PROTOCOL FOR A SYSTEMATIC REVIEW TO EVALUATE THE EVIDENCE FOR AN ASSOCIATION BETWEEN OCCUPATONAL EXPOSURE TO CANCER CHEMOTHERAPY AGENTS AND ADVERSE HEALTH OUTCOMES

Project Leader: Kembra Howdeshell, PhD, Office of Health Assessment and Translation (OHAT), DNTP

Summary: OHAT is conducting a systematic review to evaluate the evidence for an association between occupational exposure to cancer chemotherapy agents and adverse health outcomes. The protocol is detailed in this document.

BACKGROUND AND SIGNIFICANCE

Background

Many cancer chemotherapy agents are cytotoxic drugs, and many of these agents are known mutagens and/or developmental toxicants (2013). Occupational exposure to cancer chemotherapy agents may occur in medical, veterinary, and manufacturing settings among personnel involved in the production, preparation and administration of these agents (Couch and West 2012, Couch *et al.* 2013, Kiffmeyer *et al.* 2013, Kopp *et al.* 2013) as well as other workers involved with the care of patients administered chemotherapy (Hon *et al.* 2014). Potential routes of occupational exposure may include dermal, ingestion, and inhalation. While levels of such exposures are thought to be much lower than those administered to cancer patients, occupational exposure likely involves more than one chemotherapy agent or specific combination therapy, and it may occur more frequently and over a longer period of time. Furthermore, occupational exposures are often unrecognized due to lack of systematic environmental monitoring and biomonitoring programs (OSHA 1999).

Evidence for exposure began appearing in the 1970s with reports of elevated mutagenic activity in the urine of health care workers who prepared and administered such agents (reviewed in Connor and McDiarmid (2006)). Subsequent studies reported elevated levels of biomarkers of exposure such as chromosome aberrations, sister chromatid exchanges, and DNA damage in workers handling these agents, as well as direct identification of chemotherapy agents or their metabolites in workers' urine. The monitoring of workplace contamination was implemented following the establishment of guidelines for safe handling of hazardous drugs in the 1980s and 1990s by national health care worker agencies in multiple countries, including the Occupational Safety and Health Administration in the United States (OSHA 1999). Beginning in the 1990s, numerous publications have documented surface contamination of safety cabinets, countertops, floors, and equipment with chemotherapy agents. While improved handling procedures and engineering controls have reduced contamination, Connor et al. (2012) reported that surface contamination persists in pharmacy and nursing areas of some hospital-based cancer centers. In addition, potential occupational exposure to cancer chemotherapy agents has increased with: (1) greater usage of chemotherapy for non-cancer disease conditions and (2) the development of new surgical techniques involving administration of chemotherapy directly into the peritoneal cavity (Villa *et al.* 2015).

The association between occupational exposure to cancer chemotherapy and adverse health effects was evaluated in two systematic reviews as well as a recent literature review by Connor *et al.* (2014). A systematic review and meta-analysis by Dranitsaris (2005) evaluated the literature on cancer, pregnancy outcomes, and acute toxic effects in nurses, pharmacists, or pharmacy technicians/assistants who work with cytotoxic drugs. They reported that there was insufficient literature to reach conclusions on cancer, acute effects, congenital malformations or stillbirth; however, they did find a significant association between occupational exposure and spontaneous abortions (Dranitsaris *et al.* 2005). The other systematic review (Quansah and Jaakkola 2010) and the recent literature review by Connor *et al.* (2014) evaluated the association between occupational exposure to chemotherapy agents in nurses or healthcare workers, respectively, and the adverse pregnancy outcomes. Both reviews concluded that occupational exposure was associated with an increased incidence of congenital malformations and spontaneous abortions, but that the significance of the findings were limited by low sample size and heterogeneity in the study designs (Quansah and Jaakkola 2010, Connor *et al.* 2014). Because reproductive health outcomes have been addressed in recent reviews, OHAT is focusing its systematic review of the evidence for an association between occupational exposure to cancer chemotherapy and any non-reproductive health outcomes, including, but not limited to,: cancer, immune system effects, and acute effects. The review will also include an evaluation of the association between occupational exposure to cancer chemotherapy agents and biomarkers of effect (specifically, structural chromosomal aberrations, DNA damage as measured by Comet assay, and micronuclei induction). These three assays test for a broad range of types of DNA damage induced by drugs or chemicals that induce genetic toxicity. Chromosomal aberrations and micronuclei induction have been identified as good predictors of future cancer incidence (Bonassi *et al.* 2004, Norppa *et al.* 2006, Bonassi *et al.* 2007). The Comet assay is becoming more widely used in human biomonitoring to measure DNA damage as a biomarker of effect following exposure to chemical known to induce genetic toxicity (Collins *et al.* 2014). However, the causality of DNA damage detected by Comet assay to subsequently induce cancer has yet to be determined (reviewed in Collins *et al.* (2014)).

