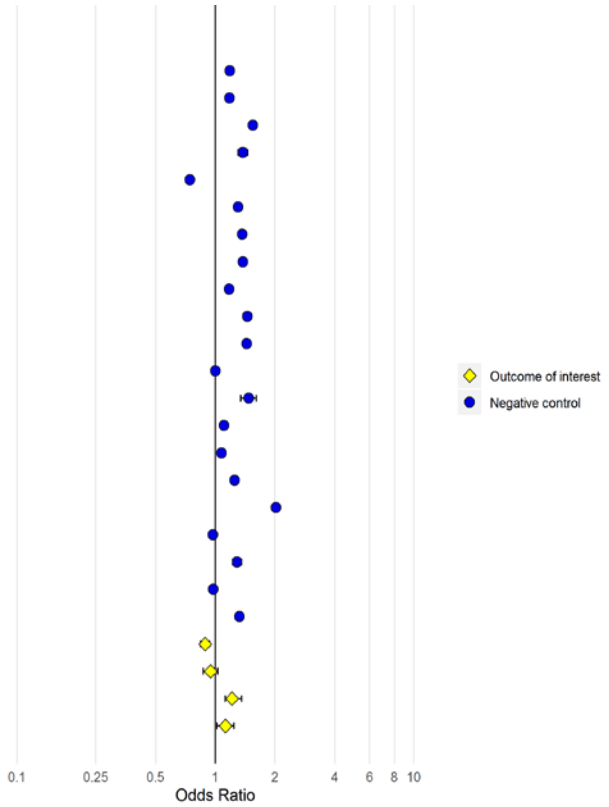
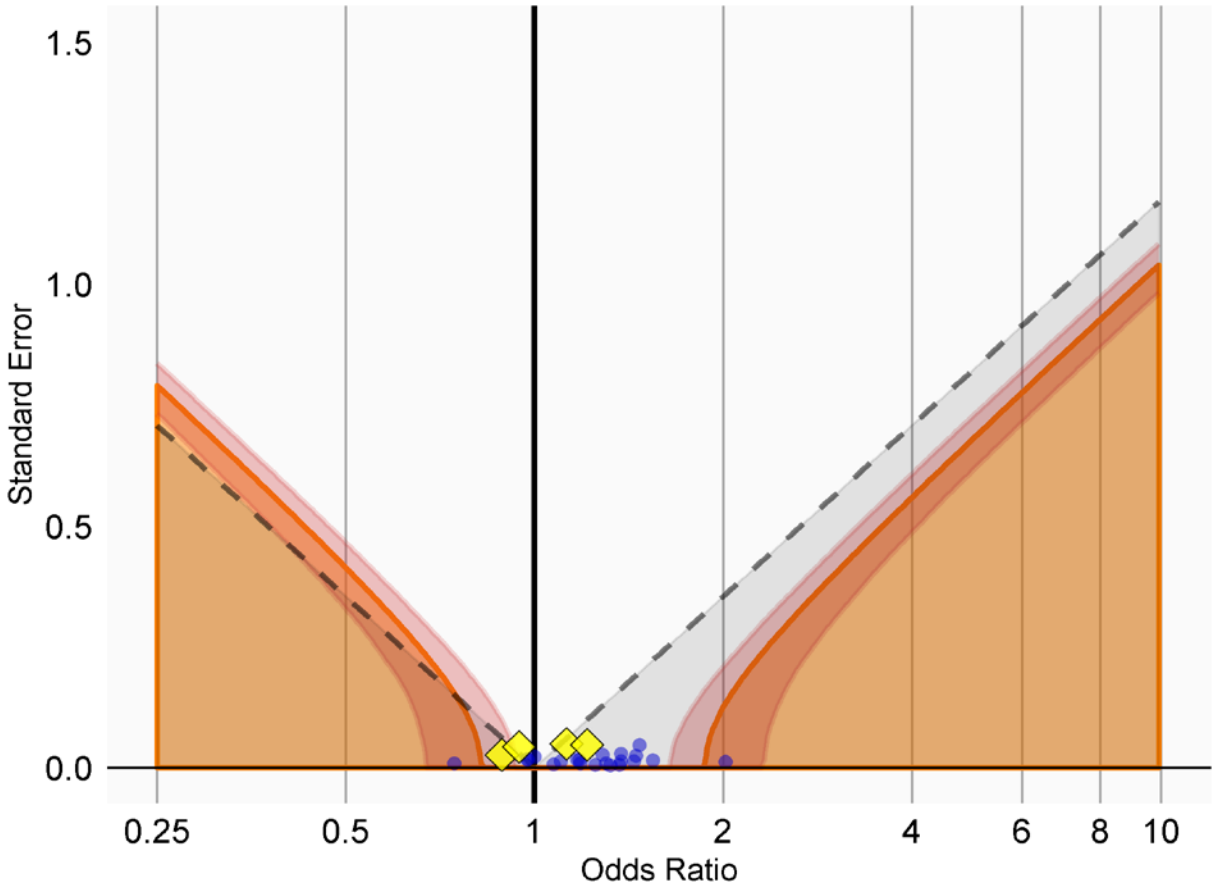


Figure 44: Bias plot, showing effect size on the x-axis, and standard error (related to the width of the confidence interval) on the y-axis. Blue dots indicate negative controls, yellow diamonds indicate outcomes of interest. Estimates below the dashed lines have p < 0.05 using traditional p-value calculation. Estimates in the orange area have calibrated p < 0.05. The pink area denotes the 95% credible interval around the boundary of the orange area.



Analysis 4: Sampling, year delay, adj. for age, sex, year, BMI, alcohol, smoking & diabetes

Figure 45: Forest plot showing point estimate and 95% confidence intervals for all negative controls and outcomes of interest.

Outcome	Odds Ratio (95% CI)
Achilles tendinitis	1.09 (1.07-1.10)
Breath smells unpleasant	1.14 (1.11-1.18)
Bronchiectasis	1.59 (1.55-1.64)
Foot-drop	1.30 (1.27-1.37)
Hemangioma	0.73 (0.71-0.75)
Impaired glucose tolerance	1.06 (1.04-1.08)
Impotence	1.16 (1.15-1.17)
Incontinence of feces	1.33 (1.30-1.37)
Interpersonal relationship finding	1.06 (1.03-1.10)
Irritability and anger	1.27 (1.21-1.34)
Joint stiffness	1.33 (1.29-1.36)
Loss of sense of smell	0.98 (0.93-1.02)
Osteitis deformans	1.47 (1.41-1.61)
Perforation of tympanic membrane	1.05 (1.02-1.07)
Pure hypercholesterolemia	0.94 (0.93-0.95)
Respiratory symptom	1.14 (1.13-1.15)
Restless legs	1.84 (1.80-1.88)
Simple goiter	0.94 (0.91-0.98)
Skin sensation disturbance	1.20 (1.14-1.27)
Talipes planus	0.93 (0.90-0.97)
Urinary symptoms	1.24 (1.23-1.25)
Lymphoma	0.89 (0.84-0.94)
Multiple myeloma	0.98 (0.90-1.07)
Primary Liver Cancer	1.06 (0.93-1.19)

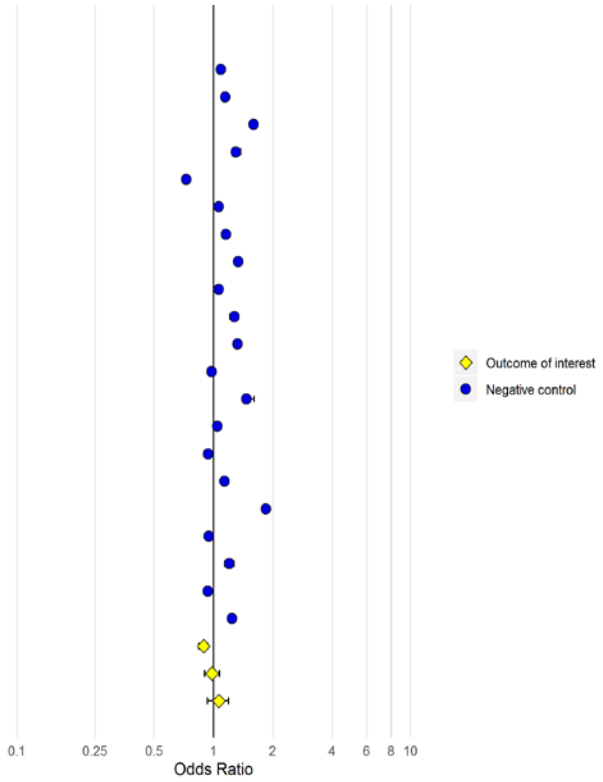
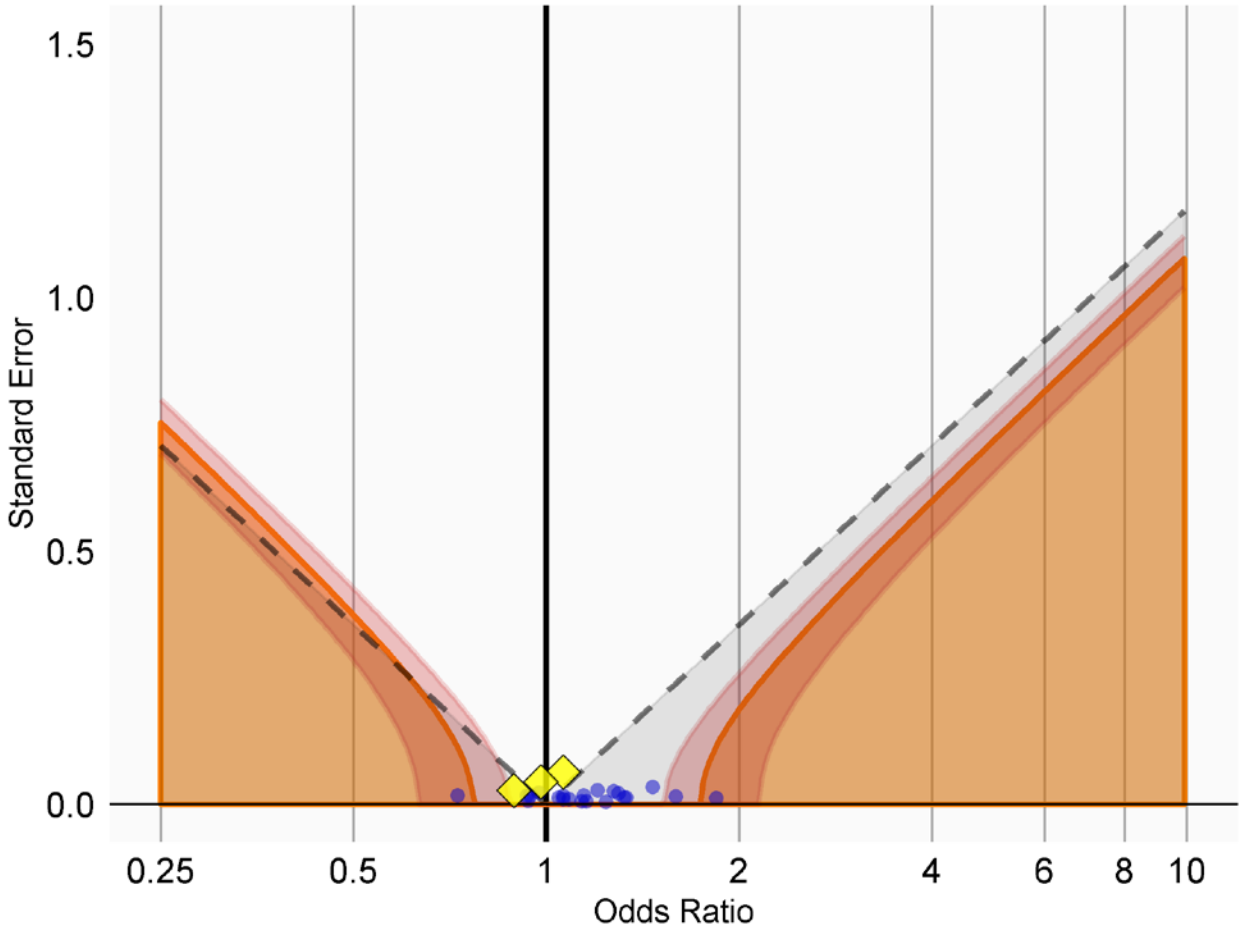


Figure 46: Bias plot, showing effect size on the x-axis, and standard error (related to the width of the confidence interval) on the y-axis. Blue dots indicate negative controls, yellow diamonds indicate outcomes of interest. Estimates below the dashed lines have p < 0.05 using traditional p-value calculation. Estimates in the orange area have calibrated p < 0.05. The pink area denotes the 95% credible interval around the boundary of the orange area.



Analysis 5: Matching, all time prior

Figure 47: Forest plot showing point estimate and 95% confidence intervals for all negative controls and outcomes of interest.

Outcome	Odds Ratio (95% CI)
Achilles tendinitis	1.65 (1.62-1.69)
Atrophic vaginitis	1.38 (1.35-1.41)
Breath smells unpleasant	1.74 (1.66-1.83)
Bronchiectasis	2.11 (2.02-2.21)
Disorders of initiating and maintaining sleep	2.49 (2.46-2.53)
Ear problem	1.57 (1.52-1.61)
Falls	2.50 (2.43-2.56)
Foot-drop	2.59 (2.41-2.78)
Hemangioma	1.30 (1.24-1.36)
Hyperthyroidism	1.43 (1.39-1.48)
Impaired glucose tolerance	1.60 (1.57-1.64)
Impingement syndrome of shoulder region	2.48 (2.39-2.57)
Impotence	1.69 (1.67-1.72)
Incontinence of feces	2.19 (2.12-2.26)
Interpersonal relationship finding	1.86 (1.78-1.93)
Irregular periods	1.36 (1.33-1.38)
Irritability and anger	2.64 (2.45-2.86)
Joint stiffness	2.27 (2.20-2.33)
Loss of sense of smell	1.38 (1.31-1.46)
Mixed hyperlipidemia	1.58 (1.53-1.63)
Osteitis deformans	2.53 (2.25-2.85)
Panic attack	2.10 (2.06-2.15)
Perforation of tympanic membrane	1.70 (1.63-1.76)
Premature menopause	1.45 (1.35-1.56)
Prolapse of female genital organs	1.36 (1.33-1.38)
Pure hypercholesterolemia	1.38 (1.36-1.40)
Respiratory symptom	1.70 (1.68-1.73)
Restless legs	2.85 (2.76-2.94)
Restlessness and agitation	2.54 (2.38-2.71)
Rosacea	1.24 (1.22-1.26)
Simple goiter	1.32 (1.27-1.38)
Skin sensation disturbance	1.89 (1.77-2.00)
Snapping thumb syndrome	1.69 (1.62-1.76)
Talipes planus	1.82 (1.72-1.93)
Urinary symptoms	1.88 (1.86-1.90)
Lymphoma	1.61 (1.52-1.70)
Multiple myeloma	2.92 (2.67-3.21)
Primary Liver Cancer	2.14 (1.91-2.40)
Renal Cell Carcinoma	2.02 (1.82-2.25)

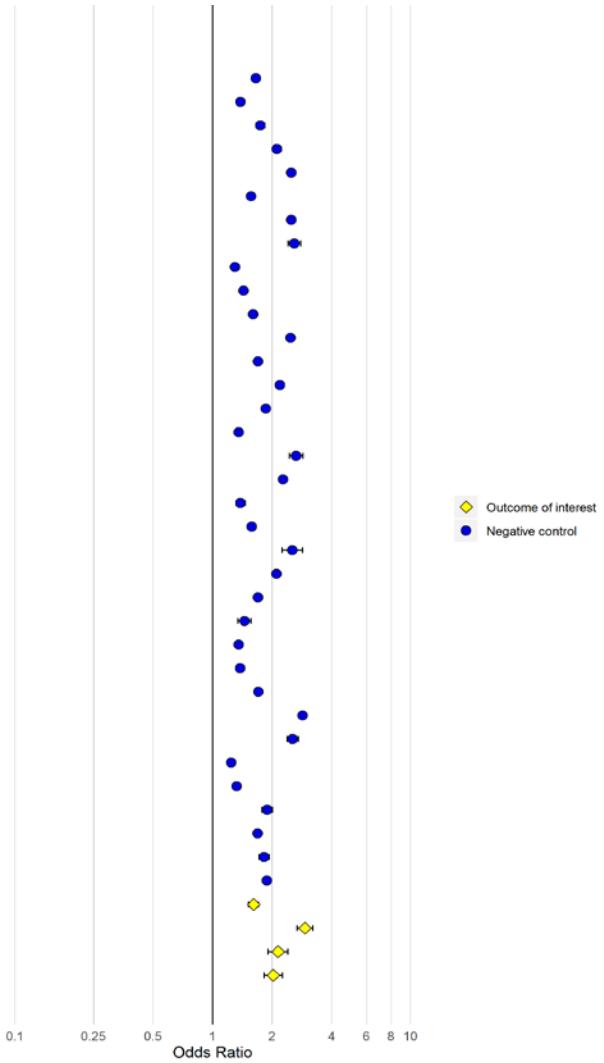
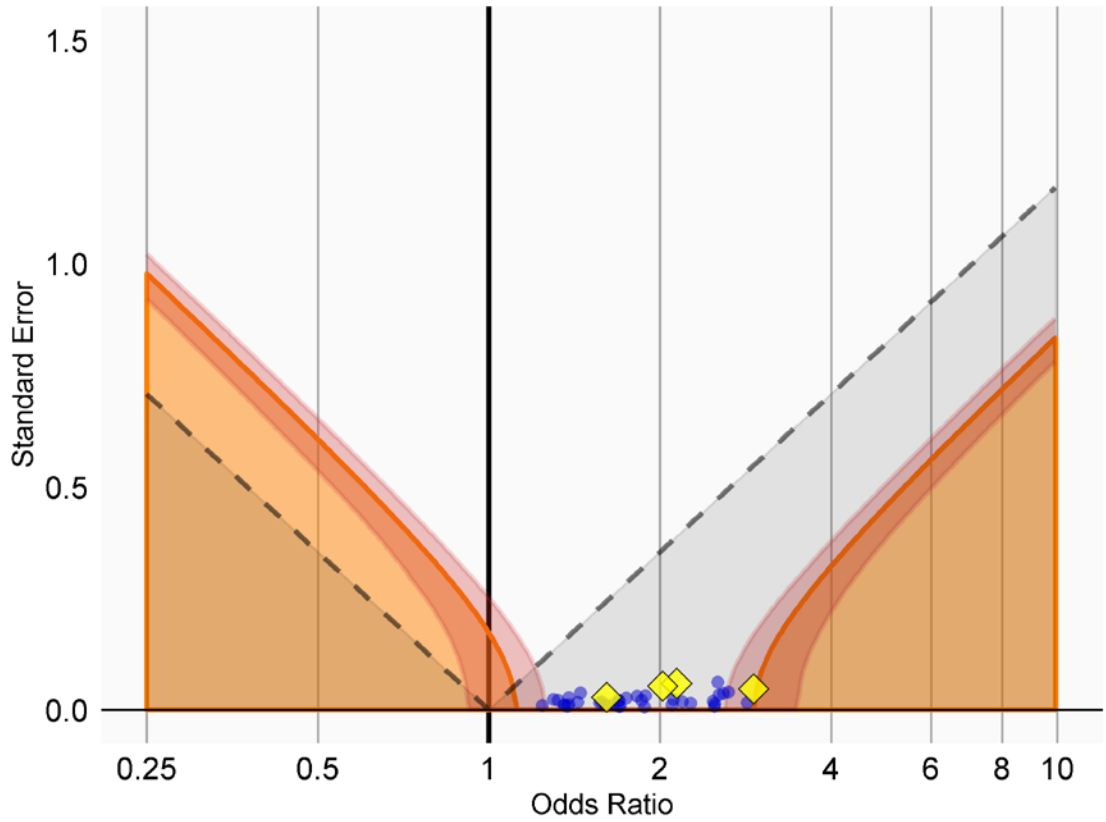


Figure 48: Bias plot, showing effect size on the x-axis, and standard error (related to the width of the confidence interval) on the y-axis. Blue dots indicate negative controls, yellow diamonds indicate outcomes of interest. Estimates below the dashed lines have p < 0.05 using traditional p-value calculation. Estimates in the orange area have calibrated p < 0.05. The pink area denotes the 95% credible interval around the boundary of the orange area.



Analysis 6: Matching, all time prior, adj. for BMI, alcohol, smoking & diabetes

Figure 49: Forest plot showing point estimate and 95% confidence intervals for all negative controls and outcomes of interest.

Outcome	Odds Ratio (95% CI)
Achilles tendinitis	1.52 (1.49-1.56)
Atrophic vaginitis	1.45 (1.42-1.48)
Breath smells unpleasant	1.71 (1.62-1.79)
Bronchiectasis	2.10 (2.01-2.20)
Disorders of initiating and maintaining sleep	2.38 (2.35-2.41)
Ear problem	1.51 (1.46-1.55)
Falls	2.41 (2.35-2.48)
Foot-drop	2.47 (2.30-2.66)
Hemangioma	1.28 (1.22-1.34)
Hyperthyroidism	1.37 (1.33-1.42)
Impaired glucose tolerance	1.36 (1.33-1.39)
Impingement syndrome of shoulder region	2.35 (2.26-2.44)
Impotence	1.48 (1.46-1.50)
Incontinence of feces	2.10 (2.03-2.18)
Interpersonal relationship finding	1.76 (1.69-1.84)
Irregular periods	1.32 (1.29-1.35)
Irritability and anger	2.34 (2.16-2.53)
Joint stiffness	2.17 (2.11-2.24)
Loss of sense of smell	1.35 (1.27-1.42)
Mixed hyperlipidemia	1.33 (1.29-1.38)
Osteitis deformans	2.47 (2.19-2.79)
Panic attack	1.94 (1.90-1.98)
Perforation of tympanic membrane	1.62 (1.56-1.68)
Premature menopause	1.45 (1.34-1.56)
Prolapse of female genital organs	1.35 (1.32-1.38)
Pure hypercholesterolemia	1.24 (1.22-1.26)
Respiratory symptom	1.56 (1.54-1.58)
Restless legs	2.66 (2.58-2.75)
Restlessness and agitation	2.55 (2.39-2.73)
Rosacea	1.24 (1.21-1.26)
Simple goiter	1.29 (1.24-1.35)
Skin sensation disturbance	1.82 (1.71-1.93)
Snapping thumb syndrome	1.54 (1.47-1.60)
Talipes planus	1.71 (1.61-1.81)
Urinary symptoms	1.81 (1.78-1.83)
Lymphoma	1.59 (1.50-1.68)
Multiple myeloma	2.97 (2.70-3.26)
Primary Liver Cancer	1.84 (1.63-2.07)
Renal Cell Carcinoma	1.87 (1.68-2.08)

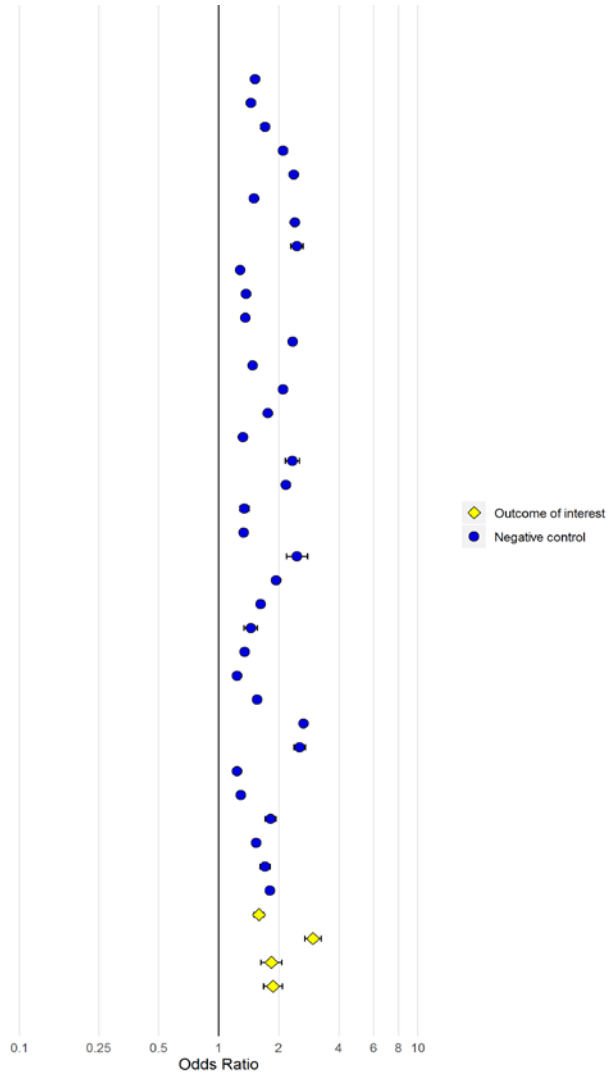
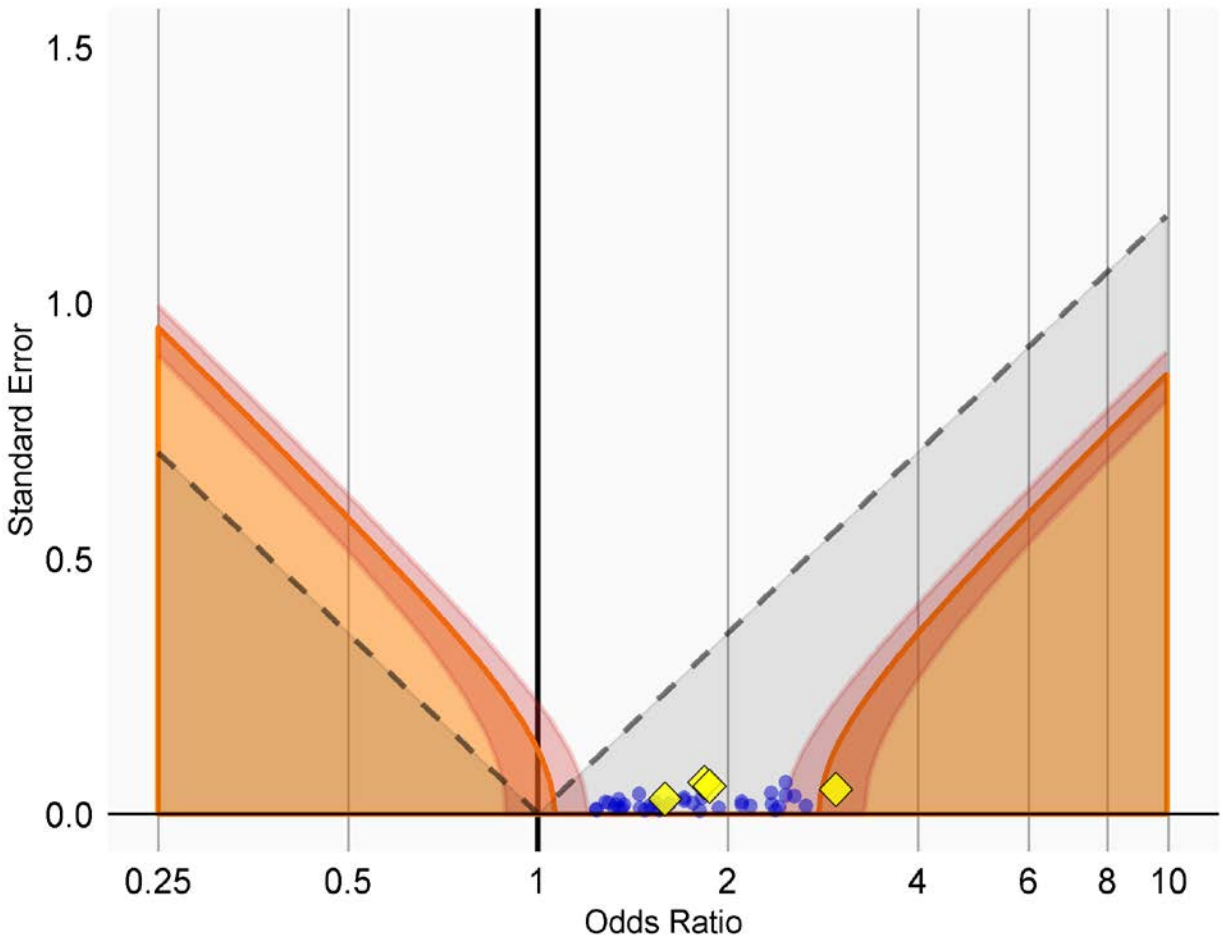


Figure 50: Bias plot, showing effect size on the x-axis, and standard error (related to the width of the confidence interval) on the y-axis. Blue dots indicate negative controls, yellow diamonds indicate outcomes of interest. Estimates below the dashed lines have p < 0.05 using traditional p-value calculation. Estimates in the orange area have calibrated p < 0.05. The pink area denotes the 95% credible interval around the boundary of the orange area.



Analysis 7: Matching, year delay

Figure 51: Forest plot showing point estimate and 95% confidence intervals for all negative controls and outcomes of interest.

Outcome	Odds Ratio (95% CI)
Achilles tendinitis	1.56 (1.53-1.60)
Atrophic vaginitis	1.37 (1.34-1.40)
Breath smells unpleasant	1.64 (1.56-1.73)
Bronchiectasis	2.02 (1.94-2.12)
Disorders of initiating and maintaining sleep	2.20 (2.17-2.23)
Ear problem	1.54 (1.49-1.58)
Falls	2.09 (2.04-2.14)
Foot-drop	1.96 (1.83-2.10)
Hemangioma	1.27 (1.21-1.33)
Hyperthyroidism	1.36 (1.32-1.41)
Impaired glucose tolerance	1.55 (1.52-1.58)
Impingement syndrome of shoulder region	2.04 (1.97-2.11)
Impotence	1.62 (1.60-1.65)
Incontinence of feces	2.03 (1.96-2.10)
Interpersonal relationship finding	1.79 (1.72-1.87)
Irregular periods	1.34 (1.31-1.37)
Irritability and anger	2.55 (2.36-2.75)
Joint stiffness	1.89 (1.84-1.95)
Loss of sense of smell	1.33 (1.26-1.40)
Mixed hyperlipidemia	1.52 (1.46-1.57)
Osteitis deformans	1.75 (1.56-1.96)
Panic attack	1.99 (1.94-2.03)
Perforation of tympanic membrane	1.60 (1.54-1.66)
Premature menopause	1.38 (1.28-1.49)
Prolapse of female genital organs	1.33 (1.30-1.36)
Pure hypercholesterolemia	1.33 (1.31-1.35)
Respiratory symptom	1.67 (1.64-1.69)
Restless legs	2.62 (2.55-2.70)
Restlessness and agitation	2.05 (1.93-2.19)
Rosacea	1.23 (1.21-1.25)
Simple goiter	1.28 (1.23-1.33)
Skin sensation disturbance	1.70 (1.60-1.81)
Snapping thumb syndrome	1.61 (1.55-1.68)
Talipes planus	1.70 (1.60-1.80)
Urinary symptoms	1.79 (1.76-1.81)
Lymphoma	1.25 (1.18-1.32)
Multiple myeloma	1.42 (1.31-1.54)
Primary Liver Cancer	1.60 (1.43-1.78)
Renal Cell Carcinoma	1.53 (1.38-1.69)

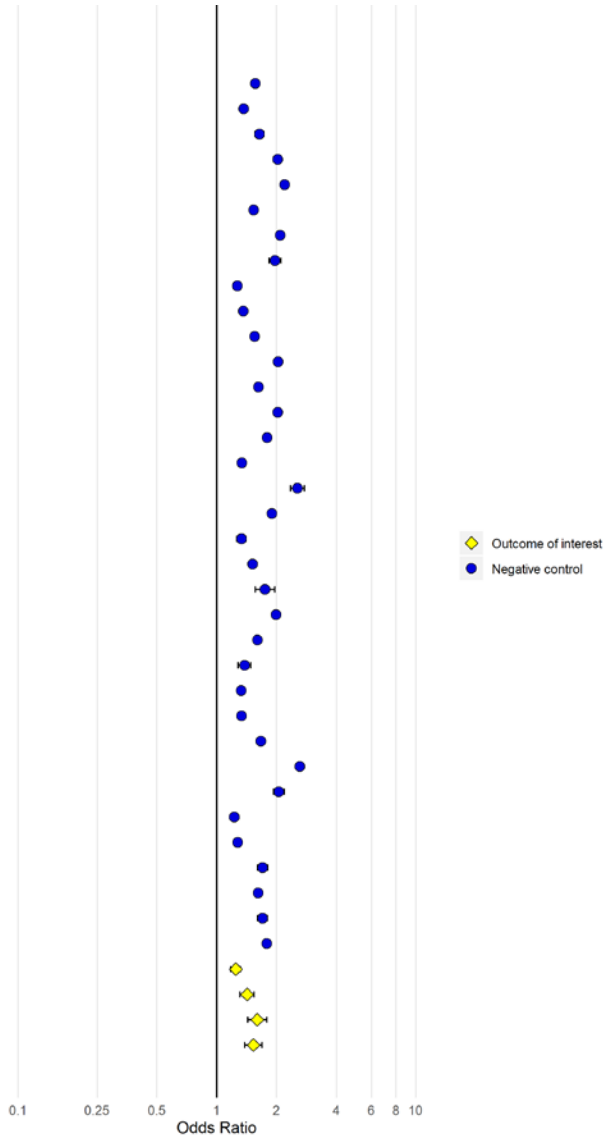
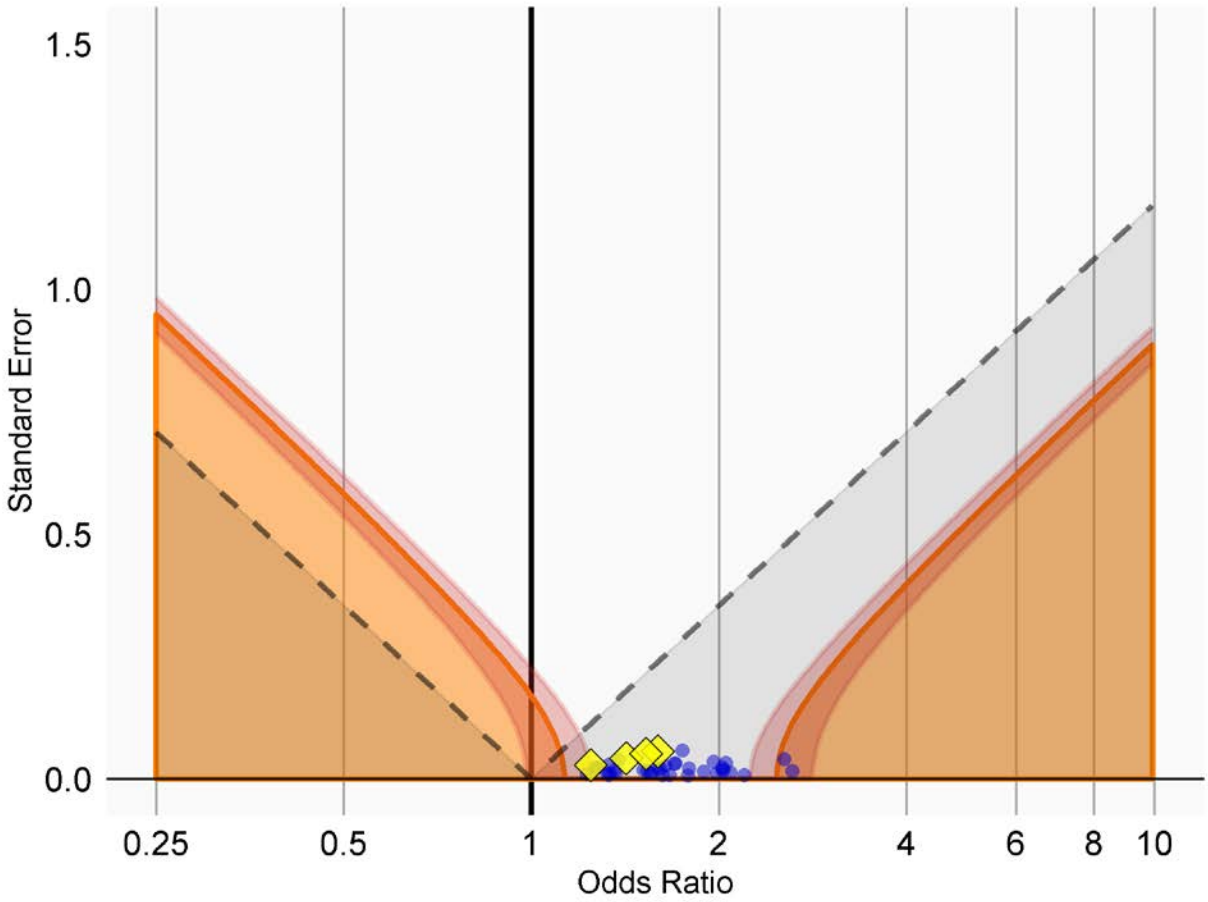


Figure 52: Bias plot, showing effect size on the x-axis, and standard error (related to the width of the confidence interval) on the y-axis. Blue dots indicate negative controls, yellow diamonds indicate outcomes of interest. Estimates below the dashed lines have p < 0.05 using traditional p-value calculation. Estimates in the orange area have calibrated p < 0.05. The pink area denotes the 95% credible interval around the boundary of the orange area.



Analysis 8: Matching, year delay, adj. for BMI, alcohol, smoking & diabetes

Figure 53: Forest plot showing point estimate and 95% confidence intervals for all negative controls and outcomes of interest.

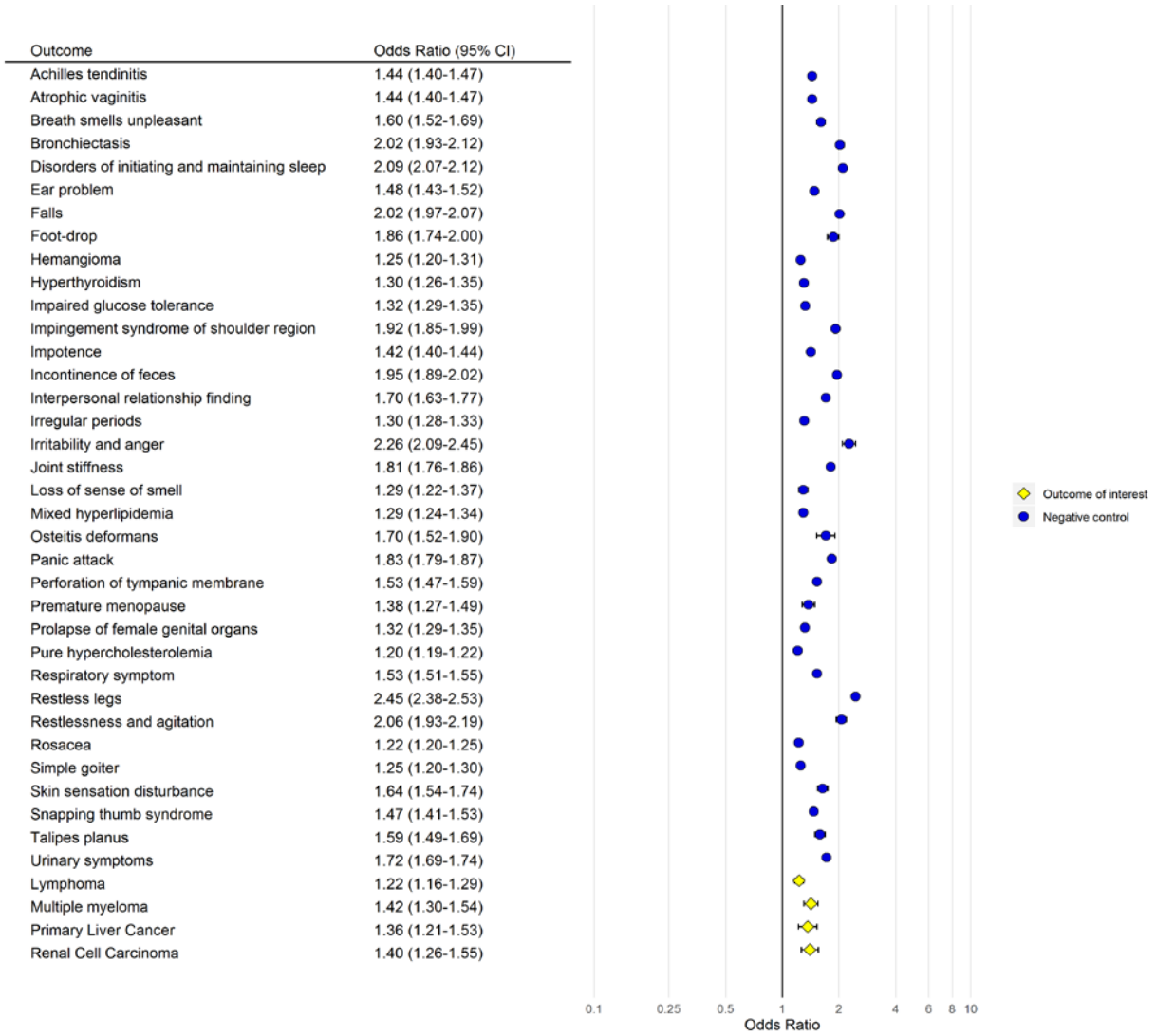
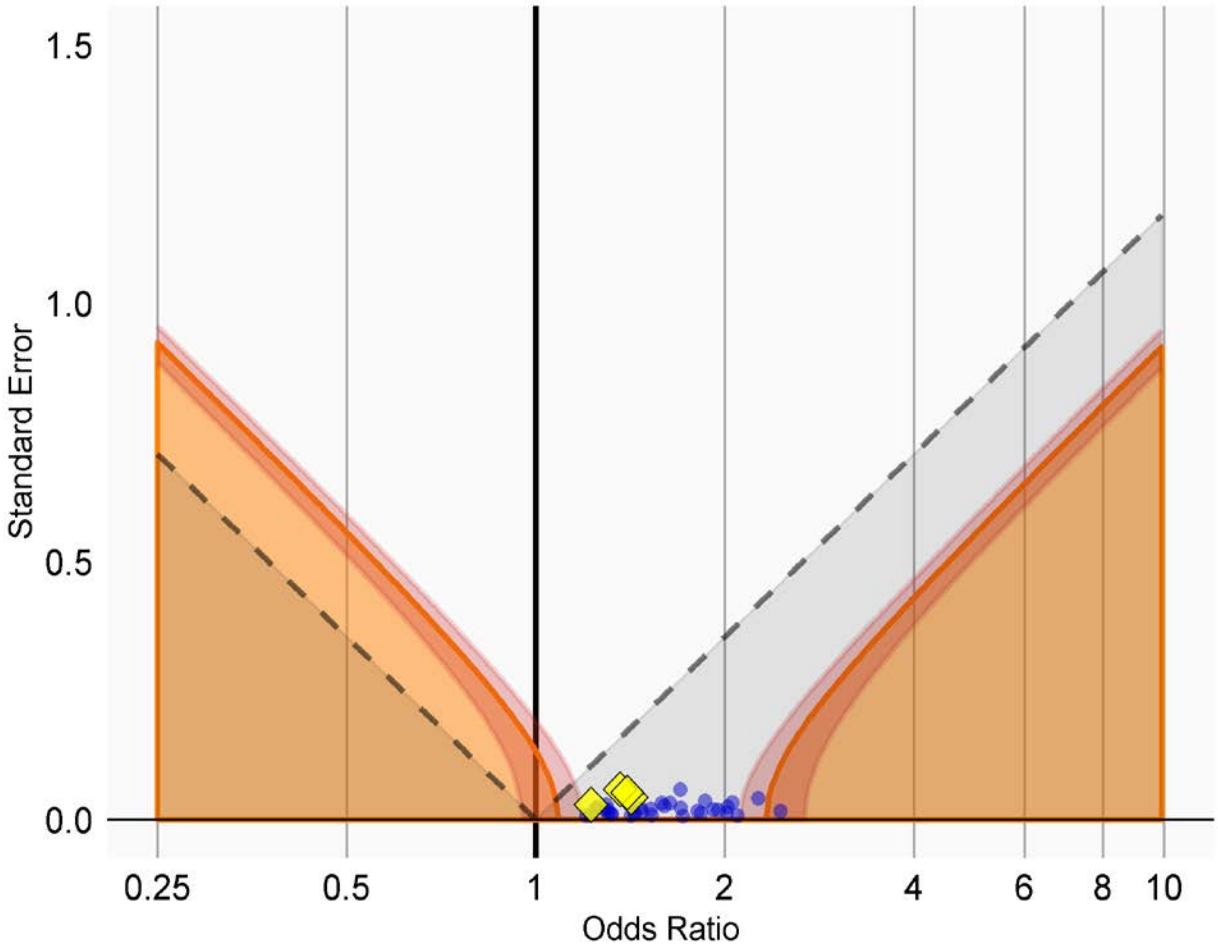


Figure 54: Bias plot, showing effect size on the x-axis, and standard error (related to the width of the confidence interval) on the y-axis. Blue dots indicate negative controls, yellow diamonds indicate outcomes of interest. Estimates below the dashed lines have p < 0.05 using traditional p-value calculation. Estimates in the orange area have calibrated p < 0.05. The pink area denotes the 95% credible interval around the boundary of the orange area.



8.4.2 Quantification of bias in cohort designs

Cohort studies

- Similarly, in the cohort study designs the negative control outcomes (blue points) show considerable error.
- The RRs of the outcomes of interest fell within the range of the negative controls.

Table 23. Count and fraction of negative controls (for which there was enough data to compute an estimate) having a (two-sided) p < 0.05.

Analysis ID	Description	Controls with estimate	Controls significant	Fraction significant (p < 0.05)
9	No delay	32	14	43.8%
10	Delay	28	8	28.6%

Table 24. Characteristics of the Cohort Study Emulating Walter 2011 Using CPRD

	High Users	None Users	
Characteristic	**% (n = 5,284)**	**% (n = 84,567)**	**Std. diff**
Age group			
75-79	32	41.7	-0.2
80-84	33.4	33	0.01
85-89	23.1	17.5	0.14
90-94	8.3	5.4	0.12
95-99	2.7	1.8	0.06
Gender: female	75.8	52.1	0.51
Medical history: General			
Acute respiratory disease	7.6	3.2	0.2
Chronic liver disease	0	0	0.01
Chronic obstructive lung disease	3.3	1.3	0.13
Crohn's disease	0	0	0.01
Dementia	0.9	0.9	0
Depressive disorder	1.5	0.5	0.1
Diabetes mellitus	2.7	2.0	0.05
Gastroesophageal reflux disease	0.3	0.1	0.03
Gastrointestinal hemorrhage	0.9	0.5	0.05
Hyperlipidemia	0.9	0.8	0.01
Hypertensive disorder	3.1	3.7	-0.03
Lesion of liver	0.1	0.0	0.02
Obesity	0.1	0.1	0.02
Osteoarthritis	7.9	1.5	0.31
Pneumonia	0.6	0.2	0.06
Psoriasis	0.3	0.3	0.01
Renal impairment	10.6	6.3	0.15
Rheumatoid arthritis	0.3	0.1	0.06
Urinary tract infectious disease	4.9	1.8	0.18
Visual system disorder	10.6	7.4	0.11
Medical history: Cardiovascular disease			

	High Users	None Users	
Characteristic	**% (n = 5,284)**	**% (n = 84,567)**	**Std. diff**
Atrial fibrillation	1.4	1.3	0.01
Cerebrovascular disease	1.1	1.0	0.01
Coronary arteriosclerosis	0.2	0.1	0.02
Heart disease	5	3.2	0.09
Heart failure	1.1	0.5	0.07
Ischemic heart disease	1.7	1.0	0.06
Peripheral vascular disease	1.8	1.2	0.05
Pulmonary embolism	0	0.1	-0.03
Venous thrombosis	0.7	0.5	0.02
Medical history: Neoplasms			
Malignant neoplastic disease	1.1	1.0	0
Medication use			
Agents acting on the renin-angiotensin system	48.2	32	0.33
Antibacterials for systemic use	51.8	26.1	0.55
Antidepressants	25.8	6.0	0.56
Antiepileptics	5.2	1.3	0.22
Antiinflammatory and antirheumatic products	23.3	8.2	0.42
Antineoplastic agents	1.2	0.2	0.12
Antipsoriatics	0.3	0.2	0.02
Antithrombotic agents	15.3	8.4	0.21
Beta blocking agents	27.5	21.7	0.13
Calcium channel blockers	27.1	20.0	0.17
Diuretics	60.0	32.1	0.58
Drugs for acid related disorders	46.8	14.9	0.74
Drugs for obstructive airway diseases	10.5	4.5	0.23
Drugs used in diabetes	9.8	6.4	0.13
Immunosuppressants	1.4	0.3	0.12
Lipid modifying agents	44.9	30.6	0.30
Opioids	3.8	0.10	0.27
Psycholeptics	30.8	8.6	0.58

Analysis 9: No delay

Figure 55:Forest plot showing point estimate and 95% confidence intervals for all negative controls and outcomes of interest.

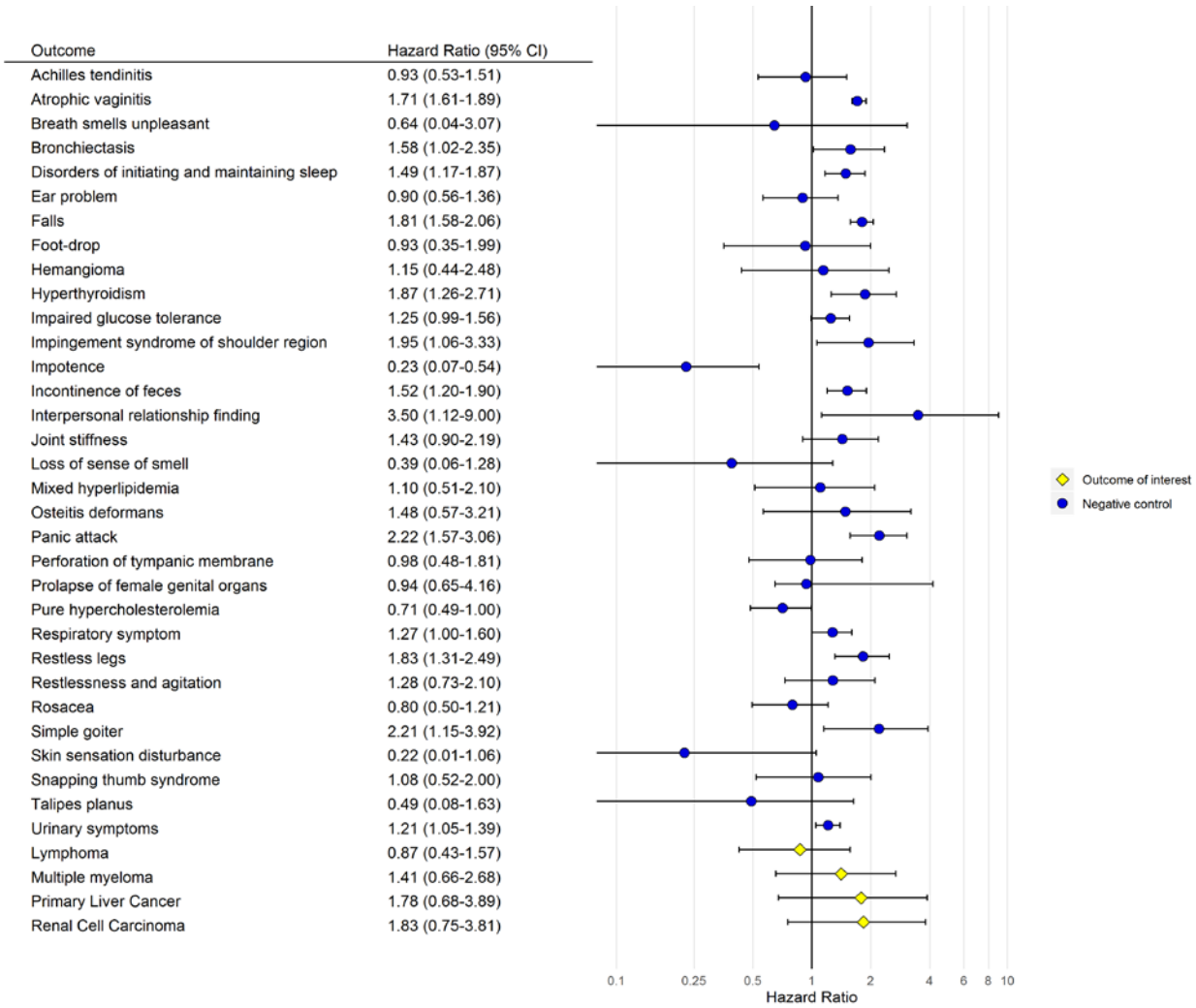
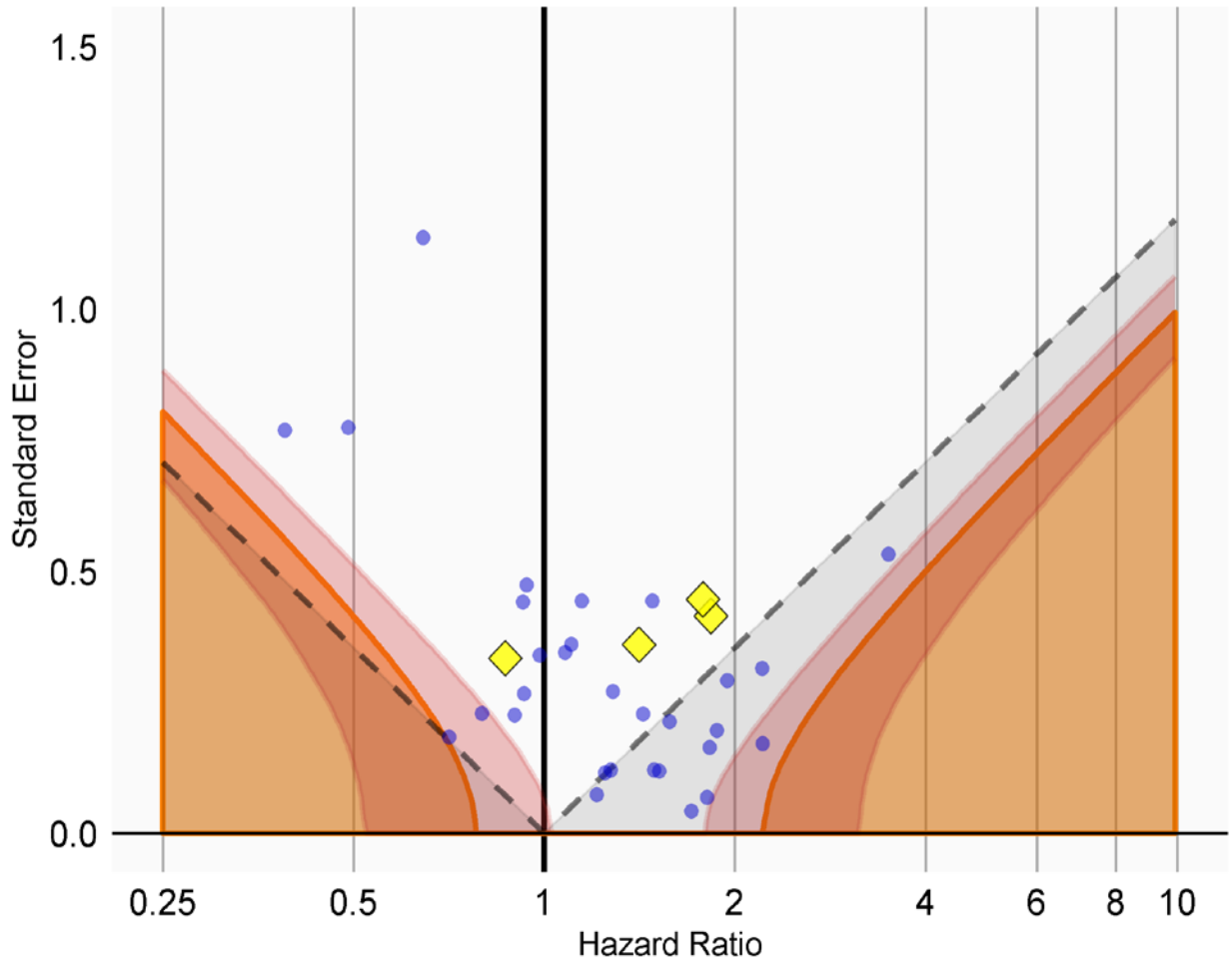


Figure 56: Bias plot, showing effect size on the x-axis, and standard error (related to the width of the confidence interval) on the y-axis. Blue dots indicate negative controls, yellow diamonds indicate outcomes of interest. Estimates below the dashed lines have p < 0.05 using traditional p-value calculation. Estimates in the orange area have calibrated p < 0.05. The pink area denotes the 95% credible interval around the boundary of the orange area.



Analysis 10: Delay

Figure 57: Forest plot showing point estimate and 95% confidence intervals for all negative controls and outcomes of interest.

Outcome	Hazard Ratio (95% CI)
Achilles tendinitis	1.18 (0.61-2.06)
Atrophic vaginitis	1.74 (1.40-2.28)
Bronchiectasis	1.46 (0.84-2.39)
Disorders of initiating and maintaining sleep	1.23 (0.86-1.71)
Ear problem	0.80 (0.42-1.39)
Falls	1.80 (1.53-2.11)
Foot-drop	0.80 (0.24-2.01)
Hemangioma	0.60 (0.10-1.98)
Hyperthyroidism	1.82 (1.07-2.92)
Impaired glucose tolerance	1.13 (0.84-1.48)
Impingement syndrome of shoulder region	1.95 (0.92-3.72)
Incontinence of feces	1.75 (1.32-2.28)
Interpersonal relationship finding	4.00 (1.12-11.14)
Joint stiffness	1.38 (0.74-2.37)
Loss of sense of smell	0.66 (0.11-2.24)
Mixed hyperlipidemia	0.81 (0.19-2.22)
Panic attack	2.05 (1.24-3.22)
Perforation of tympanic membrane	1.26 (0.48-2.71)
Pure hypercholesterolemia	0.61 (0.31-1.08)
Respiratory symptom	0.95 (0.67-1.31)
Restless legs	1.57 (1.00-2.36)
Restlessness and agitation	1.06 (0.49-2.00)
Rosacea	0.80 (0.42-1.37)
Simple goiter	2.47 (1.06-5.06)
Skin sensation disturbance	0.38 (0.02-1.84)
Snapping thumb syndrome	1.01 (0.35-2.32)
Talipes planus	1.12 (0.18-3.91)
Urinary symptoms	1.19 (1.00-1.41)
Lymphoma	1.07 (0.48-2.08)
Multiple myeloma	1.46 (0.60-3.00)
Primary Liver Cancer	2.12 (0.62-5.44)
Renal Cell Carcinoma	1.19 (0.29-3.36)

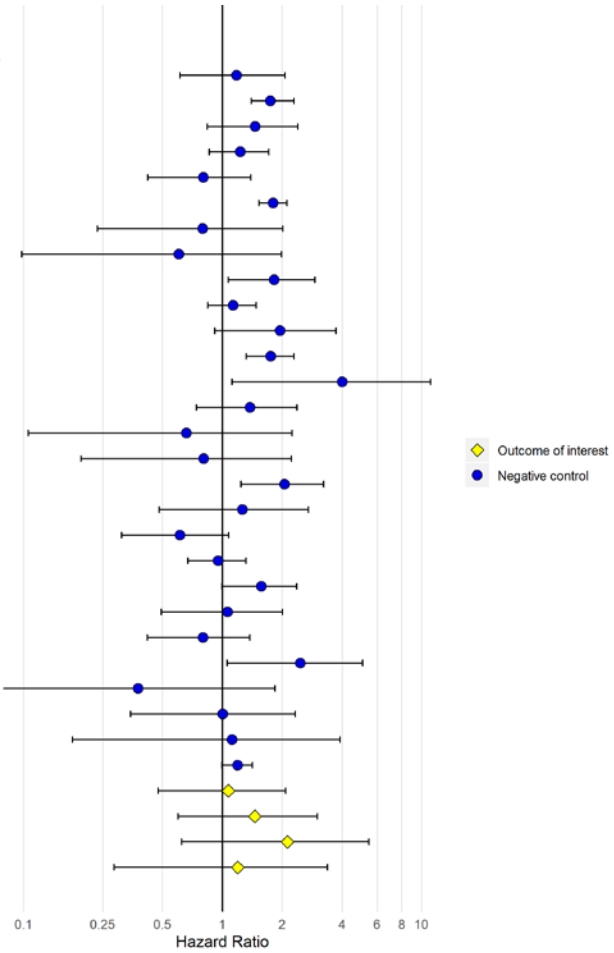
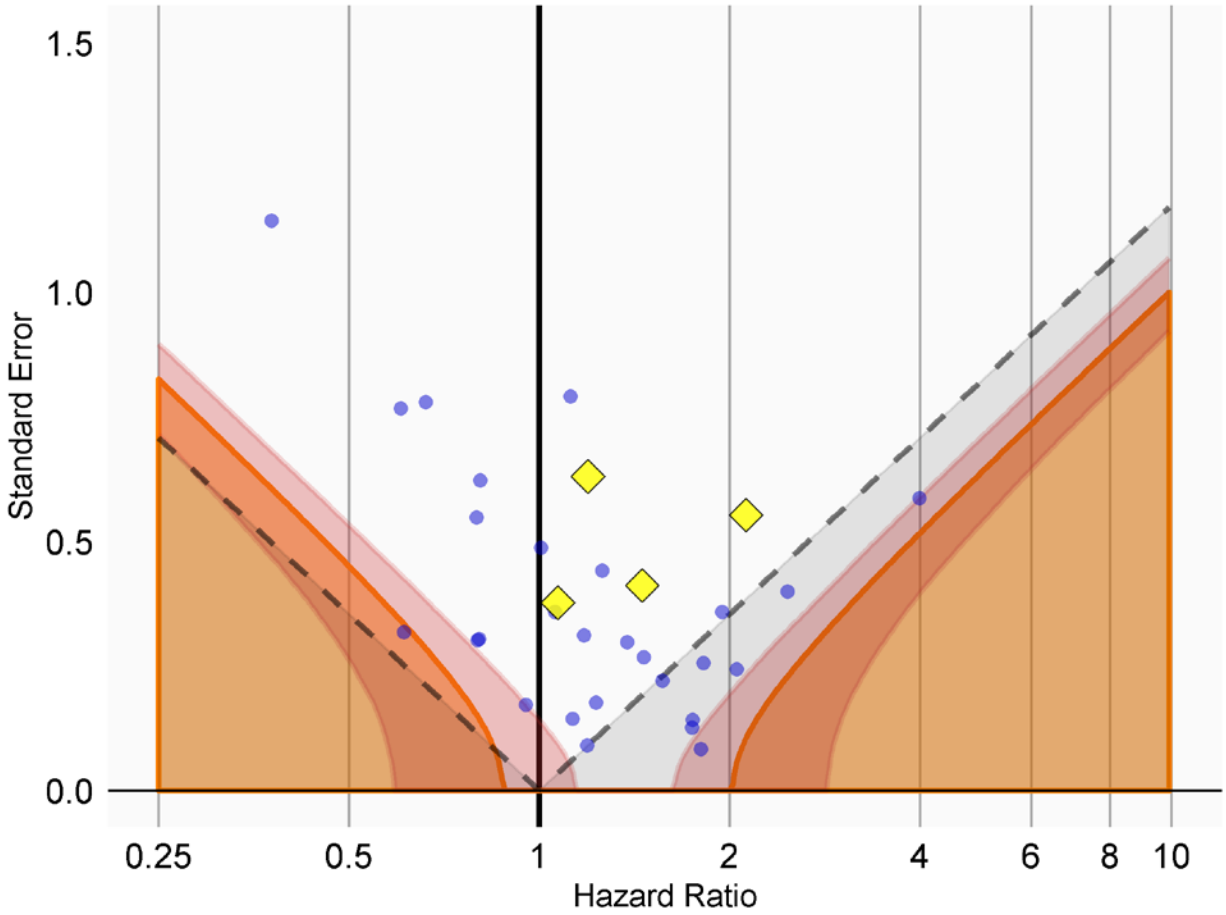


Figure 58: Bias plot, showing effect size on the x-axis, and standard error (related to the width of the confidence interval) on the y-axis. Blue dots indicate negative controls, yellow diamonds indicate outcomes of interest. Estimates below the dashed lines have p < 0.05 using traditional p-value calculation. Estimates in the orange area have calibrated p < 0.05. The pink area denotes the 95% credible interval around the boundary of the orange area.