Significance

This OHAT evaluation will complement the recent NIOSH review of reproductive health (Connor *et al.* 2014) by evaluating the non-reproductive adverse health outcomes associated with occupational exposure to cancer chemotherapy agents. The OHAT evaluation will use our recently developed systematic review and evidence integration methodology that involves a rigorous evaluation of risk of bias of each included study, and an assessment a range of additional factors over the body of evidence, to rate our confidence in the literature. This evaluation will review all relevant published studies on adverse health outcomes, biomarkers of effect, and organ system function in humans, and published in English and non-English language. The results of the evaluation will inform recommendations to protect worker health as well as identification of research gaps and data needs to better understand the adverse health effects associated with direct or indirect handling of cancer chemotherapy agents in the occupational setting. This review will also inform advice to family and friends who act as caregivers of patients administered cancer chemotherapy for medical conditions. Finally, data management will be conducted in a manner that permits public sharing of the data extracted from included studies.

OVERALL OBJECTIVE AND SPECIFIC AIMS

The overall objective of this evaluation is to develop hazard identification conclusions about whether occupational exposure to cancer chemotherapy is associated with adverse health outcomes (e.g.,

cancer, immune system effects, and acute effects) and related health effects (e.g., genetic toxicity) by considering evidence from human studies.

Specific aims:

- Identify literature reporting the effects of occupational exposure to cancer chemotherapy agents on adverse health outcomes, including primary health outcomes (e.g., cancer, immune effects, acute effects, kidney and liver toxicity, *etc*.) and secondary health outcomes (e.g., genetic biomarkers of effects (specifically, chromosomal aberrations, DNA damage, and micronuclei induction), immune function assays, liver and kidney function markers, *etc*.) in human studies
- Extract data on potential health effects from relevant studies (data extraction files of the included studies will be shared upon release of final report)
- Assess the internal validity ("risk of bias") of individual studies using pre-defined criteria
- Synthesize the evidence using a narrative approach or meta-analysis (if appropriate) considering limits on data integration (i.e., heterogeneity, sample size, *etc.*)
- Rate confidence in the body of evidence for human studies according to one of four statements: (1) High, (2) Moderate, (3) Low, or (4) Very Low, or No Evidence Available
- Translate confidence ratings into level of evidence of health effects for each type of health outcome, separately, for human studies according to one of four statements: (1) High, (2) Moderate, (3) Low, or (4) Inadequate
- Use the level of evidence ratings for human health effects data and consider the degree of support from other sources (e.g., for cancer: IARC, RoC) to reach one of five possible hazard identification categories: (1) Known, (2) Presumed, (3) Suspected, (4) Not Classifiable, or (5) Not Identified to be a Hazard to Humans

To address our overall objective we developed a PECO statement (**P**opulation, **E**xposure(s), **C**omparator(s), and **O**utcome(s)) (**Table 1. Population, Exposure, Comparator, and Outcome (PECO) Statement**), which is used as an aid to develop the evaluation question, develop the search terms, and the inclusion/exclusion criteria for our systematic review (Higgins and Green 2011, AHRQ 2014).

Table 1. Population, Exposure, Comparator, and Outcome (PECO) Statement	
PECO Element	**Evidence**
Population	Men and women who come into contact with cancer chemotherapy agents in their workplace
Exposure	Cancer chemotherapy agents, including anti-neoplastic agents, synthetic hormones (e.g., tamoxifen), monoclonal antibodies and other targeted therapies (e.g., imatinib)
Comparator	A comparison population exposed to lower levels (or no exposure/exposure below detection levels) of cancer chemotherapy agents in their workplace
Outcomes	Primary health outcomes: Any non-reproductive adverse health effect (e.g., cancer, acute effects, immune system effects, liver and kidney toxicity, *etc.*) Secondary health outcomes: Genetic biomarkers of effect (specifically, chromosomal aberrations, micronuclei induction, and DNA damage measured by Comet Assay), and functional changes in immune system, liver, kidney or other organ systems
Study type	No restrictions

The overall objective and PECO statement were based on a series of problem formulation steps that included (1) input from an evaluation team with expertise in occupational health, toxicology, and systematic review, and information science; (2) deliberation with NTP staff and consultation with scientists at other Federal agencies; and (4) a public review of a concept document by the NTP Board of Scientific Counselors at the 16-18 April 2014 meeting (http://ntp.niehs.nih.gov/go/9741).