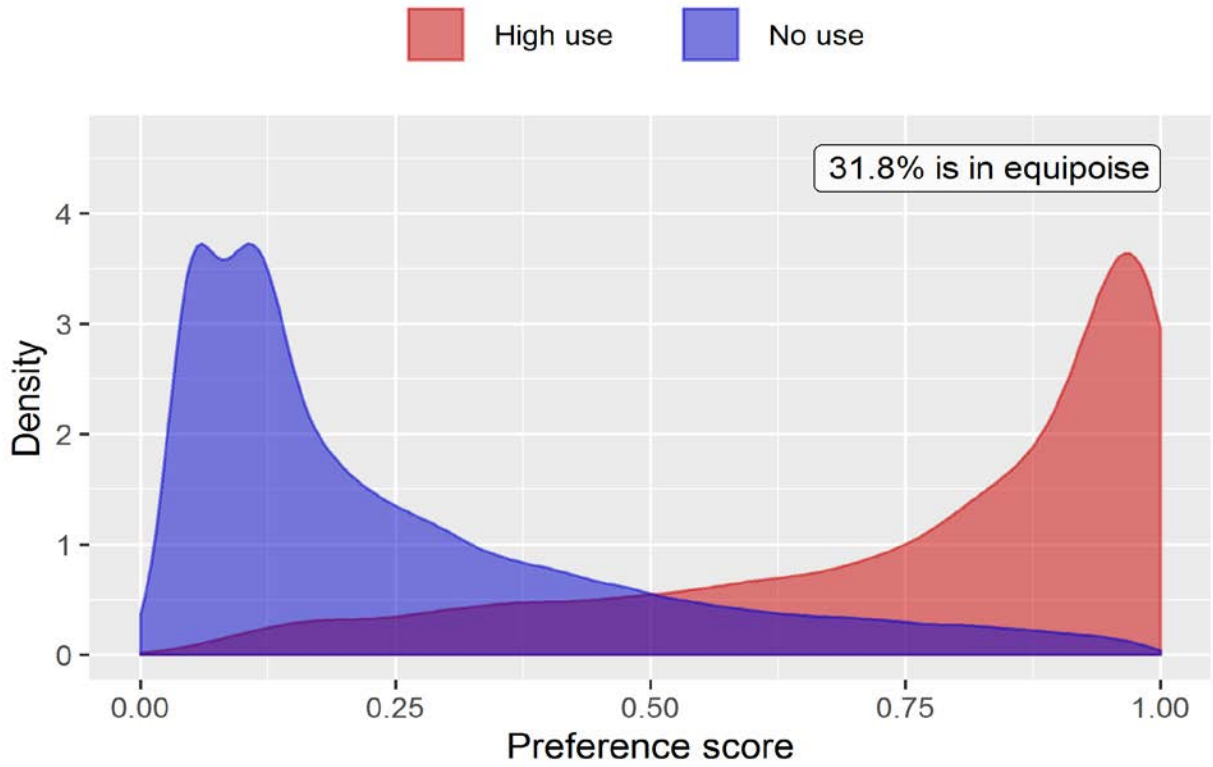


Propensity analysis

In our emulation of the Walter et al. (2011) study we also fitted a propensity model to evaluate to what extent the 2 exposure groups are comparable. This model was fitted by included a large set of covariates (all prior drugs, drug classes, diagnoses, procedures, etc.), and using a regularized logistic regression (Walter et al., 2011b).

Figure 59 shows the preference score distribution. The preference score is a transformation of the propensity score to account for the different sizes of the 2 exposure groups. (Walker et al., 2016)

Figure 59: Preference score distribution where the preference score is a transformation of the propensity score to account for the different sizes of the 2 exposure groups



Propensity score plot tells us that the cohorts are very different and there is very little overlap. For most people their treatment assignment was highly predictable. This means the data can reliably predict who will be prescribed acetaminophen or not. This reinforces the notion of channeling to the drug based on existing comorbidities/medications/treatments.

The region around the preference score value of 0.5 is where individuals are equally likely to receive a prescription of acetaminophen (which we define as clinical equipoise). Ideally, the region of thought to be in clinical equipoise, which is between 0.25 and 0.75 on the graph, would have the highest density of patients in both groups, or at least 50% of the patients. However, the large peaks of probability at either end of the plot show that this is very clearly not the case, since only 32% of the cohorts fall in the region of clinical equipoise. Instead, the plot shows that the 2 groups are very different. Thus, the potential for bias is quite high.

Note that the usual rule-of-thumb is that all covariates must have a standardized difference of mean ≤ 0.10 for us to consider 2 groups 'balanced'. There are 1,312 covariates that do not meet our rule-of-thumb for balance. All these unbalanced covariates have a positive standardized difference of the mean (except 'gender = MALE'), indicating that the high-use group is already 'sicker' at baseline on all these dimensions. For example, the high-users are more often exposed to antibiotics, diuretics, drugs for acid-related disorders, and antidepressants than non-users.

(These drug classes, rather than individual diagnosis codes, have the largest standardized difference of mean, likely because they represent entire disease areas).

8.4.3 Summary of Studies to Quantify Epidemiology Bias Using Study Designs in Literature

Our main objective was to quantify bias when using the study designs observed in literature, and to determine if any association observed for a given cancer outcome was outside the range of the systematic error observed using negative controls. We conducted case-control studies varying design features related to selection of controls, exposure to acetaminophen prior to the index date and the covariates controlled. For each variant outcome models were fit for 37 negative controls and 4 outcomes of interest. In each scenario the extent to which the negative controls varied from a RR of 1 was considerable and the number of statistically significant outcomes in each case was more than the 5% expected in unbiased studies. The RRs for the outcomes of interest fell within the range of the negative controls. This suggests that there is too much error to discern a statistically significant effect of the magnitude observed in the outcomes of interest here and in all the studies in the review.

The conclusions for the cohort studies were similar to the case-control studies. Two design feature variants were tested: including and excluding the first 2 years of follow-up time (exposures and outcomes). There was substantial error evidenced again by the number of statistically significant negative controls that exceeded 5%. The RRs for the outcomes of interest were within the range of error in the negative controls, as seen in the forest plots.

The results from this study reinforce what has been discussed above and seen in the published literature. The negative controls showed that despite the fact that these designs all attempt to adjust for confounding and other forms of bias, the extent of systematic error was substantial. The error is due to channeling bias, protopathic bias, and residual confounding and RRs for cancer outcomes of interest were within the range of the negative controls.

8.5 Scientific Accuracy and Completeness Issues Identified in the HID for Animal Carcinogenicity, Genetic Toxicology and Mode of Action Studies

There are a significant number of scientific accuracy and completeness issues in the HID that we have identified, and these are documented in the sections that follow. There are a number of examples in the HID where data and interpretations are framed in a manner that does not allow the reader to know whether they are re-analysis and interpretation by the authors of the HID or the results/conclusions of the original study authors themselves. Therefore, we request that the CIC please carefully review the scientific and quality issues to help in the evaluation of the data before making a decision.

8.5.1 Scientific Accuracy and Completeness Issues Identified in Carcinogenicity Studies in Animals

HID Specific Comments on Cited Mouse Carcinogenicity Studies

HID Assessment of Amo and Matsuyama (1985)

The HID noted that increased tumors were observed in the Amo and Matsuyama 1985 study in B6C3F1 mice (OEHHA, 2019): p. 121. Specifically, the HID noted that "a statistically significant increase in hepatocellular adenoma or carcinoma combined was observed in the high-dose group compared to controls (Amo and Matsuyama, 1985). In addition, a statistically significant increase in pituitary adenomas was observed in the high-dose group, with a significant dose-related trend (p = 0.01)" (OEHHA, 2019): p. 121.

Review of the HID Assessment of Amo and Matsuyama (1985)

The approach that the HID took in reaching their conclusion above is not scientifically valid and is not consistent with the conclusions of Amo and Matsuyama. The study authors conclude that "[t]he results of the present tests show that feeding the maximum tolerated dose of acetaminophen (0.6% diet) held no carcinogenic hazard for B6C3F1 mice" (Amo and Matsuyama, 1985). It appears that OEHHA performed an independent analysis of the data without accounting for the background incidences of these tumor types in B6C3F1 mice. In the Amo and Matsuyama (1985) study, the incidence of liver adenoma/carcinoma was 8/50 (16%), and pituitary adenomas was 9/50 (18%) in female mice at the highest dose tested. According to a review of the spontaneous neoplasm incidences in B6C3F1 mice in the 2-year carcinogenicity studies, "in untreated female B6C3F1 the most frequently occurring neoplasms were **liver adenoma/carcinoma (23.6%)**, malignant lymphoma (20.9%), and **pituitary gland adenoma/carcinoma (14.8%)**" (Haseman et al., 1998): p. 428). The reported tumor incidence in the high dose B6C3F1 female mice in Amo and Matsuyama (1985) are within background levels for both liver adenoma/carcinomas and pituitary adenomas. Further, there was an abnormally low background incidence of the tumor types in the female B6CF1 mice at the lower acetaminophen concentrations, and a high occurrence of liver adenoma/carcinomas in the male control mice in this study (13/43, see table below). In addition, in the only GLP guideline preclinical bioassay (NTP, 1993), no significant increase in tumors were observed at any dose tested in male or female B6C3F1 mice. These doses overlap with the doses tested in the Amo and Matsuyama (1985) study and a comparison of the liver tumor results in the two studies is summarized in Table 1. In male mice, there were ***decreases*** in the incidences of liver tumors, expressed as adenomas and carcinomas combined, at the high dose in both studies, and the decrease was statistically significant in the NTP (1993) cancer bioassay. In female mice, the small increase in the liver tumors at the high dose in the Amo and Matsuyama (1985) study was not observed in the NTP (1993) cancer bioassay, i.e., the incidence of liver tumors was virtually the same in the control and high dose groups (Table 25). The B6C3F1 mouse strain is highly

susceptible to liver tumors, and liver tumors are the most common type of tumor induced in B6C3F1 mice by exposures to test materials in NTP cancer bioassays. Moreover, there is no evidence that acetaminophen causes liver tumors in male or female B6C3F1 mice in the NTP cancer bioassay.

Table 25: Comparison of the incidence of liver adenoma and carcinoma combined among male and female B6C3F1 mice in Amo and Matsuyama (1985) and NTP (1993)

Study	Concentration of acetaminophen in the diet			
	0 ppm	600 ppm	3000 ppm	6000 ppm
Amo and Matsuyama (1985)				
Male mice	13/43 (30%)	ND	12/39 (31%)	6/45 (13%)
Female mice	2/49 (4%)	ND	2/46 (4%)	8/50 (16%)
NTP (1993)				
Male mice	16/50 (32%)	9/50 (18%)	10/50 (20%)	7/50 (14%)
Female mice	3/49 (6%)	4/50 (8%)	7/50 (14%)	3/49 (6%)

The HID also described a statistically significant increase in benign pituitary gland tumors in the females, but not the males, at the high dose in the Amo and Matsuyama (1985) study (based on its statistical re-evaluation of the pituitary tumor data) (Table 26). In comparison, in the NTP (1993) cancer bioassay, no difference in the incidence of pituitary gland adenomas was observed at the same high dose compared to controls among either male or female B6C3F1 mice (Table 26). Considered collectively, these data provide no clear or consistent evidence of an increase in benign tumors of the pituitary gland in male or female B6C3F1 mice.

Table 26: Comparison of the incidence of pituitary gland adenomas among male and female B6C3F1 mice in Amo and Matsuyama (1985) and NTP (1993)

Study	Concentration of acetaminophen in the diet			
	0 ppm	600 ppm	3000 ppm	6000 ppm
Amo and Matsuyama (1985)				
Male mice	0/43 (0%)	ND	1/39 (3%)	1/45 (2%)
Female mice	2/49 (4%)	ND	3/46 (7%)	9/50 (18%)
NTP (1993)				
Male mice	0/48 (0%)	0/39 (0%)	0/39 (0%)	0/46 (0%)
Female mice	14/46 (30%)	16/43 (37%)	7/42 (17%)	14/45 (31%)

In conclusion, there was no evidence of a carcinogenic hazard in this study and the approach that the HID took by performing statistical analyses without accounting for historical background control tumor incidence is not a scientifically valid approach.

HID Assessment of Flaks and Flaks (1983)

The HID noted that increased tumors were observed in the Flaks and Flaks 1983 study in IF mice (OEHHA, 2019): p. 113). Specifically, the HID noted that "[i]n the 18-month study of acetaminophen in IF strain male mice (Flaks and Flaks, 1983), statistically significant increases in hepatocellular adenoma, carcinoma, and adenoma and carcinoma combined were observed in the high dose group (500 mg/kg/day), with significant positive trends. Despite significant mortality in the high-dose group within the first 48 hours of the study, 87% of the surviving high-dose males developed liver tumors (20/23). In the 18-month study conducted in IF female mice (Flaks and Flaks, 1983), statistically significant increases in hepatocellular adenoma, and adenoma and carcinoma combined were observed in the high dose group (500 mg/kg/day), with significant positive trends" (OEHHA, 2019): p. 121).

Review of the HID Assessment of Flaks and Flaks (1983)

OEHHA presented the results of this study without providing critical information that would impact the relevance of these results to carcinogenicity hazard assessment for humans. In the Flaks and Flaks (1983) study there were only tumors present at chronic hepatotoxic and lethal doses (i.e. levels above the Maximum Tolerated Dose), which is not relevant to humans and would disqualify this as a valid test for carcinogenicity per accepted ICH, OECD and NTP Regulatory Guidances. Specifically, 37 out of 60 males, and 13 out of 60 female mice did not survive the 18-month study duration. According to the US EPA, "significant increases in mortality from effects other than cancer generally indicate that an adequate high dose has been exceeded" (EPA, 2005). The underlying reason for this guidance is that cytotoxicity can occur especially at doses that exceed the MTD. It is critical for the CIC to be aware that, as a general principle, the high dose in an animal carcinogenicity study should not exceed the MTD (EPA, 2005; FDA, 2010; ICH, 2011; NTP, 1993, 2011; OECD, 2012).

Given that the MTD was exceeded at the highest dose in this study, the observed responses at that dose should not be considered for acetaminophen hazard characterization. Of note, the lower administered dose (250 mg/kg) did not result in an increased incidence of any tumors. Furthermore, in the only GLP guideline preclinical bioassay (NTP, 1993), no significant increase in tumors were observed at any dose tested in male or female mice. In addition, there was no mortality at the highest dose tested in the GLP guideline preclinical bioassay. In addition, the IF mouse is not a strain typically recommended or used for carcinogenicity testing by any regulatory or scientific organization. In fact, a PubMed search did not identify any other published long-term

carcinogenicity study of any substance conducted in IF mice by these or any other investigators; we found no evidence of a historical control database for the IF mouse.

In conclusion, there was no evidence of a carcinogenic hazard in this study and the approach that OEHHA took by assessing carcinogenicity in mice administered acetaminophen at doses above the MTD is not scientifically valid.

HID Assessment of Weisburger et al. (1973)

According to the HID, an increase in tumors was observed in the Weisburger et al. (1973) study in NIH mice (OEHHA, 2019): p. 113). Specifically, OEHHA noted that "[i]n an 11-month study of acetaminophen in male Swiss mice, the incidence of hepatocellular adenomas and carcinomas (combined) was elevated, but not significantly different from controls in treated mice (control: 0/27 vs. treated: 3/26), as was the incidence of urinary bladder papillomas (control: 0/27 vs. treated: 2/20) (Weisburger et al., 1973). Interpretation of these findings is complicated by not only the short study duration, but also the fact that these data represent the combined observations from three experiments (rates of survival were low in each experiment, due to high levels of fighting-related mortality" (OEHHA, 2019).

Review of HID Comments

The primary focus of this study was to determine the effects of dietary acetaminophen and acetanilide (as competitive inhibitors of sulfation) on the incidence of tumors initiated by two known carcinogens, N-2-fluorenylacetamide (FAA) and N-hydroxy-2-fluorenylacetamide (N-OHFAA). This was not a study designed to determine the carcinogenic potential of acetaminophen, but a study to determine if acetaminophen administration may prevent the tumor promotion/initiation activity of known genotoxic carcinogens. Specifically, mice in this study were administered acetaminophen in their diet alone or were concomitantly administered one of two known carcinogens. According to reported study results, mice exposed to the known carcinogens (FAA and N-OHFAA) exhibited liver tumors, cysts, antecedent lesions, and tumors in the urinary bladder. The authors noted that the administration of acetaminophen had "no effect" on the mammary tumor incidence from N-OHFAA and did not alter bladder tumors induced by FAA in mice (Weisburger et al., 1973): p. 235). In fact, the study noted that acetaminophen administration decreased bladder tumor incidences cause by N-OHFAA in male mice. No liver tumors or urinary bladder tumors were observed in any female mice administered acetaminophen alone. As noted in the HID, the incidence of liver tumors (3/26), and urinary bladder tumors (2/26) in male mice administered acetaminophen are not statistically significant and are considered to be within the background rates for these tumors in NIH mice. The variability of liver and bladder tumors with AAF or N-OH-AAF by themselves, make evaluation of an acetaminophen effect essentially impossible. In addition, any tumor "promoter" generally produces tumors by itself, albeit at lower incidence and longer time than after pretreatment with

DNA reactive carcinogen like AAF (Cohen and Ellwein, 1991) and the absence of any effect with acetaminophen supports that it is neither a tumor initiator or promotor.

Furthermore, this study only used a single dose level of acetaminophen (11000 ppm in the diet), which is above the MTD for mice. Therefore, the non-statistically significant tumor findings in this study are not relevant to humans, and the observed responses at this dose should not be considered for acetaminophen hazard characterization. In conclusion, there was no evidence of a carcinogenic hazard in this study. The approach that the HID took by not discussing the negative results of the acetaminophen alone, not highlighting that this was a tumor promotion study (see (OEHHA, 2019) p. 121) and by assessing carcinogenicity in a tumor promoter study in which acetaminophen was administered at doses above the MTD is not scientifically appropriate or valid.

Other studies conducted in mice that were negative for carcinogenicity:

Several other studies have been conducted in mice that have **NO** acetaminophen treatment-related tumor findings following long-term dietary exposure. These include the following studies that are analyzed in the main response document:

- Wright (1967) – Note: this study was considered by the HID to be inadequate due to significant treatment related mortality, limited study design duration and reporting; top dose level is lower than the Weisburger study above.
- Hagiwara and Ward (1986)
- NTP (1993)

HID Specific Comments on Cited Preclinical Rat studies

HID Assessment of NTP (1993)

The HID noted that increased mononuclear cell leukemia (MCL) was observed in female Fischer 344 (F344)/N rats (OEHHA, 2019). Specifically, the HID noted that "[i]n the NTP two-year studies of acetaminophen (NTP, 1993), statistically significant increases in mononuclear cell leukemia (MNCL) were observed in female F344/N rats in the high-dose group compared to controls, with a positive dose-related trend. Among all females with MNCL, the proportion of animals dying before week 100 increased with dose (2/9 or 22% in controls; 4/17 or 24% in low-dose; 7/15 or 47% in mid-dose; 14/24 or 58% in high-dose group). In controls with MNCL, the leukemia was often observed only in the spleen and liver, with infrequent involvement of more than one additional organ, while in treated females with MNCL there was an increase in multiple organ involvement (defined as spleen and liver, plus two or more additional organs) [3/9 (33%) in controls; 16/17 (94%) in the low-dose; 12/15 (80%) in the mid-dose; 21/24 (88%) in the high-dose]. The control incidence of MNCL, 9/48 (18.8%), was similar to the laboratory historical control incidence of 16.5% (66/399; range 6–28%) and the historical control incidence reported for all NTP studies available at that time (20.8%, 425/2043; range 6–40%) (NTP, 1993)."

NTP (1993) noted the following about the MNCLs:

"On average, leukemias were detected one month earlier in the high-dose group than in the controls, suggesting a shortening of neoplasm latency. In addition, there was an increase in the extent of multiple organ involvement in the organ distribution of mononuclear cell leukemia in groups of exposed female rats compared to controls."

NTP concluded that there was "equivocal evidence" of carcinogenic activity in female rats, "based on increased incidences of MNCL." In reaching this conclusion, NTP noted the "generally high and variable background rate of this neoplasm in Fischer rats, and the lack of concordance of this study result with a lifetime study of acetaminophen in Fischer rats in Japan (NTP, 1993)" (OEHHA, 2019): p. 128).

Review of the HID Assessment of NTP (1993)

MCL, a distinct form of large granular lymphocyte leukemia (LGLL) is a cancer that occurs only in rats and essentially only in one strain, the Fischer F344 rat, is not relevant for humans, and has not been found in other rodent species (e.g. mice and hamsters) (Caldwell, 1999; Maronpot et al., 2016). The only potential human counterpart for LGLL is an extremely rare but aggressive leukemia that, unlike the F344 MCL, has a viral etiology (Caldwell, 1999; Maronpot et al., 2016). Thus, the evidence indicates that MCL, a spontaneous tumor that occurs at high incidence in aging F344 rats, is distinct from human large LGL and, therefore, MCL data should not be used in assessing potential human health hazards (Maronpot et al., 2016). This conclusion is based on an extensive review of the NTP experience by members of the NTP.

In addition, when evaluating the NTP 1993 study, the IARC Working Group noted "the high and variable incidence of mononuclear cell leukemia between and within studies with Fischer rats and considered that this was not a treatment-related effect" (IARC, 1999). In fact, in 2006, concern of a high background incidence of MCL in Fischer rats was one of the reasons why the NTP decided to discontinue the use of that strain for 2-year chronic toxicity and carcinogenicity bioassays (King-Herbert and Thayer, 2006; Maronpot et al., 2016).

Multiple researchers have noted that the high and variable background incidence of MCL in F344 rats impacts the ability to determine the relevance of potential treatment-related increases to human health risk (Caldwell, 1999; Lington et al., 1997; Thomas et al., 2007). It is important to note that the background incidence of MCL in male and female F344/N rats has increased over time. Specifically, the average MCL incidence in male and female F344 rats over time has been reported to be:

- 1970-1979: males 28.5%, females 19.6% (Haseman, 1983)
- 1977-1987: males 33.6%, females 20.2% (Haseman et al., 1990)
- 1980-1989: males 46.7%, females 26.8% (NTP, 1994)
- 1990-1996: males 50.5%, females 28.1% (Haseman et al., 1998)

At the time of the NTP study, the control incidence of MCL in female F344 rats used in 2-year cancer bioassays was 28.1%. In the 1993 NTP acetaminophen 2-year cancer bioassay, the reported control incidence of MCL was 9/50 (18%), which is substantially lower than the background incidence during this time period.

Of note, the HID stated that "statistically significant increases in mononuclear cell leukemia (MNCL) were observed in female F344/N rats, in the high-dose group compared to controls, with a positive dose-related trend" (OEHHA, 2019): p. 128). In the 1993 NTP study, the reported incidence of MCL in female rats was:

- Control: 9/50
- 600 PPM : 17/50
- 3000 PPM : 15/50
- 6000 PPM: 24/50

Based on the data, female F344 rats ingesting a diet of 3000 PPM (5 times higher than the next lowest dose, 600 PPM in the diet) had an MCL incidence of 15/50. This was a lower incidence than the group that received 600 PPM acetaminophen in the diet. As a result, based on the available data, it is unclear how one could conclude that "a positive dose-related trend" was observed in this study.

Furthermore, in the NTP studies, there were no statistically significant increases for any other tumor types in either mice or rats. It is important to note that in the only GLP guideline study, no statistically significant increase for liver carcinoma/adenoma, pituitary adenoma/carcinoma, or bladder tumors were observed. In addition, no bladder calculi were observed in any study aside from Flaks et al. (1985), described below.

HID Assessment of Flaks et al. (1985)

According to the HID, increased incidence of hepatic tumors were observed in male and female Leeds rats (OEHHA, 2019). Specifically, OEHHA noted that "[i]n 18-month feed studies of acetaminophen in Leeds rats, statistically significant increases in the incidences of hepatocellular adenomas were observed in both the male rat study, and the female rat study, with positive dose-related trends (Flaks et al., 1985). In males, statistically significant increases in urinary bladder transitional cell papilloma and transitional cell papilloma and carcinoma combined were also seen in the high-dose group, with positive dose-related trends. In females, a statistically significant increase in urinary bladder transitional cell papilloma and carcinoma combined was seen in the mid-dose group" (OEHHA, 2019): p. 128-129).

Review of HID Assessment of Flaks et al. (1985)

The HID did not mention a number of important points related to this study that would impact its relevance to human carcinogenicity potential of acetaminophen:

1. Flaks et al., (1985) observed statistically significant increases in benign, but not malignant, bladder tumors in males at the high dose (10,000 ppm) only and in females at the low dose (5000 ppm) only and in liver "neoplastic nodules" in both sexes at the high dose only. The CIC listing criteria specifically specifies reporting of increases in malignant tumors, not benign tumors (OEHHA, 2001). The HID states: "In males, statistically significant increases in urinary bladder transitional cell papilloma and transitional cell papilloma and carcinoma combined were also seen in the high-dose group, with positive dose-related trends" (OEHHA, 2019). However, it is important to recognize that the statistically significant increase in combined tumors was due to an increase in the benign tumors, not malignant tumors, since there was never more than a single male rat with bladder carcinoma in any dose group. The HID also states: "In females, a statistically significant increase in urinary bladder transitional cell papilloma and carcinoma combined was seen in the mid-dose group."(OEHHA, 2019). Actually, this sentence refers to the findings at the low dose, since there was no mid-dose group in this study, and once again, the statistically significant increase in combined tumors is attributable to benign, not malignant, tumors since there was never more than a single female rat with bladder carcinoma in any group.
2. Leeds rat is not a strain typically used or recommended for carcinogenicity testing by any regulatory or scientific organization. A PubMed search did not identify any other published long-term carcinogenicity study of any substance conducted in Leeds rats by any other investigators; we found no evidence of a historical control database for the Leeds rat.
3. In their assessment of Flaks et al. (1985), the IARC Working Group "noted that in the study in rats in which tumors were induced (Flaks et al., 1985) no tumors were found in either male or female controls, which is a highly unusual finding and raises questions about the interpretation of the findings" (IARC, 1999): p. 415). It is not appropriate to use historical control data from other strains of rats to aid in the evaluation of the findings from this study as the HID describes on page 129 of the document. The incidence of these findings is highly variable from strain to strain and among various testing facilities.
4. The terminology used in Flaks et al. (1985) to diagnose the liver tumors was "neoplastic nodule". This was a term that was used previously for lesions that were thought to be liver tumors but on further review were found to be either foci of cellular alteration, hepatocyte hyperplasia or hepatocellular adenoma. Foci of cellular alteration and hepatocyte hyperplasia are not neoplastic changes. Furthermore, it is not possible to determine from the published manuscript exactly what the authors were reporting. Due to the confusion in the presentation of proliferative hepatocellular tumors, the NTP held an expert panel review of lesions diagnosed as neoplastic nodule and published their recommendations (Maronpot et al., 1986). The major and significant suggested change was to replace the term neoplastic nodule with hepatocellular hyperplasia and hepatocellular adenoma, and then to re-evaluate the results and implications of shifting back to more conventional diagnostic terms.
5. Although the authors of the Flaks et al., (1985) publication, stated that the proliferative findings reported in the urinary bladder were not coincident with the presence of bladder calculi, this conclusion is of questionable accuracy. The lack of a strong correlation between calculi at necropsy and proliferative uroepithelial changes have been described many times and is due to several factors (Cohen et al., 2007), including spontaneous voiding of the calculi prior to termination and dissolution in the fixative. If the investigators did not examine for

calculi or crystalluria during the course of the study or at necropsy, they may have been missed. If looked for carefully, calculi or crystalluria would likely have been found in many more animals (Phang and Rinde, 1993). Additionally, the microphotograph that was included in the publication as a papilloma in the bladder was not a neoplasm, but rather urothelial papillary hyperplasia that is typical of the type of proliferative urothelial change seen secondary to bladder calculi and is a reversible lesion (Phang and Rinde, 1993; Shirai et al., 1986; Shirai et al., 1995). Most importantly, since it was reported by the authors that no tumors were found at any site in the control groups, the study findings should be regarded with caution. This is highly unusual, and without any historical control data to support that this is possible, all findings in the study should be suspect.

6. The IARC Working Group "noted that in the study in rats in which tumours were induced (Flaks et al., 1985) no tumours were found in either male or female controls, which is a highly unusual finding and raises questions about the interpretation of the findings" (IARC, 1999). The HID states that "other publications from the same laboratory corroborate the extremely low spontaneous incidence of liver and bladder neoplasms in Leeds rats" (OEHHA, 2019). However, the HID does not address the more important point that these investigators did not find any tumors in any tissues in any control group of male or female Leeds rats in the acetaminophen study or in any of their other carcinogenicity studies, which are identified in the HID. Flaks et al. (1985) did not find a single tumor in 40 control male and 40 control female Leeds rats in their acetaminophen study. The HID notes that no liver tumors were observed among 40 untreated male Leeds rats in an earlier study by Flaks et al. (1982); in fact, no tumors of any type were reported in the 40 negative control rats in this study. Finally, the HID states that no liver tumors were observed in untreated controls in a 20-month study in male Leeds rats (Flaks, 1978); once again, no tumors of any type were found among the control rats in this 1978 publication by Flaks. It appears that these investigators have never seen a tumor in any tissue or organ in a control group in any of their cancer studies using Leeds rats. This seems highly improbable and defies credibility.
7. Other limitations of the Flaks et al. (1985) study that were not mentioned by OEHHA include: limited description of methods, no description of the statistical methods, no randomized assignment of animals, no observation of clinical symptoms, no testing of diets to validate the concentration and stability of the test material, and infrequent (monthly) measurements of body weights.
8. None of the other carcinogenicity studies of acetaminophen, including the NTP cancer bioassay, reported an increase in bladder or liver tumors in rats

In summary, the HID did not address any of the key deficiencies for this highly-questionable study including that: (1) it does not meet the standard of "scientifically valid testing according to generally accepted principles," (2) reported increases in benign tumors only, and (3) is inconsistent with the results of three other carcinogenicity studies of acetaminophen in rats that did not observe increases in either bladder or liver tumors. Therefore, OEHHA should not rely on this study to characterize acetaminophen carcinogenicity.

Other Studies conducted in rats

Other studies have been conducted in rats that have no acetaminophen treatment-related tumor findings following long-term dietary exposure. These include the following studies (see main response document for detail):

- Hiraga and Fujii (1985)
- Johansson (1981a)

Summary

- In nearly all of the studies cited by the HID, there were no increases in tumors in any organ systems in the acetaminophen treated vs. control animals.
- Amo and Matsuyama (1985): The reported tumor incidence in the high dose B6C3F1 female mice in this study are within background levels for both liver adenoma/carcinomas and pituitary adenomas. The authors themselves note that "[t]he results of the present tests show that feeding the maximum tolerated dose of acetaminophen (0.6% diet) held no carcinogenic hazard for B6C3F1 mice" (Amo and Matsuyama, 1985): p. 572)
- Weisburger et al., (1973): This was not a study designed to determine the carcinogenic potential of acetaminophen, but a study to determine if acetaminophen administration may prevent the tumor promotion/initiation activity of known carcinogens. In addition, this study only used a single dose of acetaminophen (11000 ppm in the diet), which is above the MTD for mice. Therefore, the non-statistically significant tumor findings in this study are not relevant to humans, and the observed responses at this dose should not be considered for acetaminophen hazard characterization.
- Flaks and Flaks, (1983): In the 18-month carcinogenicity study, there was only an increase in tumors following administration of acetaminophen at chronic hepatotoxic doses that far exceed the MTD and therefore per ICH (OECD, NTP, and others) guidelines this is not considered an acceptable study and is not relevant to humans.
- NTP, (1993): In the 2-year cancer bioassay, there was an increase in mononuclear cell leukemia at that top dose in female rats only; this tumor type has a highly variable background incidence in the strain of rat used and is also not considered to be relevant to humans (Maronpot et al., 2016). There was no increase in tumors in male rats or male or female mice.
- Flaks et al., (1985): In the 18-month rat study, there was no dose dependence to the increase in the reported bladder tumors and the presence of calculi and are consistent with papillary urothelial hyperplasia and not tumors. Hepatic tumors only occurred following chronic dosing at high doses with evidence of hepatotoxicity indicating that the dose exceeded the MTD. In addition, there were no tumors reported in the control group, which is unprecedented in this type of study and calls into question the validity of the study.

8.5.2 Scientific Accuracy and Completeness Issues Identified in Genotoxicity Studies

8.5.2.1 Humans in vivo (p. 154-156)

The HID reported that:

- *"There are six publications reporting on genotoxicity studies of acetaminophen, conducted in different European populations (See Table 17). All the studies measured genotoxicity endpoints in peripheral blood lymphocytes (PBL); in addition, one study assessed effects in buccal mucosa cells. Endpoints measured in these studies included CAs, SCEs, MN, and UDS. In all but one set of studies individuals served as their own controls, with markers of genotoxicity assessed before and after treatment with acetaminophen. The study by Kirkland et al. (1992) used an age- and gender-matched placebo group as the comparator to the acetaminophen-treated group." (p. 154)*

The HID indicated that "Kirkland et al. (1992) used an age- and gender-matched placebo group as the comparator to the acetaminophen-treated group" implying that was the only comparison made and that the study did not compare pre-and post-dose samples for CA levels as was done in the other studies (OEHHA, 2019): p. 154). This is incorrect. Kirkland et al. (1992) compared pre- and post-dose as well as acetaminophen and placebo groups.

The HID noted that:

- *"As shown in Table 17, the ability of acetaminophen to induce CAs was assessed in PBLs of exposed humans in four studies, and the results were positive in two studies (Hongslo et al. 1991; Kocisova et al. 1988) and negative in the other two studies (Hantson et al. 1996; Kirkland et al. 1992)." (p. 155)*

However, further examination of Hongslo et al. (1991) and Kocisova et al. (1988) indicated that the HID did not report key information that does not support a positive result:

- Hongslo et al. (1991) administered acetaminophen (3 x 1g during 8 hrs) to 9 volunteers and reported a small (from 2.38% pre-dose to 5.03% 24 hrs after the first dose) but insignificant ($p<0.1$) increases in the proportion of cells with CA, including gaps. When gaps were excluded (as is normal practice) the increase was much smaller from 2.16% to 3.43% (this was not analyzed for statistical significance). Excluding gaps, the increase was primarily due to a 6-fold increase in chromatid breaks (i.e. similar to the observations of Kocisová et al., (1988), although no blood samples were taken at later sampling times). As in the Kocisová et al. (1988) study, not all volunteers showed an increase in the levels of aberrant cells, excluding gaps (7/9 volunteers showed an increase but 2/9 showed a decrease).
- Kocisová et al. (1988) reported 2 studies. In the first study, acetaminophen was administered (3 x 1g during 8 hrs) to 11 volunteers (3 males/8 females), and a small but statistically significant ($p<0.05$) increase (from 1.68% pre-dose to 2.77% at 24 hrs after the first dose) in the proportion of cells with CA (excluding gaps) was observed. However, CA frequencies were

not significantly different from pre-dose levels at later sampling times (72 or 168 hrs), and had returned to below pre-dose levels by 168 hrs. Thus, the increase in the proportion of cells with CA was transient, which is unusual since in other longitudinal studies CA levels tend to remain increased for periods of weeks or months (Kucerova et al., 1980; Schmid et al., 1985). In the same publication a second study with the same volunteers was performed 1 week later with the same dosing schedule, except that each dose of acetaminophen was given together with 1 g of the anti-oxidant, ascorbic acid. A small but statistically significant ($p<0.05$) increase (from 1.09% pre-dose to 2.22% 72 hrs after the first dose) in the proportion of cells with CA was observed. CA levels were not significantly different from pre-dose at 24 or 168 hrs, so again the increase in the proportion of cells with CA was transient. Whilst the appearance of significant levels of CA at a single sampling time is not unusual, the fact that the peak of CA frequencies was at different times in the 2 studies with the same volunteers is unexplained, and suggests the increases may be due to chance. It is unclear whether the co-administration of ascorbic acid delayed the appearance of CA, or whether this was due to chance. It should be noted that in both studies the increased CA levels were due entirely to chromatid breaks; there were no increases in chromosome breaks or exchanges. It was most interesting that the individual responses of the volunteers in the first and second studies showed that 7 and 6, respectively, of the 11 volunteers showed an *increase* in the number of aberrant cells, whereas 4 and 5 volunteers, respectively, showed *no increase or a decrease* in the numbers of aberrant cells. Since the same volunteers were used in both studies, it was possible to see that no specific sub-group of the volunteers showed a consistent response (i.e. those that showed increased CA levels with acetaminophen alone were not the same as those showing increased CA levels with acetaminophen plus ascorbic acid). On the contrary, it was apparent that those individuals who had shown a comparatively large increase in chromatid break frequency in the first study showed a small increase or even a decrease in the second study, and vice versa. It is therefore highly implausible that the increased CA levels in these 2 studies resulted from the genotoxic effects of acetaminophen, and it is more likely they were due to chance.

The HID indicated that

- "*Acetaminophen induced SCEs in PBL in one study (Hongslo et al. 1991) and had no effect in another study (Kirkland et al. 1992)*" and that "*[i]t is possible that Kirkland et al. (1992) had a reduced ability to detect acetaminophen-related effects on PBL CAs and SCEs due to inter-individual variability between the placebo and acetaminophen- treated groups in "baseline" levels of these markers of clastogenicity." (HID: p. 155).*

However, the HID is incorrect that Kirkland et al. (1992) examined SCEs. Furthermore, the HID is also incorrect that Kirkland's results on clastogenicity (i.e., negative CA results) were due to variability between the placebo and acetaminophen- treated groups. On the contrary, the study

by Kirkland et al. (1992) examined both pre- and post-dose as well as acetaminophen and placebo groups and found no increase in CA induction for either comparison group.

Additionally, the study by Hongslo et al. (1991) administered 1 g acetaminophen three times over eight hours to human volunteers and the number of induced SCEs/chromosome was evaluated in the volunteer's lymphocytes before the treatment and 16 hours after the treatment (Hongslo et al., 1991). 0.19 SCEs/chromosome (range of 0.144-0.240 SCEs/chromosome) were observed before treatment and 0.21 SCEs/chromosome (range of 0.159 to 0.244) was reported after treatment (Hongslo et al., 1991). While the authors considered this difference to be significant, whether this response is biologically relevant is questionable (Hongslo et al., 1991). The SCE assay (OECD, 1986a) was deleted as a test guideline due to a poor understanding of the mechanisms of action that can be detected by the test and a high false positive rate (OECD, 2017). Additionally, there were alternative and more reliable assays used for determining clastogenic potential, such as the micronucleus test. Therefore, positive responses reported by the SCE assay should be interpreted with caution and more weight should be given to reliable guideline assays such the micronucleus test or CA assay.

The HID indicated that:

- *"[A]cetaminophen was shown to induce MN in human PBLs (Kocisova and Sram 1990) and buccal mucosa cells (Kocisova and Sram 1990; Topinka et al. 1989)." (p. 155)*

This is incorrect, the study found no significant increase in MN in human PBLs. Additionally, while the HID noted that acetaminophen induced MN in human buccal mucosa cells, they did not present results for the full-time course. MN induction was 0.19% pre-acetaminophen exposure and 0.23% (NS), 0.38% (Sig.), and 0.23% (NS) after 24, 72, and 168 hours post exposure. While it is not unusual for MN frequency to increase at a single sampling time, effects in a site-of-contact tissue would be expected at the first sampling time (24 hours) and not at the mid-sampling time (72 hours). The biological relevance of this response is therefore difficult to interpret.

The HID noted that:

- *"Topinka et al. (1989) also reported that acetaminophen decreased UDS in PBLs. These authors noted that acetaminophen has been shown to interfere with nucleotide excision repair in several mammalian cell types (Brunborg et al. 1995; Hongslo et al. 1993), and suggested that the decrease in UDS observed following acetaminophen treatment was the result of reduced DNA excision repair activity." (HID: p. 155).*

Topinka et al. (1989) observed a slight decrease in UDS after 24 hours but the levels returned to control levels after 72 and 168 hours demonstrating that this small effect was transient in nature. The authors state that the effect acetaminophen had on 1-methyl-3-nitro-1-nitroso-guanidine (MNNG) induced UDS in human peripheral lymphocytes from human volunteers was studied; human volunteers were administered 1 g acetaminophen three times over 8-hours (Šrám et al.,

1990; Topinka et al., 1989). Acetaminophen did not increase UDS induced by MNNG; rather, UDS induced by MNNG was decreased (Šrám et al., 1990; Topinka et al., 1989).

8.5.2.2 Human cells in vitro (p. 157-158)

The HID noted that:

- *"Acetaminophen induced DNA strand breaks in a human hepatocellular carcinoma cell line, as measured by the comet assay and by γ-H2AX staining (Bandi et al. 2014), and in liver slices, as measured by the comet assay (Jetten et al. 2014)" while "[a]cetaminophen did not induce DNA single strand breaks in cultured human skin fibroblasts in the presence of sheep seminal vesicle microsomes (Andersson et al. 1982)." (p. 157).*

However, the HID did not critically evaluate whether each study examined the effects of cytotoxicity and whether the genotoxic result was confounded by toxicity in the test system. The International Conference on Harmonization (ICH) Guidance on Genotoxicity Testing and Data Interpretation for Pharmaceuticals Intended for Human Use S2(R1) recommends that the tested doses for mammalian cell assays should not cause a greater than 50% reduction in cell growth (ICH, 2011): p. 6, 16). As noted in ICH S2(R1), genotoxic hazard identification should be carefully evaluated alongside cytotoxic effects as cellular toxicity can confound positive responses in DNA and chromosomal tests (ICH, 2011). For example, ICH 2011 indicated that "[a]s cytotoxicity increases, mechanisms other than direct DNA damage by a compound or its metabolites can lead to 'positive' results that are related to cytotoxicity and not genotoxicity" (ICH, 2011): p. 16). It was further noted that "[s]uch indirect induction of DNA damage secondary to damage to non-DNA targets is more likely to occur above a certain concentration threshold" and '[t]he disruption of cellular processes is not expected to occur at lower, pharmacologically relevant concentrations" (ICH, 2011): p. 16). DNA damaging agents are typically detected under conditions where there is only moderate levels of toxicity and even weak clastogens display positive results without exceeding 50% reduction in cell growth (ICH, 2011). Thus, the ICH recommendation of selecting the top dose that does not cause significant cytotoxicity (or 50% reduction in cell growth) in a DNA damage or cytogenetic assay should be considered during hazard evaluation. For example, Bandi et al. (2014) observed DNA damage at a cytotoxic dose, therefore these results do not represent an intrinsic genotoxic hazard but rather the response is an artifact of toxicity in the test system. Additionally, while Jetten et al. (2014) observed DNA strand breaks as measured by the comet assay, the biological significance of their measured response was unclear. As noted by the HID, "the authors did not report the doses used, instead reporting the "average BMD" observed among liver slices from five individuals; BMDs varied by 64-fold between individuals" (OEHHA, 2019), p. 158). Additionally, it was unclear whether cytotoxicity effected these results as the authors measured LDH release using a similar BMD approach with even greater variability (400-fold among individuals).

The HID noted that:

- *"In studies with human granulocytes stimulated to undergo the "respiratory burst" by treatment with phorbol myristate acetate (PMA) for 30 minutes, 14C-labelled acetaminophen was incorporated into cellular DNA and RNA, indicating the formation of DNA and RNA adducts (Corbett et al., 1989)" (HID p. 157).*

However, these conditions were likely cytotoxic as the authors utilized 100 ng/mL of PMA to stimulate the granulocytes which is well above concentrations shown to be cytotoxic (30 ng/mL) (Corbett and Corbett, 1988; Saito et al., 2005). Therefore, these results have limited biological relevance towards understanding the genotoxic potential of acetaminophen.

The HID noted that:

> *"Acetaminophen inhibits ribonucleotide reductase activity (Hongslo et al. 1991), and so its ability to impair nucleotide excision repair in human cells was investigated by Hongslo et al. (1993) and Brunborg et al. (1995). In several different human cell types, acetaminophen was found to delay the repair of single strand DNA breaks (SSBs) induced by treatment with either UV light (mononuclear blood cells, T lymphocytes, B lymphocytes, monocytes, HL-60 cells, fibroblasts) or 4-nitroquinoline n-oxide (NQO) (mononuclear blood cells). In these studies, the effect of acetaminophen on the repair of SSBs was abrogated by the addition of deoxyribonucleotides to the cell medium. Hongslo et al. (1993) and Brunborg et al. (1995) concluded that acetaminophen's ability to delay the repair of SSBs in these studies was the result of impaired nucleotide excision repair due to acetaminophen's inhibition of ribonucleotide reductase." (p. 157)*

In these studies, the reduced thymidine update has a clear threshold (i.e. only seen at supratherapeutic exposures) and is transient, reversing *in vivo* within 2 to 4 hours (Hongslo et al., 1994; Lister and McLean, 1997). There is no evidence that these effects are truly a consequence of effects on DNA repair and not a consequence of reduced cell turnover. In addition, there is no evidence that the effects are sustained with multiple dosing at therapeutic or non-toxic supratherapeutic doses and lead to sustained DNA effects at non-toxic concentrations. Additionally, the influence of cytotoxicity on ribonucleotide reductase activity or DNA repair was not examined in Hongslo et al. (1993) and Brunborg et al. (1995).

The HID noted that

- *"In one study in human PBLs, incubation with acetaminophen resulted in a slight increase in UDS (Binkova et al. 1990)" (HID: p. 157).*

However, the HID did not discuss the relevance of this positive result. For example, the *in vitro* DNA Damage and Repair/Unscheduled DNA synthesis assay (OECD, 1986b) was deleted/recalled by OECD in April 2014. OECD 482 was deleted as a test guideline due to a lack of use of the test in various legislative jurisdictions and due to the availability of other tests that showed a better performance for detecting genotoxicity (OECD, 2015). Additionally, Binkova et al. (1990) studied UDS by scintillation counting, which is not a recommended method. Further, the HID did not

report that increased UDS measured by an increased uptake of 3H-thymidine by scintillation counting could be due to changes in the rate of replication and not due to repair.

The HID noted that:

- *"Chromosomal effects of acetaminophen in human cells exposed in vitro have been observed at concentrations ranging from 1 to 1.5 mM. These effects include induction of MN in human amniotic fluid cells (Simko et al. 1998), weak induction of MN in human PBLs (Ibrulj et al. 2007), and induction of CAs and SCEs in human PBLs (Hongslo et al. 1991; Ibrulj et al. 2007; Watanabe 1982)" (p. 157).*

However, the HID did not critically evaluate the methodology utilized by these studies. For example, Simko et al. (1998) reported a higher level of MN than would be expected in the control cells and did not report whether the slides were coded prior to scoring which in turn could lead to bias when scoring. The HID reported a weakly positive response for micronuclei formation in peripheral blood lymphocytes as studied by Ibrulji et al. (2007). However, there was no statistically significant increase in the micronuclei formation. While Ibrulj et al. (2007) confirmed the ability of acetaminophen to induce chromosomal aberrations in cultured human lymphocytes, exposed continuously for 72 hrs, the dose was likely to be a cytotoxic concentration (above ICH guidance threshold - see below) of 200 µg/mL (1.3 mM), whereas negative results were obtained at 50 and 100 µg/mL. Although cytotoxicity would be expected at concentrations >1 mM, the effect on nuclear division index was small (in the region of 20% at 200 µg/mL). The chromosomal aberration results are similar to those reported by Honglso et al. (1991) in human lymphocytes exposed to acetaminophen for the last 24 hrs of a 72-hr incubation. It is important to note that almost all induced aberrations in both studies were chromatid breaks. Induction of chromosome breaks will lead to cell death. This means that chromosomal changes that pre-dispose to indicate a mutagenic or carcinogenic hazard would need to be induced at low levels of cytotoxicity, such that affected cells would survive, and would involve induction of stable chromosome rearrangements rather than (or as well as) breaks. There was no evidence of induction of unstable chromosome rearrangements, which might be indicative of a potential to form stable rearrangements. Additionally, Honglso et al. (1991) reported the number of gaps in the total aberration count. Gaps are achromatic lesions that are smaller than the width of one chromatid with minor misalignment of the chromatids (Registre and Proudlock, 2016), their biological relevance is unclear, and therefore chromosomal damage is conventionally reported "excluding gaps". Identification of gaps in samples may vary between laboratories due to differences in identification criteria, slide scoring, and variability in chromatid width due to condensation. Thus, while it is possible that in a small number of cases, aberrations identified as gaps may be breaks within a single chromosome, they are generally not considered relevant for chromosome aberration assessment. Therefore, any gaps in a chromosome or chromatid structure are recorded, but not included in a genotoxicity assessment (Registre and Proudlock, 2016). This is consistent with OECD guidelines for both *in vitro* and *in vivo* chromosomal

aberration tests which recommend excluding gaps in the frequency analysis of chromosomal aberrations (OECD, 2016a, c).

Additionally, doses that caused a positive response in Watanabe et al. (1982) also caused cytotoxic effects which could impact the observed genotoxic response. In addition, the effects of cytotoxicity were not reported in Hongslo et al (1991) and, hence, it is unknown whether the positive response observed occurred at a potentially cytotoxic dose that could impact the observed genotoxic response, as described previously. Further, the positive responses reported in Hongslo et al. (1991), Ibrulj et al. (2007), Simko et al. (1998), and Watanabe (1982) were all above the ICH recommended dose of 1 mM. ICH recommends a maximum concentration of 1 mM or 0.5 mg/ml, whichever is lower, to be tested in mammalian cell assays, when not limited by solubility (ICH, 2011). It was indicated that the "limit of 1 mM maintain[ed] the element of hazard identification, being higher than clinical exposures to known pharmaceuticals, including those that concentrate in tissues…, and [was] also higher than the levels generally achievable in preclinical studies *in vivo*" (ICH, 2011): p. 16). In studies of acetaminophen, concentrations of 1 mM and above have generally shown severe cytotoxicity and a reduction in cell number greater than 50% (Holme and Søderlund, 1986; Hongslo et al., 1990; Hongslo et al., 1988; Muller et al., 1991; NTP, 1993; Patierno et al., 1989; Sasaki et al., 1983).

In addition, the HID reported results from deleted OECD guidelines without discussing the reliability of these assays (Hongslo et al., 1991). For example, the OECD *in vitro* SCE guideline test (OECD, 1986a) was deleted in April 2014 due to a poor understanding of the mechanisms of action that can be detected by the test or their biological relevance (OECD, 2017). Further, there were alternative and more reliable assays used for determining clastogenic potential, such as the chromosomal aberration test, mouse lymphoma assay, comet assay, or micronucleus test. Thus, the positive induction of SCEs as reported by Hongslo et al. (1991) should be given negligible weight when evaluating the genotoxicity potential of acetaminophen (Hongslo et al., 1991).

8.5.2.3 Animals in vivo (p. 159-162)

The HID noted that:

- *"Acetaminophen was found to form DNA adducts in liver and kidney of mice exposed via i.p. injection in two studies (Hongslo et al. 1994; Rogers et al. 1997), and a third i.p. study in mice also reported DNA adduct formation in liver (Dybing et al. 1984). No DNA adducts were detected in two studies in rats exposed via the oral route (Dybing et al. 1984; Hasegawa et al. 1988; Hongslo and Holme 1994; Rogers et al. 1997; Williams et al. 2007)." (p. 159)*

Numerous studies examined adduct formation utilizing tritiated (^{3}H) acetaminophen for radiolabel detection (Dybing et al., 1984; Hongslo et al., 1994; Rogers et al., 1997). However, none of these studies controlled for background rates of free tritiated compound under *in vivo* conditions. It has been noted that "[t]ritiated compounds almost inevitably lead to the formation

of tritiated water of which a tritium ion can quite efficiently be incorporated into newly synthesized DNA" and "[i]n order to account for this incorporation the specific activity of the body water must be known and a comparison with control experiments with tritiated water will provide an estimate on that part of the radioactivity of DNA that is due to tritiated water" (Lutz and Schlatter, 1979): p. 299). Lutz et al. found that oral doses of about 10 mCi tritiated water per kg rat resulted in incorporation of radioactivity into DNA from the liver which increased linearly with time (Lutz and Schlatter, 1979). Furthermore, the low radioactive signal associated with DNA in Dybing et al. (1984) occurred at hepatoxic doses. The authors noted that "it is important to note that the covalent binding of paracetamol was demonstrated at a hepatotoxic dose. If this covalent binding only occurred in cells which would later die, such a DNA interaction would not lead to mutation, an event which most probably is involved in initiation of carcinogenesis" (Dybing et al., 1984): p. 29). Consistent with the notion that the aforementioned results may be due to free tritiated compound, (Rogers et al., 1997) also examined DNA adduct formation in mice using ^{32}P-postlabeling and no differences were observed for acetaminophen treated mice compared to control mice. This result was in contrast to the observed result that tritiated compound was observed in hepatic and renal tissue at all doses tested. Further, other studies that examined acetaminophen adduct formation with ^{32}P-postlabeling have demonstrated that acetaminophen does not form DNA adducts under *in vivo* conditions (Hasegawa et al., 1988; Rogers et al., 1997; Williams et al., 2007).

The HID noted that:

- *"An increase in serum levels of 8-hydroxy-2'-deoxyguanosine (8-OHdG, or 8-oxodG), a marker for oxidative DNA damage, was observed in Kunming mice administered acetaminophen by the oral route for 10 weeks (Wang et al. 2015)." (p. 159)*

Wang et al. (2015) observed hepatotoxicity in mice which were administered 400 mg/kg acetaminophen by oral gavage. Therefore, the observation of 8-hydroxy-2'-deoxyguanosine was likely confounded by this toxic response.

The HID noted that:

- *"DNA strand breaks were detected in the livers of acetaminophen-treated male B6 mice and ICR mice after single i.p. injections of 600 mg/kg bw or 300 mg/kg bw acetaminophen, respectively (Hongslo et al. 1994; Oshida et al. 2008). DNA strand breaks were not detected in the kidney or bone marrow in these studies (Hongslo et al. 1994; Oshida et al. 2008). In addition, DNA strand breaks were not detected in the liver or kidney of acetaminophen-treated male Wistar rats after a single i.p. injection of 600 mg/kg bw acetaminophen (Hongslo et al. 1994)." (p. 159)*

While a lowest effective dose (LED) of 600 mg/kg was reported based on DNA single strand breaks induced in liver cells of mice in Hongslo et al. (1994), the HID did not report whether this dose caused a hepatotoxic response. Additionally, the effect was transient in nature. Similarly, Oshida

et al. (2008) reported that 300 mg/kg bw dose of acetaminophen induced a hepatotoxic response in mice. It is notable that there are marked species differences in acetaminophen-induced hepatotoxicity (Davis et al., 1974), with mice being much more sensitive than rats. The oral LD_{50} in mice is 338 mg/kg, whereas in rats it is 1944 mg/kg. Thus, doses which far exceed the LD_{50} in mice cause only minimal necrosis in rat liver (McGill et al., 2012b). These differences are due to differences in the rate of metabolism of acetaminophen to NAPQI (Blair et al., 1980; Tee et al., 1987) and mitochondrial dysfunction (McGill et al., 2012b). The relative sensitivity of freshly isolated hepatocytes from mouse, rat and hamster reflected the hepatotoxicity seen *in vivo*, but by contrast human hepatocytes were relatively resistant to the cytotoxicity of acetaminophen (Tee et al., 1987). Thus, toxic effects (and any genotoxicity resulting from such toxicity) would be expected at lower doses in mice than in rats or humans.

The HID noted that:

- "*Acetaminophen can cause impairment of nucleotide excision repair in rodents in vivo. Hongslo et al. (1994) showed that NQO-induced DNA-repair synthesis was decreased in the liver, spleen, and kidney of male B6 mice and Wistar rats exposed to acetaminophen via i.p. injection 5 minutes before treatment with NQO (mice, 50 mg/kg; rats, 20 mg/kg). Similar to what was observed in in vitro studies with human cells (Table 18); acetaminophen increased SSBs and delayed the repair of SSBs in livers, spleens and kidneys of NQO-treated mice and rats. The authors concluded that these effects were the result of impaired nucleotide excision repair due to acetaminophen's inhibition of ribonucleotide reductase.*" (p. 159)

It has been proposed that these effects may be a result of the inhibition of ribonucleotide reductase and may explain genotoxicity effects seen at high doses (Bergman et al., 1996; Thybaud et al., 2007), but this has not been definitively demonstrated. In these studies, the reduced thymidine update has a clear threshold (i.e. only seen at supratherapeutic exposures) and is transient, reversing *in vivo* within 2 to 4 hours (Hongslo et al., 1994; Lister and McLean, 1997). There is no evidence that these effects are truly a consequence of effects on DNA repair and not a consequence of reduced cell turnover. In addition, there is no evidence that the effects are sustained with multiple dosing at therapeutic or non-toxic supratherapeutic doses and lead to sustained DNA effects at non-toxic concentrations. When viewed in the context of the negative carcinogenicity studies, the data support that this mechanism does not represent a genotoxic or carcinogenic hazard to humans.

The HID noted that:

> "*In mouse studies, acetaminophen tested positive in several chromosomal damage assays (e.g., MN, CAs and SCEs) in two strains, BALB/c and Swiss, via multiple administration routes. Increases in MN were observed in the peripheral blood cells of BALB/c mice exposed to acetaminophen via i.p. injection or in utero (Markovic et al. 2013). Increases in MN were also observed in the bone marrow cells of Swiss mice exposed to*

acetaminophen via i.p. injection (Sicardi et al. 1991). In studies of NMRI mice administered acetaminophen via gavage or i.p. injection, no increase in MN was observed in the bone marrow (King et al. 1979). Increases in CAs were observed by three different research groups in the bone marrow of Swiss mice treated with acetaminophen either orally or via i.p. injection (Giri et al. 1992; Reddy 1984; Severin and Beleuta 1995). The CAs induced by acetaminophen in mouse bone marrow included gaps, chromatid breaks, acentric fragments, and polyploid metaphases. These types of structural CAs were not statistically significantly increased in the testes of Swiss mice exposed to acetaminophen via the oral route, although other chromosomal abnormalities were observed in the testes, such as polyploidy (Reddy and Subramanyam 1985). A dose-dependent increase in SCEs was observed in the bone marrow of Swiss mice treated with acetaminophen via i.p. injection (Giri et al. 1992)." (p. 160)

The HID did not report key information for these studies. For example, the HID reported a weakly positive result for induction of micronuclei in pregnant BALB/c mice exposed to acetaminophen intraperitoneally at 60 mg/kg on days 12 and 14 of pregnancy and a positive result for offspring exposed *in utero*. For each of the micronucleus assays 1000 acridine orange-stained reticulocytes per animal were assessed. It should be noted that this is a much smaller population of cells than is currently recommended in OECD guidelines. Importantly, it is not stated that the slides were "blinded" before scoring, and therefore scorer bias cannot be excluded. Micronucleus frequencies in the dams treated with acetaminophen were increased slightly (3.25-fold) above vehicle control frequencies at 48 hrs after dosing, but were not significantly different. On the other hand, micronucleus frequencies in the blood of the pups showed a smaller increase (2.28-fold) above vehicle controls, but this was statistically significant ($p<0.05$). Glutathione peroxidase activity in the hemolysate of the new-born pups, and malondialdehyde levels in the livers of the pups, were significantly lower than in vehicle control pups. The authors speculate that the reduction of glutathione peroxidase reflected systemic oxidative stress, which is known to occur with acetaminophen treatment, while the reduction of malondialdehyde in the liver can be interpreted as an unspecific reaction to drug treatment that might have cytotoxic, and in particular hepatotoxic, effects associated with oxidative stress and lipid peroxidation. Given that mice are more sensitive than rats or humans to the hepatotoxic effects of acetaminophen, that the increases in micronucleus frequency in the dams were higher than in pups, yet were not statistically significant, and that the slides were probably not "blinded" before scoring, these results should be viewed with caution.

As discussed previously, there are marked species differences in acetaminophen-induced hepatotoxicity (Davis et al., 1974), with mice being much more sensitive than rats. Thus, toxic effects (and any genotoxicity resulting from such toxicity) would be expected at lower doses in mice than in rats or humans. As mentioned above, the oral LD_{50} in mice is 338 mg/kg, whereas in rats it is 1944 mg/kg. Thus, it is crucial that studies evaluate hepatotoxicity during the study to ensure that observed genotoxic results are not potentially due to the hepatotoxic response. As such, hepatotoxicity was not evaluated in several studies, thus, it is unknown whether these

responses are confounded by hepatotoxicity (Giri et al., 1992; Reddy, 1984; Severin and Beleuta, 1995; Sicardi et al., 1991). Specifically, Severin and Beleuta et al. (1995) reported a lowest effective dose that was greater than the LD_{50}.

Several studies administered acetaminophen via the i.p. route, and the HID did not discuss the relevance of these studies (Giri et al., 1992; Markovic et al., 2013; Severin and Beleuta, 1995; Sicardi et al., 1991). According to ICH guidelines and OECD guidelines, the route of administration should be the anticipated route of human or clinical route (ICH, 2011; OECD, 2016b, c, d). In the case of the acetaminophen, the expected clinical route is oral or intravenous administration. While intraperitoneal (i.p.) injection has been used to deliver large bolus doses of acetaminophen in numerous rodent studies (Giri et al., 1992; Hongslo et al., 1994; King et al., 1979; Markovic et al., 2013; Oshida et al., 2008; Severin and Beleuta, 1995; Sicardi et al., 1991), according to OECD 474, 475, and 489 guidelines, i.p. is generally not recommended for testing since it is not a typical relevant route of human exposure (OECD, 2016b, c, d, 2017). For example, OECD 489 guideline for In Vivo Mammalian Alkaline Comet Assay stated that "[i]ntraperitoneal injection is generally not recommended since it is not a typical relevant route of human exposure, and should only be used with specific justification (e.g. some positive control substances, for investigative purposes, or for some drugs that are administered by the intraperitoneal route)" (OECD, 2016b) p. 10). One major challenge with correlating results from i.p. exposure compared to oral or intravenous routes is likely due to differences in pharmacokinetic considerations and associated toxicity. For example, i.p. administration of acetaminophen was shown to cause hepatic and renal toxicity at doses that were non-toxic when administered orally, likely due to a higher dose and rate of acetaminophen delivery to the liver when administered via i.p. (Colin et al., 1986). Therefore, careful consideration should be given to data generated with *in vivo* studies that administered acetaminophen via i.p. injection.

As specified by the HID, "[t]he CAs induced by acetaminophen in mouse bone marrow included gaps …" (p. 160). As described above, gaps should be excluded in the frequency analysis of chromosomal aberrations. If gaps were excluded, the positive results reported for Reddy (1984) and Severin and Beleuta (1995) would be less evident. Increases in polyploidy and any evidence of other forms on aneuploidy are not necessarily considered indicative of genotoxicity (Registre and Proudlock, 2016). Thus, the positive results reported for Tsuruzaki et al. (1982) and the discussion on the positive results reported in Reddy and Subramanyam (1985) are not relevant.

The HID noted that:

> *"Chromosomal damage has also been observed in rats exposed to acetaminophen. In one oral study acetaminophen increased MN formation in the bone marrow of treated rats (Hazleton Microtest 1993, as cited by Muller and Kasper 1995). In another oral study, administration of acetaminophen to female SD rats for two weeks prior to mating and continuing through the first 11 days after mating resulted in an increase in chromosomal aneuploidy in the embryos of exposed rats, compared to controls (Muller and Kasper 1995;*

> *Tsuruzaki et al. 1982). Tsuruzaki et al. (1982) reported that the chromosomal karyotypes of the affected embryos were all mosaics, consisting of monosomy/normal or trisomy/normal cells." (p. 160)*

The HID did not report key information in these studies. For example, it was not mentioned in the HID that the Hazleton Microtest 1993 study reported that the dose administered to the rats caused a decrease in the PCE/NCE ratios which indicated a severe cytotoxic effect (Bergman et al. 1996; Marshall 1993 also the same as Hazleton Microtest (1993). It should be noted that the increase in aneuploidy in rat embryos observed by Tsuruzaki et al. (1982) was not dose dependent. In addition, there was no increase in the cells with structural chromosome abnormalities and the effects of potential hepatotoxicity are unknown as this study was in a foreign language (Tsuruzaki et al., 1982).