Key Questions and Contextual Topics

The overall objective of the evaluation can be phrased in terms of a specific research question "What is the hazard identification conclusion as to whether occupational exposure to cancer chemotherapy is associated with adverse health effects?" This research question serves as a focus of the evaluation to be answer by address the key question in **Table 2**. The evaluation also includes contextual topics, which provide background information to support the rationale or conduct of the systematic review but are not study questions addressed in the systematic review (USPSTF 2011). Sources of information for contextual questions include (1) targeted literature searches, (2) secondary reviews, (3) expert input, or (4) reports identified during the comprehensive literature screening for the key questions.

Table 2. Key Question and Contextual Topics	
Key Questions (KQ): Assessed by Systematic Review	
KQ1	What is the hazard identification category for an association between occupational exposure to chemotherapeutics and adverse non-reproductive health effects based on evidence from humans: 1) Known, (2) Presumed, (3) Suspected, (4) Not classifiable, or (5) Not identified to be a hazard to humans?
Contextual Topics (CT): Not Assessed by Systematic Review	
CT1	Summarize the known adverse health effects observed in human patients, including secondary cancers, as well as effects observed in experimental animal studies.
CT2	Summarize the range of blood or urine concentrations of parent drugs or metabolites associated with occupational exposure to cancer chemotherapy.
CT3	Summarize the level of environmental contamination (*e.g.*, surface, glove and air) reported in various work environments where cancer chemotherapy is used.

METHODS

Step 1. Problem Formulation

Problem Formulation Activities

OHAT initially considered conducting an evaluation on the evidence for an association between occupational exposure to cancer chemotherapy and adverse health effects after the association was identified as a research need during peer-review of the NTP Monograph on Developmental Effects and Pregnancy Outcomes Associated With Cancer Chemotherapy Use During Pregnancy (NTP 2013). The NTP Executive Committee[1] was informed about the potential evaluation, and solicited for input on agency interest/relevance and for names of agency technical staff that should be involved in the evaluation. An evaluation team was identified to include experts from the NTP, Food and Drug Administration (FDA), the National Cancer Institute (NCI), the National Institute of Occupational Safety and Health (NIOSH), the US Department of Labor's Office of Health and Safety (OSHA) and non-federal technical advisors (see "About this Protocol, Contributors"). The concept proposal for the evaluation was reviewed and approved by NTP's Board of Scientific Counselors in a public meeting on April 18, 2014 (http://ntp.niehs.nih.gov/go/9741). No public comments were received.

Refining the focus of the nomination

OHAT conducted an initial comprehensive inventory of the literature using a search strategy designed to identify all reported health outcomes associated with occupational exposure to cancer chemotherapy. Five databases were searched from the beginning of the database entries through June 3, 2013: PubMed, Embase, Scopus, Toxline, and Web of Science. After compiling the list of health outcomes from the initial search, OHAT was advised that NIOSH was in the process of completing a comprehensive review of reproductive health outcomes associated with occupational exposures to antineoplastic drugs

[1] The NTP Executive Committee provides programmatic and policy oversight to the NTP Director and meets once or twice a year in closed forum. Members of this committee include the heads (or their designees) from the following federal agencies: Consumer Product Safety Commission (CPSC), Department of Defense (DoD), Environmental Protection Agency (EPA), Food and Drug Administration (FDA), National Cancer Institute (NCI), National Center for Environmental Health/Agency for Toxic Substances and Disease Registry (NCEH/ATSDR), National Institute of Environmental Health Sciences (NIEHS), National Institute for Occupational Safety and Health (NIOSH), Occupational Safety and Health Administration (OSHA).

in health care settings. Thus, OHAT decided to focus the evaluation on the non-reproductive health outcomes (e.g., primary health outcomes: cancer, immune system effects, acute effects, kidney and liver toxicity, *etc.* and related secondary health outcomes (**Table 1**)) with occupational exposure to cancer chemotherapy agents.

Consideration of key scientific issues

Differences in the composition of the exposures

The composition of occupational exposure to cancer chemotherapy is complex. Regarding occupational exposure in medical settings, most cancer treatments include multiple cancer chemotherapy agents. Likewise, employees of drug manufacturers may prepare more than one type of cancer chemotherapy agent. The duration of exposure to antineoplastic agents may also differ due to number of years in a job with potential occupational exposure and the nature of the employment (e.g., oncology nurses administering cancer chemotherapy agents have a higher potential for exposure than housekeeping employees). For this evaluation, we will be evaluating exposure mainly by job description only. If possible, we may evaluate the adverse health outcomes and biomarkers of effects by broad grouping of employment (e.g., medical staff, veterinary staff, manufacturing staff, *etc.*). However, we anticipate that the majority of studies will be evaluating medical staff (e.g., oncology nurses and pharmacy technicians) and very few studies will be available on veterinarian or manufacturing employees. When available, we will comment on studies also reporting environmental contamination levels or biomonitoring levels of exposure of individual cancer chemotherapy drugs in association with health effects (see section below on *Internal dose determination*).