8.5.2.4 Animals in vitro (p. 163-167)

The HID noted that:

- "*Acetaminophen increased gene mutations in mouse lymphoma cells (Muller and Kasper 1995; Sasaki 1986; Shimane 1985), and induced small, dose-dependent increases in mutations associated with ouabain and 6-thioguanine (6TG) resistance in Chinese hamster lung V79 cells (Shimane 1985). Acetaminophen did not induce mutations in Chinese hamster ovary K1 (CHO-K1) cells (Sasaki 1986) or C3H/10T1/2 Clone 8 mouse embryo cells (Patierno et al. 1989)" (p. 163).*

Upon further examination of Sasaki (1986) and Shimane (1985), it is evident that the HID did not report key information of these studies. For example, neither Sasaki (1980) nor Shimane (1985) tested the genotoxic effects of acetaminophen on mouse lymphoma cells. The only study that used mouse lymphoma cells was (Clements, 1992) (referred by HID as Hazleton Microtest (1992), cited Muller and Kasper (1995)). Additionally, this study reported that acetaminophen was positive in a mouse lymphoma TK fluctuation assay without exogenous metabolic activation and negative with metabolic activation with rat liver S9 mix at concentrations of 3.3 to 33 mM (i.e. above the ICH recommended limit) (Bergman et al., 1996; Hazleton Microtest, 1992). Additionally, no conclusions could be drawn on the type of damage that acetaminophen caused since the size of the mutant colonies was reportedly not analyzed (Bergman et al., 1996; Hazleton Microtest, 1992). It is possible that small increases in mutation frequencies at high concentrations in this assay can be attributed to chromosomal damage rather than point mutations (Bergman et al., 1996). Shimane (1985) treated V79 cells with acetaminophen at 100, 200 and 400 µg/mL for 24 hrs, or 50, 100 and 200 µg/mL for 48 hours in the absence of metabolic activation. Solvent control treatments were only included for the 24-hr treatments. After an appropriate expression time, cultures were assessed for mutations to 6-thioguanine (6TG) and ouabain resistance. At 200 µg/mL, cytotoxicity (reduction in colony forming ability) was around 25% for the 24-hr treatment and around 40% for the 48-hr treatment, but at 400 µg/mL cytotoxicity was >50% for both treatment times. 6TG mutant frequencies increased at 200 (>2-

fold) and 400 µg/mL (>4-fold) following 24-hr treatment, but there was no statistical analysis, and no historical control data. Moreover, both of these concentrations exceed the current upper limit for testing (1 mM) according to ICH recommendations (ICH, 2011). 6TG mutant frequencies appeared also to increase at all 3 concentrations following 48-hr treatment, but since there was no solvent control for this sampling time it is not possible to assess their relevance. Ouabain-resistant mutant frequencies increased at 100 and 400 µg/mL, but not at 200 µg/mL following 24-hr treatment, so there was no dose-response. It should be noted that V79 cells are p53-deficient, and highly susceptible to misleading positive results (Fowler et al., 2012), and as such these results would be considered only of low-moderate weight.

The HID noted that:

- "*Acetaminophen produced oxidative damage in DNA, measured as 8-oxodG, in rat C6 glioma cells (Wan et al. 2004). DNA single strand breaks were slightly increased by acetaminophen in hamster lung V79 cells (Hongslo et al. 1988) and in CHO-K1 cells (Sasaki 1986), but not in a study conducted in rat hepatoma cells (Dybing et al. 1984)." (p. 163)*

The HID did not critically evaluate these studies. For example, the HID reported a LEC concentration of 1 mM for Hongslo et al. (1988); however, 1 mM acetaminophen did not cause a decrease in alkaline elution of DNA. In fact, IARC reported a LED of 10 mM for DNA damage in hamster lung V79 cells (Hongslo et al., 1988; IARC, 1999). Additionally, the studies of Hongslo et al. (1988) and Sasaki (1986) were conducted in Chinese hamster cells that have altered p53 activity. It is now known that p53-deficient rodent cells are more likely to produce "misleading" positive results (i.e. with substances that are not genotoxic or carcinogenic *in vivo*). Thus, these studies should have low weight when evaluating the genotoxic effects of acetaminophen. Further, as discussed previously, genotoxic hazard identification should be carefully evaluated alongside cytotoxic effects as cellular toxicity can confound positive responses in DNA and chromosomal tests (ICH, 2011). Positive responses were reported at cytotoxic concentrations (Hongslo et al., 1988; Sasaki, 1986) (assuming a LEC of 10 mM in Hongslo et al. 1988). Thus, the reported LECs in these studies need to be noted as being confounded by cytotoxicity and care should be taken when using these studies to evaluate the genotoxic potential of acetaminophen. Further, the positive responses reported in Hongslo et al. (1988), Sasaki (1986), and Wan et al. (2004) were all above the ICH recommended dose of 1 mM (discussed previously) for hazard identification. Specifically, Wan et al. (2004)(Wan et al., 2004) reported they tested "large doses" of acetaminophen and that "it is unlikely that low, therapeutic doses of [acetaminophen] cause oxidative damage" (Wan et al., 2004): p. 71, 75). Additionally, Wan et al. (2004) did not examine the effects of cytotoxicity on 8-oxodG formation in rat C6 glioma cells, therefore these results have limited relevance.

The HID noted that:

- "*Acetaminophen has been shown in several studies to alter UDS in rodent cells in vitro. Acetaminophen was found to increase UDS in six assays tested in mouse or rat hepatocytes (Dybing et al. 1984; Holme and Soderlund 1986), to decrease UDS in rat, hamster, and guinea pig hepatocytes and in hamster lung cells, and to have no effect in one study of rat primary hepatocytes. UDS assays measure DNA repair synthesis, and as discussed by Madle et al. (1994), the results of UDS assays can be impacted by several factors, including detection methods (autoradiography vs. liquid scintillation), specificity of the blockade of replicative DNA synthesis, metabolic capacity of the test system (determined by genetic and environmental factors), and the presence of solvents (DMSO has been shown to affect Cyp2e1 activity)" (p. 163)*

The HID reported results from deleted (archived) guidelines without discussing the reliability of these assays. As described previously, the *in vitro* SCE and Unscheduled DNA synthesis guidelines were deleted in April 2014 due, respectively, to unclear biological relevance and the availability of other tests that showed a better performance for detecting *in vitro* genotoxicity marrow. Thus, care needs to be taken when evaluating the induction of UDS (Dybing et al., 1984; Holme and Søderlund, 1986; Hongslo et al., 1988; Milam and Byard, 1985; Sasaki, 1986) as a positive response that is used to justify the genotoxic potential of acetaminophen. These results should be given less weight than other endpoints such as induction of mutations, micronuclei and chromosome aberrations. Further, while the HID noted that *in vitro* UDS assays are impacted by several factors (p. 163), no discussion was provided regarding what studies were potentially impacted by these factors. For example, increased UDS measured by the increased uptake of ^{3}H-thymidine by scintillation counting could be due to replicating cells and not due to repair. Thus, care should be taken when evaluating the positive UDS responses reported in Holme and Søderlund (1986) and Dybing et al. (1984).

The HID noted that:

- "*Acetaminophen can cause impairment of nucleotide excision repair in rodent cells in vitro. Hongslo et al. (1988) showed that UV-induced DNA-repair synthesis was decreased in hamster lung cells exposed to 0.1 mM acetaminophen and completed blocked at concentrations greater than 1 mM, as a result of the inhibition of nucleotide excision repair. Similar to what was observed in in vitro studies with human cells (Table 18) and in vivo studies in mice and rats (Table 19), acetaminophen increased SSBs after UV pre-treatment in rat hepatocytes and in NQO-treated rat testicular cells (Brunborg et al. 1995). Brunborg et al. (1995) concluded that acetaminophen's ability to delay the repair of SSBs was the result of impaired nucleotide excision repair due to acetaminophen's inhibition of ribonucleotide reductase" (p. 163-164).*

There are several studies showing a potential inhibitory effect of acetaminophen on reparative and replicative DNA synthesis *in vitro* and *in vivo* using a thymidine uptake assay. It has been proposed that this may be a result of the inhibition of ribonucleotide reductase and may explain genotoxic effects seen at high doses (Bergman et al., 1996; Thybaud et al., 2007). In these studies,

the reduced thymidine update has a clear threshold (i.e. only seen at supratherapeutic exposures) and is transient, reversing *in vivo* within 2 to 4 hours (Hongslo et al., 1994; Lister and McLean, 1997). There is no evidence that these effects are truly a consequence of effects on DNA repair rather than a consequence of reduced cell turnover. In addition, there is no evidence that the effects are sustained with multiple dosing at therapeutic or non-toxic supratherapeutic doses or lead to sustained DNA effects at non-toxic concentrations. When viewed in the context of the negative carcinogenicity studies, the data support that this mechanism does not represent a genotoxic or carcinogenic hazard to humans.

The HID noted that:

- "*In addition, Wan et al. (2004) reported that acetaminophen significantly impaired the DNA incision activity of 8-oxoguanine DNA glycosylase/AP lyase (Ogg1), a DNA repair enzyme specific for 8-oxodG, in the nuclei of rat glioma cells" (p. 164).*

Wan et al. (2004) reported a lowest effective concentration (LEC) that was greater than the ICH recommended dose of 1 mM (discussed previously) for hazard identification. Specifically, the authors reported they were testing "large doses" of acetaminophen and that "it is unlikely that low, therapeutic doses of [acetaminophen] cause oxidative damage" (Wan et al., 2004): p. 71, 75).

The HID noted that:

- "*Chromosomal effects of acetaminophen in mouse, rat, and hamster cells exposed in vitro have been observed at concentrations ranging from 0.1 mM to >1 mM, with numerous positive findings observed between 0.03 – 0.5 mM. Among twenty-four chromosomal damage assays, acetaminophen increased either MN, CAs or SCEs in twenty-three. The one study that did not observe an effect was an assay for MN in rat primary hepatocytes (Muller-Tegethoff et al. 1995)" (p. 164).*

The HID did not account for critical limitations in several studies when reporting the results of the potential genotoxic effect of acetaminophen in human cell *in vitro* systems. For example, the HID reported that 23 studies/assays showed an increase in MN, CA, or SCEs. However, it should be noted that 20 of these 23 studies/assays were conducted in Chinese hamster cells (Holme et al., 1988; Hongslo et al., 1988; Ishidate et al., 1978; Ishidate et al., 1988; Matsumura et al., 1982; Muller et al., 1991; NTP, 1993; Sasaki, 1986; Sasaki et al., 1980; Sasaki et al., 1983; Shimane, 1985). It is now known that p53-deficient rodent cells are more likely to produce "misleading" positive results (i.e. with substances that are not genotoxic or carcinogenic *in vivo*), particularly for clastogenicity, than p53-competent human cells (Fowler et al., 2012). It is therefore not uncommon to find positive clastogenicity results in p53-deficient Chinese hamster cell lines (CHO, CHL, V79) with substances that are negative in p53-competent human lymphocytes or human TK6 cells, or to find positive results at lower concentrations in Chinese hamster cells than in human cells. Thus, more weight should be given to results in p53-competent human cells than

p53-deficient hamster cells. These 20 studies in hamster cell lines should be given less weight and care should be taken when interpreting these positive results in these 20 studies. Further, Sasaki et al. (1983) and Sasaki (1986) are reported as two independent studies; however the results reported for chromosome aberrations were identical in the 2 publications and thus should be considered as one result.

The HID did not discuss whether a study that investigated chromosome aberrations included gaps in their count. As discussed above, the biological relevance of gaps is unclear, and conventionally they are not considered relevant for chromosome aberration assessment (results should be reported "excluding gaps"). As such, careful consideration should be given to data generated with CA assays that reported data which included gap analysis. Several studies included chromosome gaps in the total aberration count (Hongslo et al., 1990; Sasaki, 1986; Sasaki et al., 1983; Shimane, 1985). Further, the increased responses are less evident (or even not significant) when gaps are excluded from these studies. The HID reported results from deleted guidelines without discussing the reliability of these assays. As described previously, the *in vitro* Sister Chromatid Exchange assay was deleted in April 2014 due to the availability of other tests that showed a better performance for detecting genotoxicity, and due to a poor understanding of the mechanisms of action that can be detected by the test (OECD, 2017). Thus, care needs to be taken when evaluating the induction of SCEs (Holme et al., 1988; Hongslo et al., 1990; Hongslo et al., 1988; NTP, 1993; Sasaki, 1986; Shimane, 1985) as a positive response that is used to justify the genotoxic potential of acetaminophen. These results should be given less weight than other endpoints such as induction of micronuclei and chromosome aberrations.

As discussed previously, it is crucial for studies to evaluate the cytotoxicity of acetaminophen in parallel with the genotoxicity studies, as cytotoxicity can confound the genotoxic response. The latest OECD guidelines urge caution in evaluating positive responses seen only at levels of toxicity close to or above the recommended limits. Positive responses were reported at cytotoxic concentrations in several studies (Holme et al., 1988; Hongslo et al., 1990; NTP, 1993; Sasaki, 1986; Sasaki et al., 1983). Thus, the reported LECs in these studies need to be noted as potentially confounded by cytotoxicity and care should be taken when using these studies to evaluate the genotoxic potential of acetaminophen. Additionally, cytotoxicity was not reported, or it is unknown whether it was measured, in several other studies (Ishidate et al., 1978; Matsushima et al., 1999; Muller et al., 1991; Sasaki et al., 1980). Thus, it is unknown whether these studies were confounded by cytotoxicity. In addition, the HID did not discuss the relevance of a maximum dose. However, as discussed previously, ICH recommends a maximum concentration of 1 mM to maintain hazard identification. Several of the positive responses noted were at concentrations above 1 mM (Dunn et al., 1987; Holme et al., 1988; Muller et al., 1991; NTP, 1993; Sasaki, 1986). Thus, if the HID had considered a maximum dose of 1 mM for hazard identification, most of these studies would not have been considered reliable in the evaluation of acetaminophen as a

potential genotoxic compound. In addition, the HID did not report that several studies did not report coding the slides prior to scoring, which can lead to bias (Sasaki, 1986; Sasaki et al., 1980; Sasaki et al., 1983).

8.5.2.5 Non-mammalian species and acellular systems (p. 167-170)

The HID noted that:

- *"DNA strand breaks were found in Dreissena polymorpha, a freshwater zebra mussel, treated with acetaminophen at concentrations as low as 5 nM for 24- 96 hours. At 96 hours acetaminophen also induced MN formation in this model (Parolini et al. 2010)." (p. 167)*
- *"Reddy and Subramanyam (1981) reported that acetaminophen induced CAs in onion roots treated at room temperature for 2, 6, 12, 18, 24, 48 , or 72 hours." (p. 167)*

Genetic effects identified *in vivo* are generally considered more important than responses from *in vitro* tests, in particular *in vitro* tests in cell lines susceptible to misleading positive results or in non-mammalian systems (other than the Ames test) for which no recommended testing guidelines are available. As stated in the recent OECD Genetic Toxicology Guidance Document "assays conducted in mammalian cells are preferred because they are considered more relevant" (OECD, 2015): p. 4). Therefore, results in non-mammalian test systems, such as mussels and plants, should not be considered as being as relevant (i.e. not be given the same weight) as results from mammalian systems and the Ames test.

The HID noted that:

- *"Using cell-free systems, Rogers et al. (1997) reported the binding of [3H]-acetaminophen to calf thymus DNA, either in the presence of horseradish peroxide (HRP) and hydrogen peroxide (H2O2), or in the presence of rat liver microsomes. The level of DNA binding observed with the HRP-H2O2 system was 200-fold greater than that observed with rat liver microsomes. These results are consistent with the hypothesis that peroxidase-mediated metabolism of acetaminophen can produce DNA-reactive radical intermediates. Additionally, acetaminophen formed adducts with purified deoxyribonucleic acid (type I) in the presence, but not the absence of mouse liver microsomes (Dybing et al. 1984).*
- *Plattner et al. (2012) reported the non-enzymatic formation of covalent adducts of acetaminophen to guanosine, as detected by electrochemistry/liquid chromatography /mass spectrometry. These investigators observed that the first step of adduct formation involved the conversion of both guanosine and acetaminophen into radical forms via one-electron-one-proton reactions, and showed that these radicals reacted with each other to form four different guanosine-acetaminophen-2H isomers" (p. 167).*

As noted above, tritiated substances can release tritium that can be incorporated into normal DNA synthesis, leading to higher levels of background radioactivity. Therefore, caution should be taken when interpreting tritiated compound results. Rogers et al. (1997) examined DNA adduct formation in mice using ^{32}P-postlabeling and no differences were observed for acetaminophen treated mice compared to control mice. This result was in contrast to the observed result that

tritiated label was observed with hepatic and renal tissue at all doses tested. Therefore, the results with purified DNA by Rogers et al. (1997) and Dybing et al. (1984) or guanosine by Plattner et al. (2012) should be interpreted with caution as *in vivo* studies have demonstrated that acetaminophen does not form DNA adducts when measured by the sensitive ^{32}P-postlabeling technique (Hasegawa et al., 1988; Rogers et al., 1997; Williams et al., 2007).

8.5.2.6 In vitro cell transformation (p. 188)

The HID noted:

- *"Cell transformation assays are designed to detect a change in the growth pattern of cells that is indicative of loss of contact inhibition, a phenotype that is characteristic of cancer cells.*

  Patierno et al. (1989) studied in vitro cell transformation of C3H/10T1/2 clone 8 mouse embryo fibroblast (10T1/2) cells exposed to acetaminophen. These cells are considered to be similar to BALB/3T3 and Swiss/3T3 cells, as they are stable in culture and highly sensitive to post-confluence inhibition of cell division (Reznikoff et al. 1973). C3H/10T1/2 cells, together with other immortalized aneuploid mouse cells, represent one of the two major types of systems used for in vitro cell transformation assays, the other type being primary diploid cells, such as Syrian Hamster Embryo cells (Creton et al. 2012).

  In this study, Patierno et al. (1989) treated 10T1/2 cells with acetaminophen at concentrations ranging from 0.5 – 2.0 mg/mL (3.3 to 13 mM) for either 24 hours without S-9 or 3 hours with Arochlor 1254-induced hamster liver S-9. In the absence of S-9 acetaminophen induced a small, but dose-dependent increase in the number of type II morphologically transformed foci. A greater number of type II transformed foci were induced by acetaminophen in the presence of S-9. Similar cell transformation results were observed with the carcinogen phenacetin (of which acetaminophen is a major metabolite). Several metabolites of acetaminophen (and phenacetin) were also tested in C3H/10T1/2 cells (NAPQI, PAP, p-benzoquinone), and each were found to be inactive in the cell transformation assay. Patierno et al. (1989) characterized the type II foci induced by acetaminophen and phenacetin as atypical (weak) non-neoplastic morphologically transformed cells that "did not exhibit any other classical parameters of neoplastic transformation, such as increased saturation density or anchorage independence." (p. 188)

Patierno et al. (1989) indicated that the "results suggest that metabolic intermediates of high concentrations of phenacetin and acetaminophen induce a low frequency of nonneoplastic morphological transformation of 10T½ mouse embryo cells" (Patierno et al., 1989): p. 1038). Further, the authors noted that "[e]ven though the mixed clones reformed weak type II foci when maintained at confluence, they did not exhibit any other classical parameters of neoplastic transformation, such as increased saturation density or anchorage independence" (Patierno et al., 1989): p. 1043). Therefore, the results by Patierno et al. (1989) suggest that acetaminophen does not cause neoplastic transformation in this *in vitro* assay.

8.5.3 Scientific Accuracy and Completeness Issues Identified in Structure Activity Considerations, Toxcast, High Throughput Screening and KCC Assessment

The HID is not transparent in selection of HTS assays related to carcinogenicity, and the assessment cannot be reproduced without additional information. High throughput screening (HTS) data are publicly available from the ToxCast/Tox21 screening programs. Through collaboration of multiple U.S. agencies, thousands of chemicals have been analyzed in hundreds of *in vitro* assays. The output from these HTS assays comprise a wide range of endpoints related to molecular or cellular events that could potentially be components of a mechanistic pathway associated with a similarly wide range of toxicological outcomes. When data from *in vitro* models are used as part of evaluation for potential health hazard, it is well-recognized that consideration of aspects related to assay methodology (e.g., cytotoxicity) and context both between assays as well as with other evidence streams (particularly that from *in vivo* studies) are critical in the interpretation of observed activity (Becker et al., 2017; Judson et al., 2016; Wikoff et al., 2019). Consideration of these aspects was not fully apparent in the HID, as described below.

The HID has relied upon the key characteristics of carcinogens (KCCs) approach to organize the HTS data as part of the evaluation. As described in the HID, only a subset of the HTS data are considered relevant to these characteristics – these assays are manually selected based on "mapping" assay endpoints from the HTS data to specific KCCs. However, the HID does not utilize publicly available mappings, does not provide documentation of the mappings used to select KCC-relevant assays, nor is any information provided for the criteria or procedure for assembling such assay endpoint-to-KCC mappings. The HID authors cite private email communication for the procurement of the most up-to-date mappings that have been determined by the International Agency for Research on Cancer (IARC) and enumerates the number of mapped assays without further context (OEHHA, 2019), p. 201).

Further, the presentation of active HTS data without mapping to KCCs is extraneous. Tables 33 and 34 of the HID present all active HTS data regardless of a possible anchor to carcinogenicity and thus have the capacity to suggest a greater level of activity than is relevant to the given assessment.

The impact of the lack of transparency with the selection of HTS data is especially important in consideration of the fact that the HID only reports data with activity (i.e., assays relevant to KCC that were inactive were not reported). This has substantial impact on the interpretation as it is common for multiple assay endpoints to measure a similar biological signal and understanding both activity and inactivity across all assay endpoints measured for a given biological process or mechanism is important. Without provision of the assay endpoint-to-KCC mapping or reporting of any inactivity data, data integration is regarded as incomplete for developing weight-of-the-evidence conclusions regarding activity.

The HID is not consistent in applying criteria related to the reliability of individual assays: regarding assay-specific data quality issues (reported as flags in the summary files and in the dashboard), the HID appears to appropriately account for both data quality issues and for loss of cell viability when such was tested in an assay. However, for assays in which cell viability was not directly tested in addition to the assay target, no other consideration of cytotoxicity appears to have been included. This represents an inconsistent consideration of cytotoxicity across the assays included in the evaluation. It is unclear why such an approach was taken as data are available for the incorporation of this critical component of cytotoxic interference in all assays. Specifically, a battery of cell viability assays is included in ToxCast/Tox21, and the cytotoxic range for each chemical has been determined based on these assays. Further, the potential assay interference due to cytotoxicity is characterized by Z-scores for each chemical and assay endpoint pair, which indicate the distance between the assay AC_{50} value (i.e. the concentration eliciting 50% maximal activity) and concentrations eliciting cytotoxicity, as described in Judson *et al.*, (2016). Assay endpoint activity should be taken in context of this "cytotoxic signal burst" information for understanding the activity of test articles in assays that do not also have data for direct measures of cytotoxicity. Z-scores and other cytotoxicity data are available via the summary files download:

https://epa.figshare.com/articles/ToxCast_and_Tox21_Summary_Files/6062479

This could have been utilized in a more consistent evaluation. One such example (detailed below) is the inclusion of an assay endpoint that measures the activation of nuclear factor erythroid 2-like 2 (NFE2L2), an oxidative stress-related transcription factor (relevant to KCC #5 – "Induces oxidative stress"), for p-benzoquinone. Activity of the test article was well above the global cytotoxic burst range for this compound as determined by the battery of cell viability assays, and as demonstrated by a Z-score <0 (-0.69), while a Z-score cut-off of ≥3 has been suggested as a criterion for activity below the cytotoxicity concentration range (Judson et al., 2016).

Similarly, consideration of other assay validity criteria was also not consistent. The authors of the HID applied the criterion of a "pass" grade for chemical sample QC for inclusion in the assessment, based on analytical testing for identity (molecular weight) and purity at receipt and after 4 months of storage. These analytical data are only available for assays from the Tox21 program. While it is logical to exclude data collected using chemical samples that have quality issues, including data for which quality metrics are unavailable is questionable. A more consistent approach would be to also exclude data for which chemical sample quality control (QC) metrics are unavailable, or to seek additional information for any samples with issues concerning chemical QC, as additional analytical data on a sample-to-sample basis may be available from the US EPA.

A newer version of the HTS data are available and, thus, the data presented in the HID represent an evaluation of outdated data. While it is understandable that March 7th (dashboard version

3.0.5 integrating data from the invitroDB_v3 release) were utilized in the HID given the workflow, the newer data released on August 9th, 2019 (dashboard version 3.0.9 integrating data from the invitrodb_v3.2 release) (EPA, 2019) should be relied upon in the continued evaluation by the CIC. The IARC KCC mapping referenced in the HID was current as of May 24th, 2018; thus, the mapping is unlikely to include all assay endpoints that are currently available for acetaminophen and metabolites.

None of the HTS data are formally integrated across KCCs (e.g., no information on activity relative to inactivity within a KCC, nor are the HTS data formally integrated with other evidence streams in the HID). Viewed in context of the preclinical findings, which would account for many of the limitations in interpretation of *in vitro* assays, as well as account for activity associated with metabolites (even following chronic exposure to very high doses), the activity observed in the BeleutaHTS data are without biological significance. Numerous preclinical assays demonstrate a lack of adversity associated with the molecular or cellular signals obtained in the ToxCast/Tox21 assays. In addition, there is no evidence that the metabolite concentrations utilized in the *in vitro* assays are relevant to therapeutic exposures in humans. Additional assessment of the concentrations used in the HTS assays using methods such as IVIVE are required to interpret the findings.

In addition, the HID also highlights the following characteristics that are associated with acetaminophen: (1) it forms an electrophilic reactive metabolite, (2) has the potential to cause oxidative stress, (3) has the potential to be genotoxic, and (4) has the potential to alter DNA repair. However, they neglect to highlight that 2-4 have a threshold and only occur under cytotoxic conditions and are only observed in certain model systems.

Specific Comments on HTS data for acetaminophen and two rodent metabolites

For the purposes of comment preparation, independent analyses based on publicly available mappings with expert curation were conducted. For such, all concentration and Z-score values, data quality flags, and assay descriptions reflect data from the most current HTS data release.

Acetaminophen (APAP)

HTS data for acetaminophen were not discussed in the HID due to the fact that all 5 of the assay endpoints in which acetaminophen was active were flagged by the ToxCast screening program for issues with data quality. The most recent version of the ToxCast data contains data for 665 assay endpoints for acetaminophen, of which 4 were considered as "active" in the ToxCast dashboard and in the HID, in contrast to the 636 assay endpoints (with 5 active) accounted for in the HID from an earlier version of the data. The 4 assay endpoints with activity in the more recent release are all included in the list of active assay endpoints from the earlier release; in other words, no 'new' activity was published in the update.

The availability of data that demonstrate inactivity in KCC-relevant assay endpoints may provide valuable contextual information. While the authors of the HID do not provide a list of the assay endpoints that were mapped to KCCs, it is safely assumed that many among the 665 assay endpoints in which acetaminophen was tested are relevant to one or more KCC. For example, using mappings based on publicly available information as described above, over 250 assay endpoints with primary read-out data (i.e., not including assay endpoints that provide contextual information, such as cell viability measures or counterscreen/specificity assays) did not have flags for data quality issues, and were mapped to one or more KCCs, were inactive for acetaminophen. That is, the majority of the HTS data for acetaminophen are inactive.

Cytotoxicity information was not included in the HID for acetaminophen. Acetaminophen was not cytotoxic at concentrations up to 100 µM, the highest concentration tested, as determined by a battery of cell viability assays included in the ToxCast data.

The HID surmised that "the inactivity of acetaminophen in the ToxCast assays may be due to the lack of metabolic activation in the testing systems." Accordingly, HTS data for two metabolites, p-benzoquinone and p-aminophenol, were also evaluated and included in the HID, comments on these metabolites are below. This statement essentially constitutes unsupported speculation, as no data or citations were provided in the HID showing that any of these metabolites besides NAPQI are formed at any appreciable levels in humans. In addition, given that they have only been detected in rodents, the negative NTP carcinogenicity studies demonstrate that if they are formed, they do not cause cancer in rodents at the levels that they were formed in the cancer bioassays. Therefore, the carcinogenicity and genotoxicity data for the metabolites should not be considered in the hazard assessment of the carcinogenicity of acetaminophen.

In conclusion, because there is no evidence that these metabolites are formed in humans, any potential effects associated with these metabolites are not relevant to the acetaminophen carcinogenicity hazard assessment.

Nonetheless, there are a number of issues identified with the review of HTS data as presented in the HID that are addressed below.

p-Benzoquinone

p-Benzoquinone is postulated to be a metabolite of acetaminophen in mice by indirect evidence (Pascoe et al., 1988). There is no evidence that it is formed in humans. This compound was tested in the ToxCast/Tox21 program.

General cytotoxicity information was not included in the HID for p-benzoquinone. p-Benzoquinone was cytotoxic *in vitro*, with a median cytotoxic concentration of 36.53µM and a lower bound of 7.98µM, as determined by the ToxCast screening program analysts based on a battery of cell viability assays included in the ToxCast data.

The newer version of the ToxCast data contains data for 580 assay endpoints for p-benzoquinone, of which 103 were considered "active" in the dashboard and in the HID, in contrast to the 556 assay endpoints (with 105 active) accounted for in the HID from an earlier version of the data. 102 of the 103 assay endpoints with activity in the more recent release are included in the list of 105 active assay endpoints from the earlier release; in other words, a single 'new' endpoint with activity was reported, which is not relevant to the KCC.

After excluding all assay data that were collected using chemical samples with sub-optimal sample QC metrics, or for which data quality flags were reported, the HID reported 28 active assays. Only 7 assay endpoints assigned as "active" in the ToxCast dashboard were determined to be relevant to the KCCs. As such, the presentation of the active assays not linked to carcinogenicity is extraneous (particularly when it is considered that the inactive data are also not provided, recognizing the important context provided by such).

The first assay endpoint is "ATG_NRF2_ARE_CIS_up", which measures the activation of nuclear factor erythroid 2-like 2 (NFE2L2 or Nrf2), an oxidative stress-related transcription factor (relevant to KCC #5 – "Induces oxidative stress"). p-Benzoquinone has an AC_{50} of 51.89µM, which is above the cytotoxic lower bound and median, and is well above the cytotoxic burst range as demonstrated by a Z-score <0 (-0.69). Thus, the p-benzoquinone is not considered active in this assay at sub-cytotoxic concentrations. Further, p-benzoquinone was inactive in 12 other assays related to oxidative stress, as mapped by TS staff (number of inactive oxidative stress assays using the mappings that the HID used is unknown). However, only 1 such assay was without data quality flags or potential chemical QC issues: "ATG_NRF1_CIS_up" that also measures an oxidative stress-related transcription factor.

The second assay listed in the HID is "ATG_ERa_Trans_up," an inducible reporter assay for the *ESR1* gene, relevant for KCC #8 – "Modulates receptor-mediated effects." The AC_{50} value for p-benzoquinone in this assay is 29.96µM, which is above the cytotoxic lower bound, but below the cytotoxic median. Applying a Z-score criterion of ≥3 as recommended by Judson *et al.* (2016) would deem this assay inactive (the Z-score is 0.39). For context, p-benzoquinone was inactive in 7 other assays related to estrogen receptor activity that were without data quality flags or chemical QC issues.

The final four assays listed in the HID are all related to KCC #10 – "Alters cell proliferation, cell death or nutrient supply": BSK_3C_Proliferation_down, BSK_CASM3C_Proliferation_down, BSK_hDFCGF_Proliferation_down, and BSK_SAg_Proliferation_down. These assay endpoints can all be considered active, as there were no data quality or chemical QC issues, and the AC_{50} values are all below the cytotoxicity median concentration and also below the cytotoxic lower bound in all cases except for 1. The Z-scores are all at least 2.55. While p-benzoquinone was active in these 4 assay endpoints, the assay is a measure of loss of cell viability, as opposed to cellular proliferation. Examples of signals relevant to KCC #10 is described in (Smith et al., 2016) are:

"Increased proliferation, decreased apoptosis, changes in growth factors, energetics and signaling pathways related to cellular replication or cell cycle control, angiogenesis." Thus, the loss in cell viability as indicated by the activity in these assays does not demonstrate a signal related the cancer mechanism category intended by KCC #10.

Overall, p-benzoquinone does not appear to induce activity related to any of the KCCs when HTS data with flags for data quality issues or sub-optimal chemical sample QC are applied, as described by the HID, and when appropriate integration of cytotoxicity data and direction of bioactivity measures are included in the evaluation of the data.

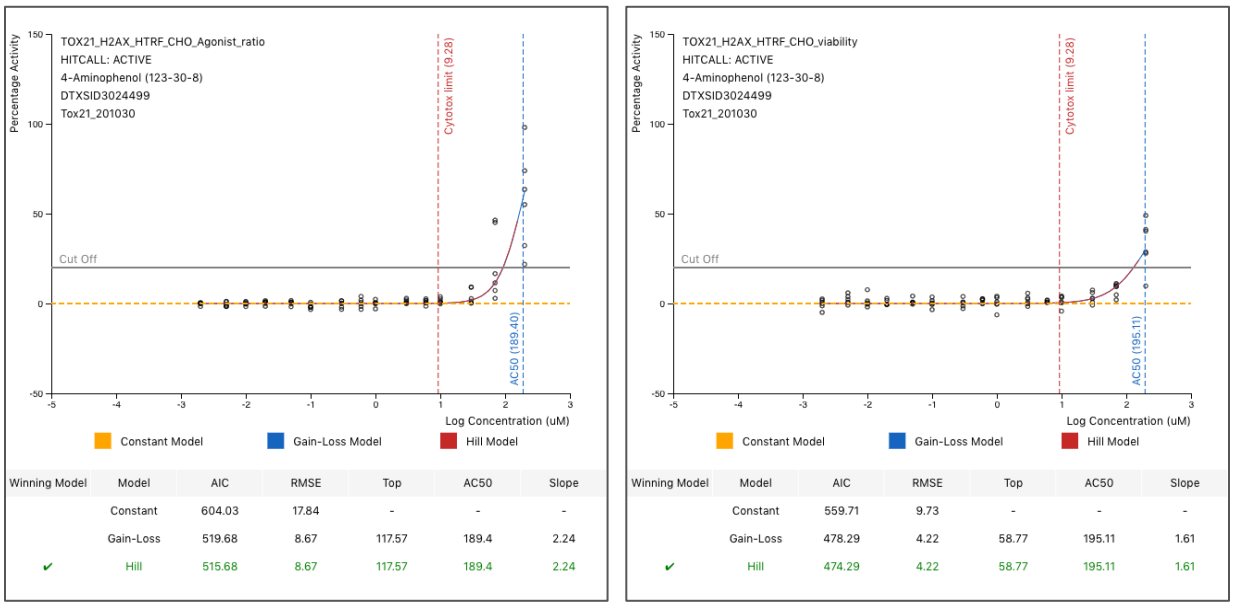
p-Aminophenol (PAP)

p-Aminophenol (PAP), which was tested in the ToxCast/Tox21 program, is posited as a metabolite of APAP in rats from *in vivo* and *in vitro* evidence (Gemborys and Mudge, 1981; Mugford and Tarloff, 1995; Newton et al., 1982). However, PAP has not been confirmed as a metabolite in humans.

General cytotoxicity information was not included in the HID for PAP. PAP was cytotoxic *in vitro*, with a median cytotoxic concentration of 42.49µM and a lower bound of 9.28µM, as determined by the ToxCast screening program analysts based on a battery of cell viability assays included in the ToxCast data. After excluding all assay data that was collected using chemical samples with sub-optimal sample QC metrics, or for which data quality flags were reported, the authors of the HID reported 13 active assays, of which 12 were considered "active" as relevant to a KCC.

The first assay endpoint is "TOX21_H2AX_HTRF_CHO_Agonist_ratio", a measure of phosphorylation of histone H2A.X at serine 139, a marker of DNA double strand breaks, in Chinese hamster ovary cells (relevant to KCC #2 – "Is genotoxic"). PAP has an AC_{50} of 189.4µM in this assay, and an AC_{50} of 195.11 µM for cell viability loss in the same assay ("TOX21_H2AX_HTRF_CHO_viability"). While the activity occurs at a concentration below cell viability loss, the concentrations are remarkably close and considerably high in general and well above the overall median cytotoxic concentration for PAP. The HID did not include the cell viability measure for this assay, based on the fact that the data were flagged as having "Less than 50% efficacy," without any other data issues. While PAP can be considered to be active in the assay, the evidence is considered weak at best due to the potential cytotoxic interference based on visual inspection (Figure 60) and consideration contextual data.

Figure 60: Concentration-response curves for PAP in the H2AX agonist (left) and H2AX viability (right) assays.



Winning Model	Model	AIC	RMSE	Top	AC50	Slope
	Constant	604.03	17.84	-	-	-
	Gain-Loss	519.68	8.67	117.57	189.4	2.24
✔	Hill	515.68	8.67	117.57	189.4	2.24

Winning Model	Model	AIC	RMSE	Top	AC50	Slope
	Constant	559.71	9.73	-	-	-
	Gain-Loss	478.29	4.22	58.77	195.11	1.61
✔	Hill	474.29	4.22	58.77	195.11	1.61

Source:https://comptox.epa.gov/dashboard/dsstoxdb/results?search=DTXSID3024499#invitrodb-bioassays-toxcast-tox21, October 3rd, 2019.

PAP was listed in the HID as active in 10 assays targeting various receptors, relevant KCC #8 – "Modulates receptor-mediated effects." PAP was active for aryl hydrocarbon receptor agonism at a concentration lower than that at which cell viability was lost in the assays (AC_{50} value of 17.42µM for agonism vs. 42.33µM for loss of cell viability). PAP was active for androgen receptor (AR) antagonism in two assays at concentrations lower than that at which cell viability was lost in the same assay. In one of these assays, the HID did not consider the cell viability measure due to issues with the chemical sample tested in the assay. In doing so, it would be logical to defer to a Z-score cut-off criteria to understand if this assay occurred below the cytotoxic concentration range. If the Z-score cut-off of ≥3 was applied as suggested in Judson *et al*. (2016), this assay would not be considered active (Z-score was 2.34). An assay for antagonism of estrogen receptor-alpha (ER-a), PAP was active at a concentration very close to the concentration at which the same sample induced cell viability loss in the assay (AC_{50} of 74.28µM vs. 75.12µM, respectively), indicating that the activity occurs at a similar threshold for cytotoxicity. For antagonism of estrogen receptor-beta (ER-b), activity occurred at a lower concentration than loss of cell viability; however, similar to the case of one of the aryl hydrocarbon receptor assays, the HID did not consider the cell viability assay due to a single data quality flag. The Z-score for the estrogen receptor-beta antagonist activity is 0.92 (AC_{50} is 26.63µM).

PAP was active for estrogen-related receptor-alpha antagonism in two assays, both at concentrations lower than loss of cell viability within the same assay. Antagonist activity for the

peroxisome proliferator activator receptor-delta (PPAR-d), the vitamin D receptor (VDR), and the retinoic acid receptor-related orphan receptor-gamma (ROR-g) occurred at concentrations lower than loss of cell viability in the same assay; however, the HID did not include the cell viability assay information for the VDR or the PPAR-d antagonist assays, and the Z-scores for the antagonism assays were <0 as the activity occurred above the median cytotoxic concentration. Model predictions as published in the ToxCast dashboard and as published in two articles classify PAP as "inactive" for both estrogen and androgen antagonist activity (https://comptox.epa.gov/dashboard/dsstoxdb/results?search=DTXSID3024499#bioactivity-toxcast-model; (Kleinstreuer et al., 2017_ENREF_163; Mansouri et al., 2016).

Regarding specificity, PAP induced activity at lower concentrations in two specificity assays for antagonism, one for ER-a antagonism and one for AR antagonism. It is expected that a putative antagonist specific to these receptors would have a higher AC_{50} in these specificity assays rather than lower. This contextual information, together with the diversity of the receptors for which PAP exerted antagonist activity, may suggest cross-reactivity of PAP across receptors. Further analysis of positive or negative controls would be informative to confirm better understand the specificity of PAP as an antagonist to various receptors.

The final assay endpoint listed in the HID in which PAP was active is related to KCC #10 – "Alters cell proliferation, cell death or nutrient supply." This assay endpoint ("TOX21_AP1_BLA_Agonist_ratio") measures activation of the transcription factor Activator Protein-1 (AP-1), which is an important regulator of cell proliferation, differentiation, apoptosis, and angiogenesis. PAP can be considered active for this endpoint, as there were no data quality or chemical QC issues, and the AC_{50} value is below the AC_{50} for loss of cell viability tested within this assay (16.58µM vs. 50.32µM). No other assay endpoints that are measures of proliferation were available for PAP.

Overall, PAP appears to induce non-specific receptor antagonist activity, although the relationship of antagonism of the receptors tested with carcinogenic activity is not apparent in the HID. Limited evidence of the ability of PAP to alter signaling relevant to cell cycle and proliferation were apparent.

9 References

Akakpo, J.Y., Ramachandran, A., Kandel, S.E., Ni, H.M., Kumer, S.C., Rumack, B.H., Jaeschke, H., 2018. 4-Methylpyrazole protects against acetaminophen hepatotoxicity in mice and in primary human hepatocytes. Hum Exp Tox 37, 1310-1322.

Amo, H., Matsuyama, M., 1985. Subchronic and chronic effects of feeding of large amounts of acetaminophen in B6C3F1 mice. Nihon Eiseigaku Zasshi 40, 567-574.

Anderson, B.J., 2008. Paracetamol (Acetaminophen): mechanisms of action. Paediatr Anaesth 18, 915-921.

Andersson, B., Nordenskjold, M., Rahimtula, A., Moldeus, P., 1982. Prostaglandin synthetase-catalyzed activation of phenacetin metabolites to genotoxic products. Mol Pharmacol 22, 479-485.

Andrews, R.S., Bond, C.C., Burnett, J., Saunders, A., Watson, K., 1976. Isolation and identification of paracetamol metabolites. J Int Med Res 4, 34-39.

Angervall, L., Bengtsson, U., Zetterlund, C.G., Zsigmond, M., 1969. Renal pelvic carcinoma in a Swedish district with abuse of a phenacetin-containing drug. Br J Urol 41, 401-405.

Baeg, N.J., Bodenheimer, H.C., Jr., Burchard, K., 1988. Long-term sequellae of acetaminophen-associated fulminant hepatic failure: relevance of early histology. Am J Gastroenterol 83, 569-571.

Bajt, M.L., Cover, C., Lemasters, J.J., Jaeschke, H., 2006. Nuclear translocation of endonuclease G and apoptosis-inducing factor during acetaminophen-induced liver cell injury. Tox Sci 94, 217-225.

Bajt, M.L., Farhood, A., Lemasters, J.J., Jaeschke, H., 2008. Mitochondrial bax translocation accelerates DNA fragmentation and cell necrosis in a murine model of acetaminophen hepatotoxicity. J Pharm Exp Ther 324, 8-14.

Baker, J.A., Weiss, J.R., Czuczman, M.S., Menezes, R.J., Ambrosone, C.B., Moysich, K.B., 2005. Regular use of aspirin or acetaminophen and risk of non-Hodgkin lymphoma. Cancer Causes Control 16, 301-308.

Bandi, S., Viswanathan, P., Gupta, S., 2014. Evaluation of cytotoxicity and DNA damage response with analysis of intracellular ATM signaling pathways. Assay Drug Dev Technol 12, 272-281.

Baris, D., Karagas, M.R., Koutros, S., Colt, J.S., Johnson, A., Schwenn, M., Fischer, A.H., Figueroa, J.D., Berndt, S.I., Han, S., Beane Freeman, L.E., Lubin, J.H., Cherala, S., Cantor, K.P., Jacobs, K., Chanock, S., Chatterjee, N., Rothman, N., Silverman, D.T., 2013. Nonsteroidal anti-inflammatory drugs and other analgesic use and bladder cancer in northern New England. Int J Cancer 132, 162-173.

Becker, N., Fortuny, J., Alvaro, T., Nieters, A., Maynadie, M., Foretova, L., Staines, A., Brennan, P., Boffetta, P., Cocco, P.L., de Sanjose, S., 2009. Medical history and risk of lymphoma: results of a European case-control study (EPILYMPH). J Cancer Res Clin Oncol 135, 1099-1107.

Becker, R.A., Dreier, D.A., Manibusan, M.K., Cox, L.A.T., Simon, T.W., Bus, J.S., 2017. How well can carcinogenicity be predicted by high throughput "characteristics of carcinogens" mechanistic data? Regul Toxicol Pharmacol 90, 185-196.

Bengtsson, U., Angervall, L., Ekman, H., Lehmann, L., 1968. Transitional cell tumors of the renal pelvis in analgesic abusers. Scand J Urol Nephrol 2, 145-150.

Bergman, K., Muller, L., Teigen, S.W., 1996. Series: current issues in mutagenesis and carcinogenesis, No. 65. The genotoxicity and carcinogenicity of paracetamol: a regulatory (re)view. Mutat Res 349, 263-288.

Binkova, B., Topinka, J., Sram, R.J., 1990. The effect of paracetamol on oxidative damage in human peripheral lymphocytes. Mutat Res 244, 227-231.

Bisaglia, M., Venezia, V., Piccioli, P., Stanzione, S., Porcile, C., Russo, C., Mancini, F., Milanese, C., Schettini, G., 2002. Acetaminophen protects hippocampal neurons and PC12 cultures from amyloid beta-peptides induced oxidative stress and reduces NF-kappaB activation. Neurochem Int 41, 43-54.

Blair, I.A., Boobis, A.R., Davies, D.S., Cresp, T.M., 1980. Paracetamol oxidation: synthesis and reactivity of N-acetyl-p-benzoquinoneimine. Tetrahedron Lett. 21, 4947– 4950.

Bloom, L., Boyle, K.E., Myers, A.E., Blacketer, C., Weinstein, R., 2019. Frequency of nonaspirin NSAID-relevant coexisting medical conditions in the primary-care setting: a retrospective database review. Ther Clin Risk Manag 15, 579-588.

Blough, E.R., Wu, M., 2011. Acetaminophen: beyond pain and Fever-relieving. Front Pharmacol 2, 72.

Bolt, H.M., Foth, H., Hengstler, J.G., Degen, G.H., 2004. Carcinogenicity categorization of chemicals-new aspects to be considered in a European perspective. Toxicol Lett 151, 29-41.

Brodie, B.B., Axelrod, J., 1949. The fate of acetophenetidin in man and methods for the estimation of acetophenetidin and its metabolites in biological material. J Pharmacol Exp Ther 97, 58-67.

Brown, L.M., Gridley, G., Check, D., Landgren, O., 2008. Risk of multiple myeloma and monoclonal gammopathy of undetermined significance among white and black male United States veterans with prior autoimmune, infectious, inflammatory, and allergic disorders. Blood 111, 3388-3394.

Brunborg, G., Holme, J.A., Hongslo, J.K., 1995. Inhibitory effects of paracetamol on DNA repair in mammalian cells. Mutat Res 342, 157-170.

Brusick, D., Aardema, M., Kier, L., Kirkland, D., Williams, G., 2016. Genotoxicity Expert Panel review: weight of evidence evaluation of the genotoxicity of glyphosate, glyphosate-based formulations, and aminomethylphosphonic acid. Crit Rev Toxicol 46, 56-74.

Burke, D.A., Wedd, D.J., Herriott, D., Bayliss, M.K., Spalding, D.J., Wilcox, P., 1994. Evaluation of pyrazole and ethanol induced S9 fraction in bacterial mutagenicity testing. Mutagenesis 9, 23-29.

Bush, S.H., Tollin, S., Marchion, D.C., Xiong, Y., Abbasi, F., Ramirez, I.J., Zgheib, N.B., Boac, B., Judson, P.L., Chon, H.S., Wenham, R.M., Apte, S.M., Cubitt, C.L., Berglund, A.E., Havrilesky, L.J.,

Lancaster, J.M., 2016. Sensitivity of ovarian cancer cells to acetaminophen reveals biological pathways that affect patient survival. Mol Clin Oncol 4, 399-404.

Calder, I.C., Goss, D.E., Williams, P.J., Funder, C.C., Green, C.R., Ham, K.N., Tange, J.D., 1976. Neoplasia in the rat induced by N-hydroxyphenacetin, a metabolite of phenacetin. Pathology 8, 1-6.

Caldwell, D.J., 1999. Review of mononuclear cell leukemia in F-344 rat bioassays and its significance to human cancer risk: A case study using alkyl phthalates. Regul Toxicol Pharmacol 30, 45-53.

Camus, A.M., Friesen, M., Croisy, A., Bartsch, H., 1982. Species-specific activation of phenacetin into bacterial mutagens by hamster liver enzymes and identification of N-hydroxyphenacetin O-glucuronide as a promutagen in the urine. Cancer Res 42, 3201-3208.

Caparrotta, T.M., Antoine, D.J., Dear, J.W., 2018. Are some people at increased risk of paracetamol-induced liver injury? A critical review of the literature. . Eur J Clin Pharmacol 74, 147-160.

Castelao, J.E., Yuan, J.M., Gago-Dominguez, M., Yu, M.C., Ross, R.K., 2000. Non-steroidal anti-inflammatory drugs and bladder cancer prevention. Br J Cancer 82, 1364-1369.

Chang, E.T., Zheng, T., Weir, E.G., Borowitz, M., Mann, R.B., Spiegelman, D., Mueller, N.E., 2004. Aspirin and the risk of Hodgkin's lymphoma in a population-based case-control study. J Natl Cancer Inst 96, 305-315.

Chen, W., Koenigs, L.L., Thompson, S.J., Peter, R.M., Rettie, A.E., Trager, W.F., Nelson, S.D., 1998. Oxidation of acetaminophen to its toxic quinone imine and nontoxic catechol metabolites by baculovirus-expressed and purified human cytochromes P450 2E1 and 2A6. Chem Res Toxicol 11, 295-301.

Cho, E., Curhan, G., Hankinson, S.E., Kantoff, P., Atkins, M.B., Stampfer, M., Choueiri, T.K., 2011. Prospective evaluation of analgesic use and risk of renal cell cancer. Arch Intern Med 171, 1487-1493.

Chow, W.H., McLaughlin, J.K., Linet, M.S., Niwa, S., Mandel, J.S., 1994. Use of analgesics and risk of renal cell cancer. Int J Cancer 59, 467-470.

Clark, P.M., Clark, J.D., Wheatley, T., 1986. Urine discoloration after acetaminophen overdose. Clin Chem 32, 1777-1778.

Clark, R., Borirakchanyavat, V., Davidson, A.R., Thompson, R.P., Widdop, B., Goulding, R., Williams, R., 1973. Hepatic damage and death from overdose of paracetamol. Lancet 1, 66-70.

Clements, J., 1992. Study to determine the ability of paracetamol to induce mutations at the thymidine kinase (tk) locus in mouse lymphoma L5178Y cells using a fluctuation assay, . Hazleton Microtest Report No. ZTKREUBF.001 (March 31. 1992).

Cohen, S.D., Pumford, N.R., Khairallah, E.A., Boekelheide, K., Pohl, L.R., Amouzadeh, H.R., Hinson, J.A., 1997. Selective protein covalent binding and target organ toxicity. Toxicol Appl Pharmacol 143, 1-12.

Cohen, S.M., Ellwein, L.B., 1991. Cell proliferation and bladder tumor promotion. Prog Clin Biol Res 369, 347-355.

Cohen, S.M., Ohnishi, T., Clark, N.M., He, J., Arnold, L.L., 2007. Investigations of rodent urinary bladder carcinogens: collection, processing, and evaluation of urine and bladders. Toxicol Pathol 35, 337-347.

Colin, P., Sirois, G., Chakrabarti, S., 1986. Effects of route of administration on the dose-dependent metabolism of acetaminophen in rats: relationship with its toxicity. Arch Int Pharmacodyn Ther 281, 181-191.

Corbett, M.D., Corbett, B.R., 1988. Nucleic acid binding of arylamines during the respiratory burst of human granulocytes. Chem Res Toxicol 1, 356-363.

Corbett, M.D., Corbett, B.R., Hannothiaux, M.H., Quintana, S.J., 1989. Metabolic activation and nucleic acid binding of acetaminophen and related arylamine substrates by the respiratory burst of human granulocytes. Chem Res Toxicol 2, 260-266.

Court, M.H., Duan, S.X., von Moltke, L.L., Greenblatt, D.J., Patten, C.J., Miners, J.O., Mackenzie, P.I., 2001. Interindividual variability in acetaminophen glucuronidation by human liver microsomes: identification of relevant acetaminophen UDP-glucuronosyltransferase isoforms. J Pharmacol Exp Ther 299, 998-1006.

Court, M.H., Zhu, Z., Masse, G., Duan, S.X., James, L.P., Harmatz, J.S., Greenblatt, D.J., 2017. Race, Gender, and Genetic Polymorphism Contribute to Variability in Acetaminophen Pharmacokinetics, Metabolism, and Protein-Adduct Concentrations in Healthy African-American and European-American Volunteers. J Pharmacol Exp Ther 362, 431-440.

Couto, A.C., Ferriera, J.D., Pombo-de-Oliveira, M.S., Koifman, S., 2015. Brazilian Collaborative Study Group of Infant Acute Leukemia. Pregnancy maternal exposure to analgesic medicines, and leukemia in Brazilian children below 2 years of age. . Eur J Cancer Prevention 24, 245-252.

Cover, C., Fickert, P., Knight, T.R., Fuchsbichler, A., Farhood, A., Trauner, M., Jaeschke, H., 2005a. Pathophysiological role of poly(ADP-ribose) polymerase (PARP) activation during acetaminophen-induced liver cell necrosis in mice. Toxicol Sci 84, 201-208.

Cover, C., Mansouri, A., Knight, T.R., Bajt, M.L., Lemasters, J.J., Pessayre, D., Jaeschke, H., 2005b. Peroxynitrite-induced mitochondrial and endonuclease-mediated nuclear DNA damage in acetaminophen hepatotoxicity. J Pharmacol Exp Ther 315, 879-887.

Critchley, J.A., Nimmo, G.R., Gregson, C.A., Woolhouse, N.M., Prescott, L.F., 1986. Inter-subject and ethnic differences in paracetamol metabolism. Br J Clin Pharmacol 22, 649-657.

Dahlin, D.C., Miwa, G.T., Lu, A.Y., Nelson, S.D., 1984. N-acetyl-p-benzoquinone imine: a cytochrome P-450-mediated oxidation product of acetaminophen. Proc Natl Acad Sci U S A 81, 1327-1331.

Davern, T.J., 2nd, James, L.P., Hinson, J.A., Polson, J., Larson, A.M., Fontana, R.J., Lalani, E., Munoz, S., Shakil, A.O., Lee, W.M., Acute Liver Failure Study, G., 2006. Measurement of serum acetaminophen-protein adducts in patients with acute liver failure. Gastroenterology 130, 687-694.

Davis, D.C., Potter, W.Z., Jollow, D.J., Mitchell, J.R., 1974. Species differences in hepatic glutathione depletion, covalent binding and hepatic necrosis after acetaminophen. Life Sci 14, 2099-2109.

de Morais, S.M., Uetrecht, J.P., Wells, P.G., 1992. Decreased glucuronidation and increased bioactivation of acetaminophen in Gilbert's syndrome. Gastroenterology 102, 577-586.

Derby, L.E., Jick, H., 1996. Acetaminophen and renal and bladder cancer. Epidemiology 7, 358-362.

Desler, C., Lykke, A., Rasmussen, L.J., 2010. The effect of mitochondrial dysfunction on cytosolic nucleotide metabolism. J Nucleic Acids 2010.

Desler, C., Munch-Petersen, B., Stevnsner, T., Matsui, S., Kulawiec, M., Singh, K.K., Rasmussen, L.J., 2007. Mitochondria as determinant of nucleotide pools and chromosomal stability. Mutat Res 625, 112-124.

Dierkes, G., Weiss, T., Modick, H., Kafferlein, H.U., Bruning, T., Koch, H.M., 2014. N-Acetyl-4-aminophenol (paracetamol), N-acetyl-2-aminophenol and acetanilide in urine samples from the general population, individuals exposed to aniline and paracetamol users. Int J Hyg Environ Health 217, 592-599.

Dou, X., Li, J., Danelisen, I., Trush, M., Misra, H., Zhu, H., Jia, Z., Li, Y., 2017. Acetaminophen, the Active Ingredient of Tylenol, Protects against Peroxynitrite-Induced DNA Damage: A Chemiluminometric and Electron Paramagnetic Resonance Spectrometric Study. Reactive Oxygen Species.

Du, K., Farhood, A., Jaeschke, H., 2017a. Mitochondria-targeted antioxidant Mito-Tempo protects against acetaminophen hepatotoxicity. Arch Toxicol 91, 761-773.

Du, K., Ramachandran, A., Jaeschke, H., 2016. Oxidative stress during acetaminophen hepatotoxicity: Sources, pathophysiological role and therapeutic potential. Redox Biol 10, 148-156.

Du, K., Ramachandran, A., McGill, M.R., Mansouri, A., Asselah, T., Farhood, A., Woolbright, B.L., Ding, W.X., Jaeschke, H., 2017b. Induction of mitochondrial biogenesis protects against acetaminophen hepatotoxicity. Food Chem Toxicol 108, 339-350.

Du, K., Xie, Y., McGill, M.R., Jaeschke, H., 2015. Pathophysiological significance of c-jun N-terminal kinase in acetaminophen hepatotoxicity. Expert Opin Drug Metab Toxicol 11, 1769-1779.

DuBois, R.N., Hill, K.E., Burk, R.F., 1983. Antioxidant effect of acetaminophen in rat liver. Biochem Pharmacol 32, 2621-2622.

Dunn, T.L., Gardiner, R.A., Seymour, G.J., Lavin, M.F., 1987. Genotoxicity of analgesic compounds assessed by an *in vitro* micronucleus assay. Mutat Res 189, 299-306.

Dybing, E., Holme, J.A., Gordon, W.P., Soderlund, E.J., Dahlin, D.C., Nelson, S.D., 1984. Genotoxicity studies with paracetamol. Mutat Res 138, 21-32.

Dyson, J., Jaques, B., Chattopadyhay, D., Lochan, R., Graham, J., Das, D., Aslam, T., Patanwala, I., Gaggar, S., Cole, M., Sumpter, K., Stewart, S., Rose, J., Hudson, M., Manas, D., Reeves, H.L., 2014.

Hepatocellular cancer: the impact of obesity, type 2 diabetes and a multidisciplinary team. J Hepatol 60, 110-117.

Eastmond, D.A., 1993. Induction of micronuclei and aneuploidy by the quinone-forming agents benzene and o-phenylphenol. Toxicol Lett 67, 105-118.

Eastmond, D.A., 2017. Recommendations for the evaluation of complex genetic toxicity data sets when assessing carcinogenic risks to humans. Environ Mol Mutagen 58, 380-385.

EFSA, 2011. Scientific opinion on genotoxicity testing strategies applicable to food and feed safety assessment (EFSA Scientific Committee) (Question no. EFSA-Q-2009-00782, adopted on 13 September 2011 by European Food Safety Authority, 3 October 2012, replaces the earlier version). EFSA J. 2379. Available from: http://www.efsa.europa.eu/en/efsajournal/pub/2379.htm.

El-Serag, H.B., 2012. Epidemiology of viral hepatitis and hepatocellular carcinoma. Gastroenterology 142, 1264-1273 e1261.

EPA, 2005. Guidelines for Carcinogen Risk Assessment. U.S. Environmental Protection Agency, Washington, DC.

EPA, 2019. ToxCast & Tox21 summary files from invitrodb_v3.2. Available: https://www.epa.gov/chemical-research/toxicity-forecaster-toxcasttm-data [accessed August 11th, 2019]. U.S. Environmental Protection Agency, Washington, DC.

FDA, 2010. Supervisor's Secondary Review - Pharmacology/Toxicology Review and Evaluation: acetaminophen application. Federal Drug Administration.

Fischer, V., West, P.R., Nelson, S.D., Harvison, P.J., Mason, R.P., 1985. Formation of 4-aminophenoxyl free radical from the acetaminophen metabolite N-acetyl-p-benzoquinone imine. J Biol Chem 260, 11446-11450.

Flaks, A., Flaks, B., 1982. 3-Methylcholanthrene-inhibition of hepatocarcinogenesis in the rat due to 3'-methyl-4-dimethylaminoazobenzene or 2-acetylamino-fluorene: a comparative study. Carcinogenesis 3, 981-991.

Flaks, A., Flaks, B., 1983. Induction of liver cell tumours in IF mice by paracetamol. Carcinogenesis 4, 363-368.

Flaks, B., 1978. Effects of chronic oral dosing with quinine sulphate in the rat. Pathol Res Pract 163, 373-377.

Flaks, B., Flaks, A., Shaw, A.P., 1985. Induction by paracetamol of bladder and liver tumours in the rat. Effects on hepatocyte fine structure. Acta Pathol Microbiol Immunol Scand A 93, 367-377.

Forrest, J.A., Adriaenssens, P., Finlayson, N.D., Prescott, L.F., 1979. Paracetamol metabolism in chronic liver disease. Eur J Clin Pharmacol 15, 427-431.

Forte, A.J., Wilson, J.M., Slattery, J.T., Nelson, S.D., 1984. The formation and toxicity of catechol metabolites of acetaminophen in mice. Drug Metab Dispos 12, 484-491.

Fortuny, J., Kogevinas, M., Garcia-Closas, M., Real, F.X., Tardon, A., Garcia-Closas, R., Serra, C., Carrato, A., Lloreta, J., Rothman, N., Villanueva, C., Dosemeci, M., Malats, N., Silverman, D., 2006. Use of analgesics and nonsteroidal anti-inflammatory drugs, genetic predisposition, and bladder cancer risk in Spain. Cancer Epidemiol Biomarkers Prev 15, 1696-1702.

Fortuny, J., Kogevinas, M., Zens, M.S., Schned, A., Andrew, A.S., Heaney, J., Kelsey, K.T., Karagas, M.R., 2007. Analgesic and anti-inflammatory drug use and risk of bladder cancer: a population based case control study. BMC Urol 7, 13.

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

Friedman, G.D., 1982. Phenylbutazone, musculoskeletal disease, and leukemia. J Chronic Dis 35, 233-243.

Friis, S., Nielsen, G.L., Mellemkjaer, L., McLaughlin, J.K., Thulstrup, A.M., Blot, W.J., Lipworth, L., Vilstrup, H., Olsen, J.H., 2002. Cancer risk in persons receiving prescriptions for paracetamol: a Danish cohort study. Int J Cancer 97, 96-101.

Gago-Dominguez, M., Yuan, J.M., Castelao, J.E., Ross, R.K., Yu, M.C., 1999. Regular use of analgesics is a risk factor for renal cell carcinoma. Br J Cancer 81, 542-548.

Gaylor, D.W., 2005. Are tumor incidence rates from chronic bioassays telling us what we need to know about carcinogens? Regul Toxicol Pharmacol 41, 128-133.

Gelotte, C.K., Auiler, J.F., Lynch, J.M., Temple, A.R., Slattery, J.T., 2007. Disposition of acetaminophen at 4, 6, and 8 g/day for 3 days in healthy young adults. Clin Pharmacol Ther 81, 840-848.

Gemborys, M.W., Mudge, G.H., 1981. Formation and disposition of the minor metabolites of acetaminophen in the hamster. Drug Metab Dispos 9, 340-351.

Genkinger, J.M., De Vivo, I., Stampfer, M.J., Giovannucci, E., Michaud, D.S., 2007. Nonsteroidal antiinflammatory drug use and risk of bladder cancer in the health professionals follow-up study. Int J Cancer 120, 2221-2225.

Giri, A.K., Sivam, S.S., Khan, K.A., 1992. Sister-chromatid exchange and chromosome aberrations induced by paracetamol *in vivo* in bone-marrow cells of mice. Mutat Res 278, 253-258.

Gow, P.J., Ghabrial, H., Smallwood, R.A., Morgan, D.J., Ching, M.S., 2001. Neonatal hepatic drug elimination. Pharmacol Toxicol 88, 3-15.

Graham, G.G., Scott, K.F., 2005. Mechanism of action of paracetamol. Am J Ther 12, 46-55.

Guengerich, F.P., Liebler, D.C., 1985. Enzymatic activation of chemicals to toxic metabolites. Crit Rev Toxicol 14, 259-307.