Internal dose determination

Currently, the most reliable way to determine the potential for occupational exposure to cancer chemotherapy agents is the detection of environmental contamination (e.g., surface, air, or glove contamination). However, it is difficult to estimate an internal dose from environmental contamination because exposure would vary dependent on route of exposure and the use of personal protective equipment. Biomonitoring studies of the more commonly used cytotoxic drugs can provide background information about the levels of the parent drug or metabolites in blood and urine of the occupationally exposed workers. However, the detection of parent or metabolites of these drugs is highly dependent on the timing of the blood or urine sample relative to the exposure. For this evaluation, OHAT will provide, as contextual information, a summary of the level of environmental contamination of parent drug reported in work areas where cancer chemotherapy is manufactured, prepared or administered. We will also provide, as additional contextual information, the range of blood or urine concentrations of parent drugs or metabolites reported in workers directly handling the drugs (e.g., manufacturing employees, pharmacists, nurses, or doctors) or caring for patients who were treated with the drugs (*e.g.*, hospital housekeeping, animal husbandry staff at a veterinarian clinic or research laboratory) (**Table 1**). There are many challenges in linking internal dose measurements to adverse health effects. For example, very few studies of health effects associated with occupational exposure to cancer chemotherapy agents also report internal dose determination. In addition, most internal dose measurements are likely measured at the same time that health effects are assessed; however, many health effects are the result of long term occupational exposure. Thus, for this evaluation of occupational exposure to cancer chemotherapy agents, we will be evaluating exposure mainly by job description and we will comment on the studies that also report environmental contamination levels of internal dose measurements in association with health effects.

Evidence of secondary malignancies in human cancer patients

The induction of secondary malignancies is well-recognized sequela in patients who have received antineoplastic agents for the treatment of cancer. Specifically, therapy-related acute myeloid leukemia (t-AML) accounts for up to 20% of all acute myeloid leukemia cases (Pedersen-Bjergaard *et al.* 2002). Treatment with antineoplastic cancer chemotherapy agents is also associated with the development of myelodysplastic syndrome, which is a closely related group of blood disorders that are often precursors of acute promyelogenous leukemia. t-AML and therapy related myelodysplastic syndrome have been observed following treatment with a variety of types of antineoplastic agents, including alkylating agents as well as non-alkylating agents, such as antimetabolites and topoisomerase II inhibitors (reviewed in McDiarmid *et al.* (2014)).

Evidence from experimental animal studies

Many cancer chemotherapy agents are cytotoxic drugs, and several of these agents are known developmental toxicants and/or mutagens in experiment animal studies (reviewed in (Shepard and Lemire 2004, NTP 2013). In the 12th Edition of the Report on Carcinogens, the National Toxicology Program identified 4 cancer chemotherapy agents as known to be a carcinogen and another 5 agents as reasonably anticipated to be a human carcinogen based on extensive review of human and animal literature (NTP 2014). Similarly, the International Agency for Research on Cancer (IARC) listed 6 cancer chemotherapy agents as Class 1 (Carcinogenic to humans), 7 agents as Class 2A (Probably carcinogenic to humans), and another 6 agents at Class 2B (Possibly carcinogenic to humans) based on extensive review of the human and animal data (IARC 1966, 1981, 1990, 2000). Due to the availability of thorough peer-reviewed assessments of the developmental toxicity and carcinogenicity of these agents in experimental animal studies, OHAT will not pursue a systematic review of the animal literature in the current evaluation. Instead, OHAT will consider the conclusions of the IARC and RoC for these agents based on the animal data and non-occupationally-exposed human studies (e.g., cancer patients administered cancer chemotherapy agents) as other data that may up or downgrade the hazard identification of the relevant human literature on adverse health outcomes associated with occupational exposure to cancer chemotherapy agents.