Hagiwara, A., Ward, J.M., 1986. The chronic hepatotoxic, tumor-promoting, and carcinogenic effects of acetaminophen in male B6C3F1 mice. Fundam Appl Toxicol 7, 376-386.

Hamlyn, A.N., Douglas, A.P., James, O.F., Lesna, M., Watson, A.J., 1977. Liver function and structure in survivors of acetaminophen poisoning. A follow-up study of serum bile acids and liver histology. Am J Dig Dis 22, 605-610.

Han, D., Dara, L., Win, S., Than, T.A., Yuan, L., Abbasi, S.Q., Liu, Z.X., Kaplowitz, N., 2013. Regulation of drug-induced liver injury by signal transduction pathways: critical role of mitochondria. Trends Pharmacol Sci 34, 243-253.

Hanawa, N., Shinohara, M., Saberi, B., Gaarde, W.A., Han, D., Kaplowitz, N., 2008. Role of JNK translocation to mitochondria leading to inhibition of mitochondria bioenergetics in acetaminophen-induced liver injury. J Biol Chem 283, 13565-13577.

Hantson, P., de Saint-Georges, L., Mahieu, P., Leonard, E.D., Crutzen-Fayt, M.C., Leonard, A., 1996. Evaluation of the ability of paracetamol to produce chromosome aberrations in man. Mutat Res 368, 293-300.

Hasegawa, R., Furukawa, F., Toyoda, K., Jang, J.J., Yamashita, K., Sato, S., Takahashi, M., Hayashi, Y., 1988. Study for tumor-initiating effect of acetaminophen in two-stage liver carcinogenesis of male F344 rats. Carcinogenesis 9, 755-759.

Haseman, J.K., 1983. Patterns of tumor incidence in two-year cancer bioassay feeding studies in Fischer 344 rats. Fundam Appl Toxicol 3, 1-9.

Haseman, J.K., Arnold, J., Eustis, S.L., 1990. Tumor Incidences in Fischer 344 Rats: NTP Historical Data, in: Boorman, G.A., Eustis, S.L., Elwell, M.R., Montgomery, C.A., Jr., MacKenzie, W.F. (Eds.), Pathology of the Fischer Rat - Reference and Atlas. Academic Press, Inc., San Diego, CA, pp. 555-564.

Haseman, J.K., Hailey, J.R., Morris, R.W., 1998. Spontaneous neoplasm incidences in Fischer 344 rats and B6C3F1 mice in two-year carcinogenicity studies: a National Toxicology Program update. Toxicol Pathol 26, 428-441.

Haworth, S., Lawlor, T., Mortelmans, K., Speck, W., Zeiger, E., 1983. Salmonella mutagenicity test results for 250 chemicals. Environ Mutagen 5 Suppl 1, 1-142.

Hayward, K.L., Powell, E.E., Irvine, K.M., Martin, J.H., 2016. Can paracetamol (acetaminophen) be administered to patients with liver impairment? Br J Clin Pharmacol 81, 210-222.

Hazleton Microtest, 1992. Unscheduled DNA synthesis and S- phase synthesis induced by paracetamol in rat liver, Study No. UBFl/lLU.

Heard, K., Green, J.L., Anderson, V., Bucher-Bartelson, B., Dart, R.C., 2016. Paracetamol (acetaminophen) protein adduct concentrations during therapeutic dosing. Br J Clin Pharmacol 81, 562-568.

Heard, K.J., Green, J.L., James, L.P., Judge, B.S., Zolot, L., Rhyee, S., Dart, R.C., 2011. Acetaminophen-cysteine adducts during therapeutic dosing and following overdose. BMC Gastroenterol 11, 20.

Heddle, J.A., Salamone, M.F., 1981. Chromosomal aberrations and bone marrow toxicity. Environ Health Perspect 39, 23-27.

Hill, A.B., 1965. The Environment and Disease: Association or Causation? Proc R Soc Med 58, 295-300.

Hindmarsh, K.W., Mayers, D.J., Wallace, S.M., Danilkewich, A., Ernst, A., 1991. Increased serum sulfate concentrations in man due to environmental factors: effects on acetaminophen metabolism. Vet Hum Toxicol 33, 441-445.

Hinson, J.A., Reid, A.B., McCullough, S.S., James, L.P., 2004. Acetaminophen-induced hepatotoxicity: role of metabolic activation, reactive oxygen/nitrogen species, and mitochondrial permeability transition. Drug Metab Rev 36, 805-822.

Hiraga, K., Fujii, T., 1985. Carcinogenicity testing of acetaminophen in F344 rats. Jpn J Cancer Res 76, 79-85.

Hogestatt, E.D., Jonsson, B.A., Ermund, A., Andersson, D.A., Bjork, H., Alexander, J.P., Cravatt, B.F., Basbaum, A.I., Zygmunt, P.M., 2005. Conversion of acetaminophen to the bioactive N-acylphenolamine AM404 via fatty acid amide hydrolase-dependent arachidonic acid conjugation in the nervous system. J Biol Chem 280, 31405-31412.

Hoivik, D.J., Manautou, J.E., Tveit, A., Hart, S.G., Khairallah, E.A., Cohen, S.D., 1995. Gender-related differences in susceptibility to acetaminophen-induced protein arylation and nephrotoxicity in the CD-1 mouse. Toxicol Appl Pharmacol 130, 257-271.

Holland, N., Bolognesi, C., Kirsch-Volders, M., Bonassi, S., Zeiger, E., Knasmueller, S., Fenech, M., 2008. The micronucleus assay in human buccal cells as a tool for biomonitoring DNA damage: the HUMN project perspective on current status and knowledge gaps. Mutat Res 659, 93-108.

Holm, J.B., Chalmey, C., Modick, H., Jensen, L.S., Dierkes, G., Weiss, T., Jensen, B.A., Norregard, M.M., Borkowski, K., Styrishave, B., Martin Koch, H., Mazaud-Guittot, S., Jegou, B., Kristiansen, K., Kristensen, D.M., 2015. Aniline Is Rapidly Converted Into Paracetamol Impairing Male Reproductive Development. Toxicol Sci 148, 288-298.

Holme, J.A., Hongslo, J.K., Bjornstad, C., Harvison, P.J., Nelson, S.D., 1988. Toxic effects of paracetamol and related structures in V79 Chinese hamster cells. Mutagenesis 3, 51-56.

Holme, J.A., Søderlund, E., 1986. Species differences in cytotoxic and genotoxic effects of phenacetin and paracetamol in primary monolayer cultures of hepatocytes. Mutat Res 164, 167-175.

Hongslo, J.K., Bjorge, C., Schwarze, P.E., Brogger, A., Mann, G., Thelander, L., Holme, J.A., 1990. Paracetamol inhibits replicative DNA synthesis and induces sister chromatid exchange and chromosomal aberrations by inhibition of ribonucleotide reductase. Mutagenesis 5, 475-480.

Hongslo, J.K., Brogger, A., Bjorge, C., Holme, J.A., 1991. Increased frequency of sister-chromatid exchange and chromatid breaks in lymphocytes after treatment of human volunteers with therapeutic doses of paracetamol. Mutat Res 261, 1-8.

Hongslo, J.K., Brunborg, G., Steffensen, I.L., Holme, J.A., 1993. Paracetamol inhibits UV-induced DNA repair in resting human mononuclear blood cells *in vitro*. Mutagenesis 8, 423-429.

Hongslo, J.K., Christensen, T., Brunborg, G., Bjornstad, C., Holme, J.A., 1988. Genotoxic effects of paracetamol in V79 Chinese hamster cells. Mutat Res 204, 333-341.

Hongslo, J.K., Smith, C.V., Brunborg, G., Soderlund, E.J., Holme, J.A., 1994. Genotoxicity of paracetamol in mice and rats. Mutagenesis 9, 93-100.

Hu, J.J., Lee, M.J., Vapiwala, M., Reuhl, K., Thomas, P.E., Yang, C.S., 1993. Sex-related differences in mouse renal metabolism and toxicity of acetaminophen. Toxicol Appl Pharmacol 122, 16-26.

Huff, J., 1993. Chemicals and cancer in humans: first evidence in experimental animals. Environ Health Perspect 100, 201-210.

IARC, 1990. Monograph for Paracetamol (Acetaminophen). Volume 50. World Health Organization, International Agency for Research on Cancer, Lyon, France, pp. 307-332.

IARC, 1999. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Some Chemicals that Cause Tumours of the Kidney or ladder in Rodents and Some Other Substances. Vol 73. WHO-IARC Lyon, France

Ibrulj, S., Rahmanovic, A., Haveric, S., Haveric, A., Pasic, A.D., 2007. Cytogenetic evaluation of paracetamol effects in human lymphocytes culture. Drug Chem Toxicol 30, 133-143.

ICH, 2011. ICH S2(R1). Guidance on genotoxicity testing and data interpretation for pharmaceuticals intended for human use. International Conference on harmonization of technical requirements for human use. ICH Harmonised Tripartite Guideline. Step 4, 9 November 2011.

Imamura, A., Kurumi, Y., Danzuka, T., Kodama, M., Kawachi, T., Nagao, M., 1983. Classification of compounds by cluster analysis of Ames test data. Gan 74, 196-204.

Isaka, H., Yoshii, H., Otsuji, A., Koike, M., Nagai, Y., Koura, M., Sugiyasu, K., Kanabayashi, T., 1979. Tumors of Sprague-Dawley rats induced by long-term feeding of phenacetin. Gan 70, 29-36.

Ishidate, K., Yoshida, M., Nakazawa, Y., 1978. Effect of typical inducers of microsomal drug-metabolizing enzymes on phospholipid metabolism in rat liver. Biochem Pharmacol 27, 2595-2603.

Ishidate, M., Jr., 1983. Chromosomal aberration test in vitro. Realize Inc., Tokyo, p. 31–32.

Ishidate, M., Jr., Harnois, M.C., Sofuni, T., 1988. A comparative analysis of data on the clastogenicity of 951 chemical substances tested in mammalian cell cultures. Mutat Res 195, 151-213.

Ito, N., Tsuda, H., Tatematsu, M., Inoue, T., Tagawa, Y., Aoki, T., Uwagawa, S., Kagawa, M., Ogiso, T., Masui, T., et al., 1988. Enhancing effect of various hepatocarcinogens on induction of preneoplastic glutathione S-transferase placental form positive foci in rats--an approach for a new medium-term bioassay system. Carcinogenesis 9, 387-394.

Jaeschke, H., 1990. Glutathione disulfide formation and oxidant stress during acetaminophen-induced hepatotoxicity in mice in vivo: the protective effect of allopurinol. J Pharmacol Exp Ther 255, 935-941.

Jaeschke, H., Duan, L., Akakpo, J.Y., Farhood, A., Ramachandran, A., 2018. The role of apoptosis in acetaminophen hepatotoxicity. Food Chem Toxicol 118, 709-718.

Jahr, S., Hentze, H., Englisch, S., Hardt, D., Fackelmayer, F.O., Hesch, R.D., Knippers, R., 2001. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. Cancer Res 61, 1659-1665.

Jasiewicz, M.L., Richardson, J.C., 1987. Absence of mutagenic activity of benorylate, paracetamol and aspirin in the Salmonella/mammalian microsome test. Mutat Res 190, 95-100.

Jetten, M.J., Claessen, S.M., Dejong, C.H., Lahoz, A., Castell, J.V., van Delft, J.H., Kleinjans, J.C., 2014. Interindividual variation in response to xenobiotic exposure established in precision-cut human liver slices. Toxicology 323, 61-69.

Jetten, M.J., Gaj, S., Ruiz-Aracama, A., de Kok, T.M., van Delft, J.H., Lommen, A., van Someren, E.P., Jennen, D.G., Claessen, S.M., Peijnenburg, A.A., Stierum, R.H., Kleinjans, J.C., 2012. 'Omics analysis of low dose acetaminophen intake demonstrates novel response pathways in humans. Toxicol Appl Pharmacol 259, 320-328.

Johansson, B.B., 1981a. Indomethacin and cerebrovascular permeability to albumin in acute hypertension and cerebral embolism in the rat. Exp Brain Res 42, 331-336.

Johansson, S., Angervall, L., Bengtsson, U., Wahlqvist, L., 1974. Uroepithelial tumors of the renal pelvis associated with abuse of phenacetin-containing analgesics. Cancer 33, 743-753.

Johansson, S.L., 1981b. Carcinogenicity of analgesics: long-term treatment of Sprague-Dawley rats with phenacetin, phenazone, caffeine and paracetamol (acetamidophen). Int J Cancer 27, 521-529.

Josephy, P.D., Eling, T.E., Mason, R.P., 1983. Oxidation of p-aminophenol catalyzed by horseradish peroxidase and prostaglandin synthase. Mol Pharmacol 23, 461-466.

Judson, R., Houck, K., Martin, M., Richard, A.M., Knudsen, T.B., Shah, I., Little, S., Wambaugh, J., Setzer, R.W., Kothiya, P., Phuong, J., Filer, D., Smith, D., Reif, D., Rotroff, D., Kleinstreuer, N., Sipes, N., Xia, M., Huang, R., Crofton, K., Thomas, R.S., 2016. Analysis of the Effects of Cell Stress and Cytotoxicity on In Vitro Assay Activity Across a Diverse Chemical and Assay Space. Toxicol Sci 153, 409.

Kang, A.M., Padilla-Jones, A., Fisher, E.S., Akakpo, J.Y., Jaeschke, H., Rumack, B.H., Gerkin, R.D., Curry, S.C., in press. The effect of 4-methylpyrazole on oxidative metabolism of acetaminophen in human volunteers. . J Med Tox.

Kanki, K., Nishikawa, A., Masumura, K., Umemura, T., Imazawa, T., Kitamura, Y., Nohmi, T., Hirose, M., 2005. *In vivo* mutational analysis of liver DNA in gpt delta transgenic rats treated with the hepatocarcinogens N-nitrosopyrrolidine, 2-amino-3-methylimidazo[4,5-f]quinoline, and di(2-ethylhexyl)phthalate. Mol Carcinog 42, 9-17.

Karami, S., Daughtery, S.E., Schwartz, K., Davis, F.G., Ruterbusch, J.J., Wacholder, S., Graubard, B.I., Berndt, S.I., Hofmann, J.N., Purdue, M.P., Moore, L.E., Colt, J.S., 2016. Analgesic use and risk of renal cell carcinoma: A case-control, cohort and meta-analytic assessment. Int J Cancer 139, 584-592.

Kato, I., Koenig, K.L., Shore, R.E., Baptiste, M.S., Lillquist, P.P., Frizzera, G., Burke, J.S., Watanabe, H., 2002. Use of anti-inflammatory and non-narcotic analgesic drugs and risk of non-Hodgkin's lymphoma (NHL) (United States). Cancer Causes Control 13, 965-974.

Kaye, J.A., Myers, M.W., Jick, H., 2001. Acetaminophen and the risk of renal and bladder cancer in the general practice research database. Epidemiology 12, 690-694.

Kennon-McGill, S., McGill, M.R., 2018. Extrahepatic toxicity of acetaminophen: critical evaluation of the evidence and proposed mechanisms. J Clin Transl Res 3, 297-310.

King-Herbert, A., Thayer, K., 2006. NTP workshop: animal models for the NTP rodent cancer bioassay: stocks and strains--should we switch? Toxicol Pathol 34, 802-805.

King-Herbert, A.P., Sills, R.C., Bucher, J.R., 2010. Commentary: update on animal models for NTP studies. Toxicol Pathol 38, 180-181.

King, M.T., Beikirch, H., Eckhardt, K., Gocke, E., Wild, D., 1979. Mutagenicity studies with x-ray-contrast media, analgesics, antipyretics, antirheumatics and some other pharmaceutical drugs in bacterial, Drosophila and mammalian test systems. Mutat Res 66, 33-43.

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

Kirkland, D.J., 1992. Chromosomal aberration tests in vitro: problems with protocol design and interpretation of results. Mutagenesis 7, 95-106.

Kirkland, D.J., Dresp, J.H., Marshall, R.R., Baumeister, M., Gerloff, C., Gocke, E., 1992. Normal chromosomal aberration frequencies in peripheral lymphocytes of healthy human volunteers exposed to a maximum daily dose of paracetamol in a double blind trial. Mutat Res 279, 181-194.

Kleinstreuer, N.C., Ceger, P., Watt, E.D., Martin, M., Houck, K., Browne, P., Thomas, R.S., Casey, W.M., Dix, D.J., Allen, D., Sakamuru, S., Xia, M., Huang, R., Judson, R., 2017. Development and Validation of a Computational Model for Androgen Receptor Activity. Chem Res Toxicol 30, 946-964.

Knight, T.R., Fariss, M.W., Farhood, A., Jaeschke, H., 2003. Role of lipid peroxidation as a mechanism of liver injury after acetaminophen overdose in mice. Toxicol Sci 76, 229-236.

Knight, T.R., Ho, Y.S., Farhood, A., Jaeschke, H., 2002. Peroxynitrite is a critical mediator of acetaminophen hepatotoxicity in murine livers: protection by glutathione. J Pharmacol Exp Ther 303, 468-475.

Knight, T.R., Kurtz, A., Bajt, M.L., Hinson, J.A., Jaeschke, H., 2001. Vascular and hepatocellular peroxynitrite formation during acetaminophen toxicity: role of mitochondrial oxidant stress. Toxicol Sci 62, 212-220.

Kocišová, J., Rossner, P., Binkova, B., Bavorova, H., Sram, R.J., 1988. Mutagenicity studies on paracetamol in human volunteers. I. Cytogenetic analysis of peripheral lymphocytes and lipid peroxidation in plasma. Mutat Res 209, 161-165.

Kocišová, J., Šrám, R.J., 1990. Mutagenicity studies on paracetamol in human volunteers. III. Cytokinesis block micronucleus method. Mutat Res 244, 27-30.

Kon, K., Kim, J.S., Jaeschke, H., Lemasters, J.J., 2004. Mitochondrial permeability transition in acetaminophen-induced necrosis and apoptosis of cultured mouse hepatocytes. Hepatology 40, 1170-1179.

Kreiger, N., Marrett, L.D., Dodds, L., Hilditch, S., Darlington, G.A., 1993. Risk factors for renal cell carcinoma: results of a population-based case-control study. Cancer Causes Control 4, 101-110.

Kucerova, M., Polivkova, Z., Matousek, V., 1980. Chromosomal aberrations and SCE in lymphocytes of children revaccinated against smallpox. Mutat Res 71, 263-267.

Kurata, Y., Asamoto, M., Hagiwara, A., Masui, T., Fukushima, S., 1986. Promoting effects of various agents in rat urinary bladder carcinogenesis initiated by N-butyl-N-(4-hydroxybutyl)nitrosamine. Cancer Lett 32, 125-135.

Laine, J.E., Auriola, S., Pasanen, M., Juvonen, R.O., 2009. Acetaminophen bioactivation by human cytochrome P450 enzymes and animal microsomes. Xenobiotica 39, 11-21.

Lauterburg, B.H., 2002. Analgesics and glutathione. Am J Ther 9, 225-233.

Lesna, M., Watson, A.J., Douglas, A.P., Hamlyn, A.N., James, O.F., 1976. Evaluation of paracetamol-induced damage in liver biopsies. Acute changes and follow-up findings. Virchows Arch A Pathol Anat Histol 370, 333-344.

Lewis, J.D., Strom, B.L., Kimmel, S.E., Farrar, J., Metz, D.C., Brensinger, C., Nessel, L., Localio, A.R., 2006. Predictors of recall of over-the-counter and prescription non-steroidal anti-inflammatory drug exposure. Pharmacoepidemiol Drug Saf 15, 39-45.

Lindqvist, E.K., Goldin, L.R., Landgren, O., Blimark, C., Mellqvist, U.H., Turesson, I., Wahlin, A., Bjorkholm, M., Kristinsson, S.Y., 2011. Personal and family history of immune-related conditions increase the risk of plasma cell disorders: a population-based study. Blood 118, 6284-6291.

Lindqvist, E.K., Landgren, O., Lund, S.H., Turesson, I., Hultcrantz, M., Goldin, L., Bjorkholm, M., Kristinsson, S.Y., 2017. History of autoimmune disease is associated with impaired survival in multiple myeloma and monoclonal gammopathy of undetermined significance: a population-based study. Ann Hematol 96, 261-269.

Linet, M.S., Chow, W.H., McLaughlin, J.K., Wacholder, S., Yu, M.C., Schoenberg, J.B., Lynch, C., Fraumeni, J.F., Jr., 1995. Analgesics and cancers of the renal pelvis and ureter. Int J Cancer 62, 15-18.

Lington, A.W., Dodd, D.E., Ridlon, S.A., Douglas, J.F., Kneiss, J.J., Andrews, L.S., 1997. Evaluation of 13-week inhalation toxicity study on methyl t-butyl ether (MTBE) in Fischer 344 rats. J Appl Toxicol 17 Suppl 1, S37-44.

Lipworth, L., Friis, E., Mellemgaard, A., SIgnorello, L.B., Johnsen, S.P., Nielsen, G.L., McLaughlin, J.K., Blot, W.J., Olsen, J.H., 2003. A population-based cohort study of mortality among adults prescribed paracetamol in Denmark. . J Clin Epidemiol 56, 796-801.

Lister, C.F., McLean, A.E., 1997. Inhibition of DNA synthesis by paracetamol in different tissues of the rat in vivo. Toxicology 116, 49-57.

Lowrance, W.T., Ordonez, J., Udaltsova, N., Russo, P., Go, A.S., 2014. CKD and the risk of incident cancer. J Am Soc Nephrol 25, 2327-2334.

Lutz, W.K., Schlatter, C., 1979. In vivo covalent binding of chemicals to DNA as a short-term test for carcinogenicity. Arch Toxicol Suppl, 411-415.

Mansouri, K., Abdelaziz, A., Rybacka, A., Roncaglioni, A., Tropsha, A., Varnek, A., Zakharov, A., Worth, A., Richard, A.M., Grulke, C.M., Trisciuzzi, D., Fourches, D., Horvath, D., Benfenati, E., Muratov, E., Wedebye, E.B., Grisoni, F., Mangiatordi, G.F., Incisivo, G.M., Hong, H., Ng, H.W., Tetko, I.V., Balabin, I., Kancherla, J., Shen, J., Burton, J., Nicklaus, M., Cassotti, M., Nikolov, N.G., Nicolotti, O., Andersson, P.L., Zang, Q., Politi, R., Beger, R.D., Todeschini, R., Huang, R., Farag, S., Rosenberg, S.A., Slavov, S., Hu, X., Judson, R.S., 2016. CERAPP: Collaborative Estrogen Receptor Activity Prediction Project. Environ Health Perspect 124, 1023-1033.

Manyike, P.T., Kharasch, E.D., Kalhorn, T.F., Slattery, J.T., 2000. Contribution of CYP2E1 and CYP3A to acetaminophen reactive metabolite formation. Clin Pharmacol Ther 67, 275-282.

Markovic, D., Katic, J., Stojkovic, R., Borovic, S., Zarkovic, N., Fucic, A., 2013. Lipid peroxidation, detoxification capacity, and genome damage in mice after transplacental exposure to pharmaceutical drugs. Braz J Med Biol Res 46, 1014-1020.

Maronpot, R.R., Flake, G., Huff, J., 2004. Relevance of animal carcinogenesis findings to human cancer predictions and prevention. Toxicol Pathol 32 Suppl 1, 40-48.

Maronpot, R.R., Montgomery, C.A., Jr., Boorman, G.A., McConnell, E.E., 1986. National Toxicology Program nomenclature for hepatoproliferative lesions of rats. Toxicol Pathol 14, 263-273.

Maronpot, R.R., Nyska, A., Foreman, J.E., Ramot, Y., 2016. The legacy of the F344 rat as a cancer bioassay model (a retrospective summary of three common F344 rat neoplasms). Critical Reviews in Toxicology 46, 641-675.

Martinez, A., Urios, A., Blanco, M., 2000. Mutagenicity of 80 chemicals in Escherichia coli tester strains IC203, deficient in OxyR, and its oxyR(+) parent WP2 uvrA/pKM101: detection of 31 oxidative mutagens. Mutat Res 467, 41-53.

Maruyama, H., Tanaka, T., Williams, G.M., 1990. Effects of the peroxisome proliferator di(2-ethylhexyl)phthalate on enzymes in rat liver and on carcinogen-induced liver altered foci in comparison to the promoter phenobarbital. Toxicol Pathol 18, 257-267.

Matsumura, T., Sugahara, T., Wanatabe, R., Ishihara, Y., Imanishi, H., Miyaji, T., 1982. [The clinical behavior, especially the roentgenographic findings, of a developing odontoma over a 2-year period]. Osaka Daigaku Shigaku Zasshi 27, 107-112.

Matsushima, T., Hayashi, M., Matsuoka, A., Ishidate, M., Jr., Miura, K.F., Shimizu, H., Suzuki, Y., Morimoto, K., Ogura, H., Mure, K., Koshi, K., Sofuni, T., 1999. Validation study of the *in vitro* micronucleus test in a Chinese hamster lung cell line (CHL/IU). Mutagenesis 14, 569-580.

Matsushita, K., Kijima, A., Ishii, Y., Takasu, S., Jin, M., Kuroda, K., Kawaguchi, H., Miyoshi, N., Nohmi, T., Ogawa, K., Umemura, T., 2013. Development of a Medium-term Animal Model Using gpt Delta Rats to Evaluate Chemical Carcinogenicity and Genotoxicity. J Toxicol Pathol 26, 19-27.

Mazaleuskaya, L.L., Sangkuhl, K., Thorn, C.F., FitzGerald, G.A., Altman, R.B., Klein, T.E., 2015. PharmGKB summary: pathways of acetaminophen metabolism at the therapeutic versus toxic doses. Pharmacogenet Genomics 25, 416-426.

McCredie, M., Ford, J.M., Stewart, J.H., 1988. Risk factors for cancer of the renal parenchyma. Int J Cancer 42, 13-16.

McCredie, M., Pommer, W., McLaughlin, J.K., Stewart, J.H., Lindblad, P., Mandel, J.S., Mellemgaard, A., Schlehofer, B., Niwa, S., 1995. International renal-cell cancer study. II. Analgesics. Int J Cancer 60, 345-349.

McCredie, M., Stewart, J.H., 1988. Does paracetamol cause urothelial cancer or renal papillary necrosis? Nephron 49, 296-300.

McCredie, M., Stewart, J.H., Day, N.E., 1993. Different roles for phenacetin and paracetamol in cancer of the kidney and renal pelvis. Int J Cancer 53, 245-249.

McGill, M.R., Jaeschke, H., 2013. Metabolism and disposition of acetaminophen: recent advances in relation to hepatotoxicity and diagnosis. Pharm Res 30, 2174-2187.

McGill, M.R., Jaeschke, H., 2019. Animal models of drug-induced liver injury. Biochim Biophys Acta Mol Basis Dis 1865, 1031-1039.

McGill, M.R., Lebofsky, M., Norris, H.R., Slawson, M.H., Bajt, M.L., Xie, Y., Williams, C.D., Wilkins, D.G., Rollins, D.E., Jaeschke, H., 2013. Plasma and liver acetaminophen-protein adduct levels in mice after acetaminophen treatment: dose-response, mechanisms, and clinical implications. Toxicol Appl Pharmacol 269, 240-249.

McGill, M.R., Sharpe, M.R., Williams, C.D., Taha, M., Curry, S.C., Jaeschke, H., 2012a. The mechanism underlying acetaminophen-induced hepatotoxicity in humans and mice involves mitochondrial damage and nuclear DNA fragmentation. J Clin Invest 122, 1574-1583.

McGill, M.R., Staggs, V.S., Sharpe, M.R., Lee, W.M., Jaeschke, A.H., 2014. Acute Liver Failure Study Group. Serum mitochondrial biomarkers and damage-associated molecular patterns are higher in acetaminophen overdose patients with poor outcome. . Hepatology 60, 1336-1345.

McGill, M.R., Williams, C.D., Xie, Y., Ramachandran, A., Jaeschke, H., 2012b. Acetaminophen-induced liver injury in rats and mice: comparison of protein adducts, mitochondrial dysfunction, and oxidative stress in the mechanism of toxicity. Toxicol Appl Pharmacol 264, 387-394.

McGill, M.R., Yan, H.M., Ramachandran, A., Murray, G.J., Rollins, D.E., Jaeschke, H., 2011. HepaRG cells: a human model to study mechanisms of acetaminophen hepatotoxicity. Hepatology 53, 974-982.

McGlynn, K.A., Hagberg, K., Chen, J., Graubard, B.I., London, W.T., Jick, S., Sahasrabuddhe, V.V., 2015. Statin use and risk of primary liver cancer in the Clinical Practice Research Datalink. J Natl Cancer Inst 107.

McLaughlin, J.K., Blot, W.J., Mehl, E.S., Fraumeni, J.F., Jr., 1985. Relation of analgesic use to renal cancer: population-based findings. Natl Cancer Inst Monogr 69, 217-222.

Mellemgaard, A., Niwa, S., Mehl, E.S., Engholm, G., McLaughlin, J.K., Olsen, J.H., 1994. Risk factors for renal cell carcinoma in Denmark: role of medication and medical history. Int J Epidemiol 23, 923-930.

Milam, K.M., Byard, J.L., 1985. Acetaminophen metabolism, cytotoxicity, and genotoxicity in rat primary hepatocyte cultures. Toxicol Appl Pharmacol 79, 342-347.

Miller, R.P., Roberts, R.J., Fischer, L.J., 1976. Acetaminophen elimination kinetics in neonates, children, and adults. Clin Pharmacol Ther 19, 284-294.

Miners, J.O., Bowalgaha, K., Elliot, D.J., Baranczewski, P., Knights, K.M., 2011. Characterization of niflumic acid as a selective inhibitor of human liver microsomal UDP-glucuronosyltransferase 1A9: application to the reaction phenotyping of acetaminophen glucuronidation. Drug Metab Dispos 39, 644-652.

Mitchell, J.R., Jollow, D.J., Potter, W.Z., Gillette, J.R., Brodie, B.B., 1973. Acetaminophen-induced hepatic necrosis. IV. Protective role of glutathione. J Pharmacol Exp Ther 187, 211-217.

Moysich, K.B., Bonner, M.R., Beehler, G.P., Marshall, J.R., Menezes, R.J., Baker, J.A., Weiss, J.R., Chanan-Khan, A., 2007. Regular analgesic use and risk of multiple myeloma. Leuk Res 31, 547-551.

Mugford, C.A., Tarloff, J.B., 1995. Contribution of oxidation and deacetylation to the bioactivation of acetaminophen in vitro in liver and kidney from male and female Sprague-Dawley rats. Drug Metab Dispos 23, 290-294.

Muller-Tegethoff, K., Kasper, P., Muller, L., 1995. Evaluation studies on the *in vitro* rat hepatocyte micronucleus assay. Mutat Res 335, 293-307.

Muller, L., Kasper, P., 2000. Human biological relevance and the use of threshold-arguments in regulatory genotoxicity assessment: experience with pharmaceuticals. Mutat Res 464, 19-34.

Muller, L., Kasper, P., Madle, S., 1991. Further investigations on the clastogenicity of paracetamol and acetylsalicylic acid *in vitro*. Mutat Res 263, 83-92.

Mutlib, A.E., Goosen, T.C., Bauman, J.N., Williams, J.A., Kulkarni, S., Kostrubsky, S., 2006. Kinetics of acetaminophen glucuronidation by UDP-glucuronosyltransferases 1A1, 1A6, 1A9 and 2B15. Potential implications in acetaminophen-induced hepatotoxicity. Chem Res Toxicol 19, 701-709.

Nair, A.B., Jacob, S., 2016. A simple practice guide for dose conversion between animals and human. J Basic Clin Pharm 7, 27-31.

Nakanishi, K., Kurata, Y., Oshima, M., Fukushima, S., Ito, N., 1982. Carcinogenicity of phenacetin: long-term feeding study in B6c3f1 mice. Int J Cancer 29, 439-444.

Nayak, S.S., Jain, R., Sahoo, A.K., 2011. Hepatoprotective activity of Glycosmis pentaphylla against paracetamol-induced hepatotoxicity in Swiss albino mice. Pharmaceutical biology 49, 111-117.

Nery, R., 1971a. The possible role of N-hydroxylation in the biological effects of phenacetin. Xenobiotica 1, 339-343.

Nery, R., 1971b. Some new aspects of the metabolism of phenacetin in the rat. Biochem J 122, 317-326.

Newton, J., Bailie, M., Hook, J., 1983. Acetaminophen nephrotoxicity in the rat. Renal metabolic activation in vitro. Toxicology and applied pharmacology 70, 433-444.

Newton, J.F., Braselton, W.E., Jr., Kuo, C.H., Kluwe, W.M., Gemborys, M.W., Mudge, G.H., Mudge, G.H., Hook, J.B., 1982. Metabolism of acetaminophen by the isolated perfused kidney. J Pharmacol Exp Ther 221, 76-79.

Ni, H.M., Bockus, A., Boggess, N., Jaeschke, H., Ding, W.X., 2012. Activation of autophagy protects against acetaminophen-induced hepatotoxicity. Hepatology 55, 222-232.

Ni, H.M., McGill, M.R., Chao, X., Du, K., Williams, J.A., Xie, Y., Jaeschke, H., Ding, W.X., 2016. Removal of acetaminophen protein adducts by autophagy protects against acetaminophen-induced liver injury in mice. J Hepatol 65, 354-362.

Ni, H.M., Williams, J.A., Jaeschke, H., Ding, W.X., 2013. Zonated induction of autophagy and mitochondrial spheroids limits acetaminophen-induced necrosis in the liver. Redox Biol 1, 427-432.

Nordenskjold, M., Moldeus, P., 1983. Induction of DNA-strand breaks in cultured human fibroblasts by reactive metabolites. M Annals of the NY Acad of Sci 401, 460-462.

NTP, 1993. Toxicology and Carcinogenesis Studies of Acetaminophen (CAS No.: 103-90-2) in F344/N Rats and B6C3F1 Mice (Feed Studies). National Toxicology Program Technical Report Series: 394. NIH Publication No. 93-2849. Jan., 1993, 2007/03/08 ed. National Toxicology Program (NTP), National Institutes of Health, Public Health Service, U.S. Department of Health and Human Services Research Trianlge Park, NC.

NTP, 1994. Toxicology Data Management System - Tumor Incidence in Control Animals by Route and Vehicle of Administration - F344/N Rats. National Toxicology Program, National Institutes of Health, Public Health Service, U.S. Dept. of Health and Human Services, Research Triangle Park, N.C.

NTP, 2011. Specifications for the Conduct of Studies to Evaluate the Toxic and Carcinogenic Potential of Chemical, Biological and Physical Agents in Laboratory Animals for the National Toxicology Program (NTP). National Institutes of Health, Public Health Service, U.S. Dept. of Health and Human Services, Research Triangle Park, N.C.

NTP, 2019. Organ Sites with Neoplasia.

OECD, 1986a. Genetic Toxicology: *In vitro* Sister Chromatid Exchange Assay in Mammalian Cells (479). Organisation for Economic Co-operation and Development.

OECD, 1986b. Genetic Toxicology: DNA Damage and Repair/Unscheduled DNA Synthesis in Mammalian Cells in vitro (482). Organisation for Economic Co-operation and Development.

OECD, 2012. Guidance Document 116 on the Conduct and Design of Chronic Toxicity and Carcinogenicity Studies, Supporting Test Guidelines 451, 452 And 453, Series on Testing and Assessment. Organisation for Economic Co-operation and Development.

OECD, 2015. Genetic Toxicology Guidance Document, Organisation for Economic Co-operation and Development, Paris, adopted August 31st.

OECD, 2016a. In Vitro Mammalian Chromosomal Aberration Test (473). Organisation for Economic Co-operation and Development.

OECD, 2016b. In Vivo Mammalian Alkaline Comet Assay (489). Organisation for Economic Co-operation and Development.

OECD, 2016c. Mammalian Bone Marrow Chromosomal Aberration Test (475). Organisation for Economic Co-operation and Development.

OECD, 2016d. Mammalian Erythrocyte Micronucleus Test (474). Organisation for Economic Co-operation and Development.

OECD, 2017. Overview of the set of OECD Genetic Toxicology Test Guidelines and updates performed in 2014-2015 (238). Organisation for Economic Co-operation and Development.

OECD, 2018. OECD Guideline for the testing of chemicals: carcinogenicity studies. TG-451. . Organisation for Economic Co-operation and Development.

OEHHA, 2001. Guidance Criteria for Identifying Chemicals for Listing as "Known to the State to Cause Cancer". California Offices of Environmental Health Hazard Assessment (OEHHA)/California Environmental Protection Agency, Sacramento, CA

OEHHA, 2019. Proposition 65 - Evidence on the Carcinogenicity of Acetaminophen. Reproductive and Cancer Hazard Assessment Branch Office of Environmental Health Hazard Assessment California Environmental Protection Agency, Sacramento, CA p. 373.

Ognjanovic, S., Blair, C., Spector, L.G., Robison, L.L., Roesler, M., Ross, J.A., 2011. Analgesic use during pregnancy and risk of infant leukaemia: a Children's Oncology Group study. Br J Cancer 104, 532-536.

Oldham, J.W., Preston, R.F., Paulson, J.D., 1986. Mutagenicity testing of selected analgesics in Ames Salmonella strains. J Appl Toxicol 6, 237-243.

Oshida, K., Iwanaga, E., Miyamoto-Kuramitsu, K., Miyamoto, Y., 2008. An *in vivo* comet assay of multiple organs (liver, kidney and bone marrow) in mice treated with methyl methanesulfonate and acetaminophen accompanied by hematology and/or blood chemistry. J Toxicol Sci 33, 515-524.

Pacifici, G.M., 2004. Inhibition of human liver and duodenum sulfotransferases by drugs and dietary chemicals: a review of the literature. Int J Clin Pharmacol Ther 42, 488-495.

Pascoe, G.A., Calleman, C.J., Baille, T.A., 1988. Identification of S-(2,5-dihydroxyphenyl)-cysteine and S-(2,5-dihydroxyphenyl)-N-acetyl-cysteine as urinary metabolites of acetaminophen in the mouse. Evidence for p-benzoquinone as a reactive intermediate in acetaminophen metabolism. Chem Biol Interact 68, 85-98.

Patierno, S.R., Lehman, N.L., Henderson, B.E., Landolph, J.R., 1989. Study of the ability of phenacetin, acetaminophen, and aspirin to induce cytotoxicity, mutation, and morphological transformation in C3H/10T1/2 clone 8 mouse embryo cells. Cancer Res 49, 1038-1044.

Patten, C.J., Thomas, P.E., Guy, R.L., Lee, M., Gonzalez, F.J., Guengerich, F.P., Yang, C.S., 1993. Cytochrome P450 enzymes involved in acetaminophen activation by rat and human liver microsomes and their kinetics. Chem Res Toxicol 6, 511-518.

Petkov, P.I., Patlewicz, G., Schultz, T.W., Honma, M., Todorov, M., Kotov, S., Dimitrov, S.D., Donner, E.M., Mekenyan, O.G., 2015. A feasibility study: Can information collected to classify for mutagenicity be informative in predicting carcinogenicity? Regul Toxicol Pharmacol 72, 17-25.

Phang, W., Rinde, E., 1993. US EPA Second and Third Carcinogenicity Peer Reviews of Fosetyl-Al (Aliette). U.S. Environmental Protection Agency, Washington, DC.

Phillips, D.H., Farmer, P.B., Beland, F.A., Nath, R.G., Poirier, M.C., Reddy, M.V., Turteltaub, K.W., 2000. Methods of DNA adduct determination and their application to testing compounds for genotoxicity. Environ Mol Mutagen 35, 222-233.

Piper, J.M., Tonascia, J., Matanoski, G.M., 1985. Heavy phenacetin use and bladder cancer in women aged 20 to 49 years. N Engl J Med 313, 292-295.

Plattner, S., Erb, R., Pitterl, F., Brouwer, H.J., Oberacher, H., 2012. Formation and characterization of covalent guanosine adducts with electrochemistry-liquid chromatography-mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci 883-884, 198-204.

Pommer, W., Bronder, E., Klimpel, A., Helmert, U., Greiser, E., Molzahn, M., 1999. Urothelial cancer at different tumour sites: role of smoking and habitual intake of analgesics and laxatives. Results of the Berlin Urothelial Cancer Study. Nephrol Dial Transplant 14, 2892-2897.

Portmann, B., Talbot, I.C., Day, D.W., Davidson, A.R., Murray-Lyon, I.M., Williams, R., 1974. Histopathological changes in the liver following a paracetamol overdose: correlation with clinical and biochemical parameters. . J Pathol 117, 169-181.

Potter, D.W., Hinson, J.A., 1987a. The 1- and 2-electron oxidation of acetaminophen catalyzed by prostaglandin H synthase. J Biol Chem 262, 974-980.

Potter, D.W., Hinson, J.A., 1987b. Mechanisms of acetaminophen oxidation to N-acetyl-P-benzoquinone imine by horseradish peroxidase and cytochrome P-450. J Biol Chem 262, 966-973.

Potter, D.W., Hinson, J.A., 1989. Acetaminophen peroxidation reactions. Drug Metab Rev 20, 341-358.

Powell, C.L., Kosyk, O., Ross, P.K., Schoonhoven, R., Boysen, G., Swenberg, J.A., Heinloth, A.N., Boorman, G.A., Cunningham, M.L., Paules, R.S., Rusyn, I., 2006. Phenotypic anchoring of acetaminophen-induced oxidative stress with gene expression profiles in rat liver. Toxicol Sci 93, 213-222.

Prasad, J.S., Chen, N.Q., Liu, Y.X., Goon, D.J., Holtzman, J.L., 1990. Effects of ethanol and inhibitors on the binding and metabolism of acetaminophen and N-acetyl-p-benzoquinone imine by hepatic microsomes from control and ethanol-treated rats. Biochem Pharmacol 40, 1989-1995.

Prescott, L., 1996. Paracetamol (Acetaminophen): A Critical Bibliographic Review, USA.

Prescott, L.F., 2000. Paracetamol, alcohol and the liver. Br J Clin Pharmacol 49, 291-301.

Qiu, Y., Benet, L.Z., Burlingame, A.L., 1998. Identification of the hepatic protein targets of reactive metabolites of acetaminophen in vivo in mice using two-dimensional gel electrophoresis and mass spectrometry. J Biol Chem 273, 17940-17953.

Radi, R., 2004. Nitric oxide, oxidants, and protein tyrosine nitration. Proc Natl Acad Sci U S A 101, 4003-4008.

Ramachandran, A., Jaeschke, H., 2018. Acetaminophen Toxicity: Novel Insights Into Mechanisms and Future Perspectives. Gene Expr 18, 19-30.

Ramachandran, A., Jaeschke, H., 2019. Acetaminophen Hepatotoxicity. Semin Liver Dis 39, 221-234.

Ramachandran, A., Lebofsky, M., Weinman, S.A., Jaeschke, H., 2011. The impact of partial manganese superoxide dismutase (SOD2)-deficiency on mitochondrial oxidant stress, DNA fragmentation and liver injury during acetaminophen hepatotoxicity. Toxicol Appl Pharmacol 251, 226-233.

Rashed, M.S., Myers, T.G., Nelson, S.D., 1990. Hepatic protein arylation, glutathione depletion, and metabolite profiles of acetaminophen and a non-hepatotoxic regioisomer, 3'-hydroxyacetanilide, in the mouse. Drug Metab Dispos 18, 765-770.

Raucy, J.L., Lasker, J.M., Lieber, C.S., Black, M., 1989. Acetaminophen activation by human liver cytochromes P450IIE1 and P450IA2. Arch Biochem Biophys 271, 270-283.

Reddy, G.A., 1984. Effects of Paracetamol on Chromosomes of Mouse Bone Marrow. Caryologia 37, 127-132.

Reddy, G.A., Subramanyam, S., 1985. Cytogenetic Response of Meiocytes of Swiss Albino Mice to Paracetamol. Caryologia 38, 347-355.

Registre, M., Proudlock, R., 2016. The In Vitro Chromosome Aberration Test, in: Proudlock, R. (Ed.), Genetic Toxicology Testing: A Laboratory Manual. Academic Press, pp. 207-267.

Reznikoff, C.A., Brankow, D.W., Heidelberger, C., 1973. Establishment and characterization of a cloned line of C3H mouse embryo cells sensitive to postconfluence inhibition of division. Cancer research 33, 3231-3238.

Rogers, L.K., Moorthy, B., Smith, C.V., 1997. Acetaminophen binds to mouse hepatic and renal DNA at human therapeutic doses. Chem Res Tox 10, 470-476.

Rosenberg, L., Rao, R.S., Palmer, J.R., Strom, B.L., Zauber, A., Warshauer, M.E., Stolley, P.D., Shapiro, S., 1998. Transitional cell cancer of the urinary tract and renal cell cancer in relation to acetaminophen use (United States). Cancer Causes Control 9, 83-88.

Ross, J.A., Blair, C.K., Cerhan, J.R., Soler, J.T., Hirsch, B.A., Roesler, M.A., Higgins, R.R., Nguyen, P.L., 2011. Nonsteroidal anti-inflammatory drug and acetaminophen use and risk of adult myeloid leukemia. Cancer Epidemiol Biomarkers Prev 20, 1741-1750.

Ross, R.K., Paganini-Hill, A., Landolph, J., Gerkins, V., Henderson, B.E., 1989. Analgesics, cigarette smoking, and other risk factors for cancer of the renal pelvis and ureter. Cancer Res 49, 1045-1048.

Rothman, K.J., Greenland, S., Lash, T.L., 2008. Modern Epidemiology, Third ed. Lippincott Williams & Wilkins;.

Rumack, B.H., 2004. Acetaminophen misconceptions. Hepatology 40, 10-15.

Saito, T., Takahashi, H., Doken, H., Koyama, H., Aratani, Y., 2005. Phorbol myristate acetate induces neutrophil death through activation of p38 mitogen-activated protein kinase that requires endogenous reactive oxygen species other than HOCl. Biosci Biotechnol Biochem 69, 2207-2212.

Sarich, T., Kalhorn, T., Magee, S., al-Sayegh, F., Adams, S., Slattery, J., Goldstein, J., Nelson, S., Wright, J., 1997. The effect of omeprazole pretreatment on acetaminophen metabolism in rapid and slow metabolizers of S-mephenytoin. Clin Pharmacol Ther 62, 21-28.

Sasaki, M., 1986. Enhancing effect of acetaminophen on mutagenesis. Prog Clin Biol Res 209A, 365-372.

Sasaki, M., Sugimura, K., Yoshida, M., Abe, S., 1980. Cytogenetic effects of 60 chemicals on cultured human and chinese hamster cells. Kromosomo 2, 574-584.

Sasaki, M., Yoshida, S., Hiraga, K., 1983. Additional effect of acetaminophen on the mutagenicity and clastogenicity of N-methyl-N'-nitro-N-nitrosoguanidine in cultured Chinese hamster CHO-Kl cells. Mutat Res 122, 367-372.

Sawada, M., 1985. Mutagenicity tests on phenacetin-related compounds in cultured mammalian cells. Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis 147, 273.

Schildknecht, S., Daiber, A., Ghisla, S., Cohen, R.A., Bachschmid, M.M., 2008. Acetaminophen inhibits prostanoid synthesis by scavenging the PGHS-activator peroxynitrite. FASEB J 22, 215-224.

Schmid, E., Bauchinger, M., Hauf, R., 1985. Chromosome changes with time in lymphocytes after occupational exposure to toluene. Mutat Res 142, 37-39.

Scott, D., Galloway, S.M., Marshall, R.R., Ishidate, M., Jr., Brusick, D., Ashby, J., Myhr, B.C., 1991. International Commission for Protection Against Environmental Mutagens and Carcinogens. Genotoxicity under extreme culture conditions. A report from ICPEMC Task Group 9. Mutat Res 257, 147-205.

Severin, E., Beleuta, A., 1995. Induction of chromosome aberrations *in vivo* in bone-marrow cells of mice by paracetamol. Rom J Morphol Embryol 41, 117-120.

Sharma, C.V., Mehta, V., 2014. Paracetamol: mechanisms and updates. Continuing Education in Anaesthesia Critical Care & Pain 14, 153-158.

Shertzer, H.G., Schneider, S.N., Kendig, E.L., Clegg, D.J., D'Alessio, D.A., Genter, M.B., 2008. Acetaminophen normalizes glucose homeostasis in mouse models for diabetes. Biochem Pharmacol 75, 1402-1410.

Shibata, M.A., Sano, M., Hagiwara, A., Hasegawa, R., Shirai, T., 1995. Modification by analgesics of lesion development in the urinary tract and various other organs of rats pretreated with dihydroxy-di-N-propylnitrosamine and uracil. Jpn J Cancer Res 86, 160-167.

Shimane, Y., 1985. [Mutagenicity of acetaminophen on Chinese hamster V79 cells]. Shigaku 72, 1175-1187.

Shirai, T., Ikawa, E., Fukushima, S., Masui, T., Ito, N., 1986. Uracil-induced urolithiasis and the development of reversible papillomatosis in the urinary bladder of F344 rats. Cancer Res 46, 2062-2067.

Shirai, T., Shibata, M., Takahashi, S., Tagawa, Y., Imaida, K., Hirose, M., 1995. Differences in cell proliferation and apoptosis between reversible and irreversible mucosal lesions associated with uracil-induced urolithiasis in N-butyl-N-(4-hydroxybutyl)nitrosamine-pretreated. Carcinogenesis 16, 501-505.

Shudo, K., Ohta, T., Orihara, Y., Okamoto, T., Nagao, M., Takahashi, Y., Sugimura, T., 1978. Chemical reactivity and mutagenicity of N-OH-phenacetin. Proc. Jpn. Cancer Assoc., 37th Annu. Meet. 44 (Japanese).

Sicardi, S.M., Martiarena, J.L., Iglesias, M.T., 1991. Mutagenic and analgesic activities of aniline derivatives. J Pharm Sci 80, 761-764.

Simkó, M., Kriehuber, R., Lange, S., 1998. Micronucleus formation in human amnion cells after exposure to 50 Hz MF applied horizontally and vertically. Mutation Research/Genetic Toxicology and Environmental Mutagenesis 418, 101-111.

Slattery, J.T., McRorie, T.I., Reynolds, R., Kalhorn, T.F., Kharasch, E.D., Eddy, A.C., 1989. Lack of effect of cimetidine on acetaminophen disposition in humans. Clin Pharmacol Ther 46, 591-597.

Slezak, A., Moreira, H., Szyjka, A., Oszmianski, J., Gasiorowski, K., 2017. Conditions of Prooxidant Activity of Cistus and Pomegranate Polyphenols in V79 Cell Cultures. Acta Pol Pharm 74, 670-678.

Smarr, M.M., Kannan, K., Chen, Z., Kim, S., Buck Louis, G.M., 2017. Male urinary paracetamol and semen quality. Andrology 5, 1082-1088.

Smith, M.T., Guyton, K.Z., Gibbons, C.F., Fritz, J.M., Portier, C.J., Rusyn, I., DeMarini, D.M., Caldwell, J.C., Kavlock, R.J., Lambert, P.F., Hecht, S.S., Bucher, J.R., Stewart, B.W., Baan, R.A., Cogliano, V.J., Straif, K., 2016. Key Characteristics of Carcinogens as a Basis for Organizing Data on Mechanisms of Carcinogenesis. Environ Health Perspect 124, 713-721.

Snawder, J.E., Roe, A.L., Benson, R.W., Roberts, D.W., 1994. Loss of CYP2E1 and CYP1A2 activity as a function of acetaminophen dose: relation to toxicity. Biochem Biophys Res Commun 203, 532-539.

Šrám, R.J., Kocisova, J., Rossner, P., Binkova, B., Topinka, J., Bavorova, H., 1990. Mutagenic activity of paracetamol. A study conducted on volunteers. Czech Med 13, 114-123.

Stamper, B.D., 2015. Transcriptional profiling of reactive metabolites for elucidating toxicological mechanisms: a case study of quinoneimine-forming agents. Drug Metab Rev 47, 45-55.

Steineck, G., Wiholm, B.E., Gerhardsson de Verdier, M., 1995. Acetaminophen, some other drugs, some diseases and the risk of transitional cell carcinoma. A population-based case-control study. Acta Oncol 34, 741-748.

Suzuki, Y., Goto, K., Nakayama, Y., Saratani, M., Takata, T., Okamoto, T., Okazaki, S., 2016. Evaluation of a single-dose PIGRET assay for acetaminophen in rats compared with the RBC Pig-a assay. Mutat Res 811, 16-20.

Takehara, M., Hoshino, T., Namba, T., Yamakawa, N., Mizushima, T., 2011. Acetaminophen-induced differentiation of human breast cancer stem cells and inhibition of tumor xenograft growth in mice. Biochem Pharmacol 81, 1124-1135.

Taylor, J.S., 1972. Carcinoma of the urinary tract and analgesic abuse. Med J Aust 1, 407-409.

Tee, L.B., Davies, D.S., Seddon, C.E., Boobis, A.R., 1987. Species differences in the hepatotoxicity of paracetamol are due to differences in the rate of conversion to its cytotoxic metabolite. Biochem Pharmacol 36, 1041-1052.

Thomas, J., Haseman, J.K., Goodman, J.I., Ward, J.M., Loughran, T.P., Jr., Spencer, P.J., 2007. A review of large granular lymphocytic leukemia in Fischer 344 rats as an initial step toward evaluating the implication of the endpoint to human cancer risk assessment. Toxicol Sci 99, 3-19.

Thomas, S.H., 1993. Paracetamol (acetaminophen) poisoning. Pharmacol Ther 60, 91-120.

Thummel, K.E., Lee, C.A., Kunze, K.L., Nelson, S.D., Slattery, J.T., 1993. Oxidation of acetaminophen to N-acetyl-p-aminobenzoquinone imine by human CYP3A4. Biochem Pharmacol 45, 1563-1569.

Thummel, K.E., Slattery, J.T., Nelson, S.D., 1988. Mechanism by which ethanol diminishes the hepatotoxicity of acetaminophen. J Pharmacol Exp Ther 245, 129-136.

Thybaud, V., Aardema, M., Clements, J., Dearfield, K., Galloway, S., Hayashi, M., Jacobson-Kram, D., Kirkland, D., MacGregor, J.T., Marzin, D., Ohyama, W., Schuler, M., Suzuki, H., Zeiger, E., Expert Working Group on Hazard, I., Risk Assessment in Relation to In Vitro, T., 2007. Strategy for genotoxicity testing: hazard identification and risk assessment in relation to in vitro testing. Mutat Res 627, 41-58.

Tirmenstein, M.A., Nelson, S.D., 1989. Subcellular binding and effects on calcium homeostasis produced by acetaminophen and a nonhepatotoxic regioisomer, 3'-hydroxyacetanilide, in mouse liver. J Biol Chem 264, 9814-9819.

Tomatis, L., 1979. The predictive value of rodent carcinogenicity tests in the evaluation of human risks. Annu Rev Pharmacol Toxicol 19, 511-530.

Topinka, J., Sram, R.J., Sirinjan, G., Kocisova, J., Binkova, B., Fojtikova, I., 1989. Mutagenicity studies on paracetamol in human volunteers. II. Unscheduled DNA synthesis and micronucleus test. Mutat Res 227, 147-152.

Tsuda, H., Sakata, T., Masui, T., Imaida, K., Ito, N., 1984. Modifying effects of butylated hydroxyanisole, ethoxyquin and acetaminophen on induction of neoplastic lesions in rat liver and kidney initiated by N-ethyl-N-hydroxyethylnitrosamine. Carcinogenesis 5, 525-531.

Tsuruzaki, T., Watanabe, G., Yamamoto, M., 1982. [The effects of aspirin and acetaminophen during pregnancy on chromosomal structure in rat fetuses]. Nihon Eiseigaku Zasshi 37, 787-796.

van de Pol, J.A.A., van den Brandt, P.A., Schouten, L.J., 2019. Kidney stones and the risk of renal cell carcinoma and upper tract urothelial carcinoma: the Netherlands Cohort Study. Br J Cancer 120, 368-374.

van der Leede, B.M., Weiner, S., Van Doninck, T., De Vlieger, K., Schuermans, A., Tekle, F., Geys, J., van Heerden, M., de Jonghe, S., Van Gompel, J., in press. Testing of acetaminophen in support of the international multi-laboratory in vivo rat Pig-a assay validation trial, . Environ. Mol. Mutagen.

van der Marel, C.D., Anderson, B.J., van Lingen, R.A., Holford, N.H., Pluim, M.A., Jansman, F.G., van den Anker, J.N., Tibboel, D., 2003. Paracetamol and metabolite pharmacokinetics in infants. Eur J Clin Pharmacol 59, 243-251.

van Rongen, A., Valitalo, P.A.J., Peeters, M.Y.M., Boerma, D., Huisman, F.W., van Ramshorst, B., van Dongen, E.P.A., van den Anker, J.N., Knibbe, C.A.J., 2016. Morbidly Obese Patients Exhibit Increased CYP2E1-Mediated Oxidation of Acetaminophen. Clin Pharmacokinet 55, 833-847.

Walker, A., Patrick, A., Lauer, M., Hornbrook, M., Marin, M., Platt, R., Roger, V., Stang, P., Schneeweiss, S., 2016. A tool for assessing the feasibility of comparative effectiveness research. Comparative Effectiveness Research 3, 11-20.

Walter, R.B., Brasky, T.M., White, E., 2011a. Cancer risk associated with long-term use of acetaminophen in the prospective VITamins and lifestyle (VITAL) study. Cancer Epidemiol Biomarkers Prev 20, 2637-2641.

Walter, R.B., Milano, F., Brasky, T.M., White, E., 2011b. Long-term use of acetaminophen, aspirin, and other nonsteroidal anti-inflammatory drugs and risk of hematologic malignancies: results from the prospective Vitamins and Lifestyle (VITAL) study. J Clin Oncol 29, 2424-2431.

Wan, J., Bae, M.A., Song, B.J., 2004. Acetoaminophen-induced accumulation of 8-oxodeoxyguanosine through reduction of Ogg1 DNA repair enzyme in C6 glioma cells. Exp Mol Med 36, 71-77.

Wang, C., Wu, M., Arvapalli, R., Dai, X., Mahmood, M., Driscoll, H., Rice, K.M., Blough, E., 2014. Acetaminophen attenuates obesity-related renal injury through ER-mediated stress mechanisms. Cell Physiol Biochem 33, 1139-1148.

Wang, Y., Li, D., Cheng, N., Gao, H., Xue, X., Cao, W., Sun, L., 2015. Antioxidant and hepatoprotective activity of vitex honey against paracetamol induced liver damage in mice. Food Funct 6, 2339-2349.

Watanabe, G., 1982. The Cytogenetic Effects of Aspirin and Acetaminophen on in Vitro Human Lymphocytes [Japanese]. Nippon Eiseigaku Zasshi 37, 673-685.

Weinstein, R.B., Ryan, P., Berlin, J.A., Matcho, A., Schuemie, M., Swerdel, J., Patel, K., Fife, D., 2017. Channeling in the Use of Nonprescription Paracetamol and Ibuprofen in an Electronic Medical Records Database: Evidence and Implications. Drug Saf 40, 1279-1292.

Weisburger, J.H., Weisburger, E.K., Madison, R.M., Wenk, M.L., Klein, D.S., 1973. Effect of acetanilide and p-hydroxyacetanilide on the carcinogenicity of N-2-fluorenylacetamide and N-

hydroxy-N-2-fluorenylacetamide in mice, hamsters, and female rats. J Natl Cancer Inst 51, 235-240.

Weiss, J.R., Baker, J.A., Baer, M.R., Menezes, R.J., Nowell, S., Moysich, K.B., 2006. Opposing effects of aspirin and acetaminophen use on risk of adult acute leukemia. Leuk Res 30, 164-169.

Weiss, N.S., 2016. Use of acetaminophen in relation to the occurrence of cancer: a review of epidemiologic studies. Cancer Causes Control 27, 1411-1418.

West, P.R., Harman, L.S., Josephy, P.D., Mason, R.P., 1984. Acetaminophen: enzymatic formation of a transient phenoxyl free radical. Biochem Pharmacol 33, 2933-2936.

West, S.L., Savitz, D.A., Koch, G., Sheff, K.L., Strom, B.L., Guess, H.A., Hartzema, A.G., 1997. Demographics, health behaviors, and past drug use as predictors of recall accuracy for previous prescription medication use. J Clin Epidemiol 50, 975-980.

West, S.L., Savitz, D.A., Koch, G., Strom, B.L., Guess, H.A., Hartzema, A., 1995. Recall accuracy for prescription medications: self-report compared with database information. Am J Epidemiol 142, 1103-1112.

Wikoff, D.S., Rager, J.E., Chappell, G.A., Fitch, S., Haws, L., Borghoff, S.J., 2019. A Framework for Systematic Evaluation and Quantitative Integration of Mechanistic Data in Assessments of Potential Human Carcinogens. Toxicol Sci 167, 322-335.

Williams, G.M., Iatropoulos, M.J., 1997. Inhibition by acetaminophen of intestinal cancer in rats induced by an aromatic amine similar to food mutagens. Eur J Cancer Prev 6, 357-362.

Williams, G.M., Iatropoulos, M.J., Jeffrey, A.M., Duan, J.D., Perrone, C.E., 2007. Inhibition by acetaminophen of neoplastic initiation elicited in rat liver by the DNA-reactive hepatocarcinogen N-acetyl-2-aminofluorene. Eur J Cancer Prev 16, 528-534.

Win, S., Than, T.A., Min, R.W., Aghajan, M., Kaplowitz, N., 2016. c-Jun N-terminal kinase mediates mouse liver injury through a novel Sab (SH3BP5)-dependent pathway leading to inactivation of intramitochondrial Src. Hepatology 63, 1987-2003.

Winnenburg, R., Sorbello, A., Bodenreider, O., 2015. Exploring adverse drug events at the class level. J Biomed Semantics 6, 18.

Wirth, P.J., Dybing, E., Von Bahr, C., Thorgeirsson, S.S., 1980. Mechanism of N-hydroxyacetylarylamine mutagenicity in the Salmonella test system: Metabolic activation of N-hydroxyphenacetin by liver and kidney fractions from rat, mouse, hamster, and man. Molecular pharmacology 18, 117-127.

Wright, H.N., 1967. Chronic toxicity studies of analgesic and antipyretic drugs and congeners. Tox Appl Pharm 11, 280-292.

Xie, Y., McGill, M.R., Cook, S.F., Sharpe, M.R., Winefield, R.D., Wilkins, D.G., Rollins, D.E., Jaeschke, H., 2015a. Time course of acetaminophen-protein adducts and acetaminophen metabolites in circulation of overdose patients and in HepaRG cells. Xenobiotica 45, 921-929.

Xie, Y., McGill, M.R., Dorko, K., Kumer, S.C., Schmitt, T.M., Forster, J., Jaeschke, H., 2014. Mechanisms of acetaminophen-induced cell death in primary human hepatocytes. Toxicol Appl Pharmacol 279, 266-274.

Xie, Y., McGill, M.R., Du, K., Dorko, K., Kumer, S.C., Schmitt, T.M., Ding, W.X., Jaeschke, H., 2015b. Mitochondrial protein adducts formation and mitochondrial dysfunction during N-acetyl-m-aminophenol (AMAP)-induced hepatotoxicity in primary human hepatocytes. Toxicol Appl Pharmacol 289, 213-222.

Yan, H.M., Ramachandran, A., Bajt, M.L., Lemasters, J.J., Jaeschke, H., 2010. The oxygen tension modulates acetaminophen-induced mitochondrial oxidant stress and cell injury in cultured hepatocytes. Toxicol Sci 117, 515-523.

Yang, B., Petrick, J.L., Chen, J., Hagberg, K.W., Sahasrabuddhe, V.V., Graubard, B.I., Jick, S., McGlynn, K.A., 2016. Associations of NSAID and paracetamol use with risk of primary liver cancer in the Clinical Practice Research Datalink. Cancer Epidemiol 43, 105-111.

Zapater, P., Lasso de la Vega, M.C., Horga, J.F., Such, J., Frances, R., Esteban, A., Palazon, J.M., Carnicer, F., Pascual, S., Perez-Mateo, M., 2004. Pharmacokinetic variations of acetaminophen according to liver dysfunction and portal hypertension status. Aliment Pharmacol Ther 20, 29-36.