Step 2. Search and Select Studies for Inclusion

Literature Search Strategy

A literature search strategy was developed to identify all relevant published evidence on the health effects of occupational exposure to chemotherapeutics through (1) reviewing PubMed's Medical Subject Headings (MeSH) for relevant and appropriate terms, (2) extracting key terminology from relevant reviews and a set of previously identified primary data studies that are known to be relevant to the topic ("test set"), and (3) reviewing search strategies presented in other reviews. The search strategy was run and the results are assessed to ensure that 100% of the previously identified relevant primary studies were retrieved. Five databases were searched from the beginning of the database entries through October 23, 2014: PubMed, Embase, Scopus, Toxline, and Web of Science. The search strategy was customized for each database because of differences in syntax (see Appendix 1). No publication year limits will be imposed and the literature search will be updated for a final time in early July 2015. No language restrictions will be applied. We developed the literature search in collaboration with a librarian trained in systematic review methodology.

Databases Searched

- PubMed
- Embase
- Scopus
- Toxline
- Web of Science

Searching Other Resources

We will use the following methods to find additional studies that were not identified through the electronic searches. Studies will be evaluated using the same inclusion and exclusion criteria as used for screening records retrieved from the electronic search. Relevant studies identified through these steps will be marked as "provided from other sources" in the study selection flow diagram

- Hand searching the reference lists of relevant reviews, commentaries, or other non-research articles identified during the initial search. Commentaries or letters on specific studies are also reviewed to see if they contain content that should be noted during data extraction or risk of bias assessment of the original report.
- Hand searching the reference lists of all included studies after the full text review.
- Grey literature: To ensure retrieval of the relevant literature, OHAT may try to identify relevant grey literature, which refers to publications that are not commercially published or are not readily publicly available. For this report, we considered the NIOSH webpages detailing Occupational Exposure to Antineoplastic Agents and Other Hazardous Drugs (http://www.cdc.gov/niosh/topics/antineoplastic/default.html). These documents were used as background information and their reference lists were reviewed for relevant studies that may have been missed by the literature search.

Studies identified by the public when the initial list of included studies is posted on the OHAT website (anticipated for 60-90 days prior to peer review; studies identified within 30 days of posting will be considered for inclusion) or during the public comment period when the draft Monograph is released for public comment (anticipated for 45-60 days prior to peer review).

Unpublished data

NTP only includes publicly accessible and peer-reviewed information in its evaluations. If a study is identified which may be critical to the evaluation and is not peer reviewed, the NTP's practice is to obtain external peer review if the owners of the data are willing to have the study details and results made publicly accessible. The peer review would include an evaluation of the study similar to that for peer review of a journal publication. The NTP would identify and select 2-3 scientists knowledgeable in scientific disciplines relevant for the topic as potential peer reviewers. Persons invited to serve as peer reviewers would be screened for conflict of interest (COI) prior to confirming their service. In most instances, the peer review would be conducted by letter review. The study authors would be informed of the outcome of the peer review and given an opportunity to clarify issues or provide missing details. OHAT would consider the peer-review comments regarding the scientific and technical evaluation of the unpublished study in determining whether to include the study in its evaluation. The study and its related information, if used in the OHAT evaluation, would be publicly available. OHAT would

acknowledge via a note for the report that the document underwent external peer review managed by the NTP and the names of the peer reviewers would be identified. Unpublished data from personal author communication can supplement a peer-reviewed study, as long as the information is made publicly available.

Screening Process

DistillerSR®, a web-based, systematic review software program with structured forms and procedures will be used to screen articles for relevance and eligibility to ensure standardization of process[2]. Initially, results of the literature search are assembled in EndNote software and exact article duplicates removed prior to uploading the references into the systematic review software program.

Evidence Selection Criteria

In order to be eligible for inclusion, studies must comply with the criteria specified by the PECO statement (**Table 1**). Inclusion and exclusion criteria used to screen articles for relevance and eligibility at both the title-and-abstract and full-text screening stages are summarized in **Table 3**; these criteria are used to screen articles for relevance and eligibility at both the title-and-abstract and the full-text screening stages. In addition to criteria defining the relevant population, exposure, comparator, and outcomes, **Table 3** defines criteria for relevant publications types (e.g., the report must contain original data). Studies that do not meet these criteria will be excluded. Some articles may be categorized as possible supportive material if they appear inappropriate for inclusion, but appear to contain relevant background information. Those studies would not provide evidence of health effects, or lack of a health effect; however, the background information could provide context or other information (e.g., exposure or metabolism data) that would be useful when evaluating confidence in bodies of evidence and integration of the human data from the included studies with supporting evidence (e.g., IARC or RoC conclusions on the cancer inducing potential of specific cancer chemotherapy agents).

[2]DistillerSR® (http://systematic-review.net/) is a proprietary project management tool for tracking studies through the screening process and storing data extracted from these studies using user-customized forms.

Table 3. Inclusion and Exclusion Criteria to Determine Study Eligibility	
Inclusion Criteria	**Exclusion Criteria**
Participants/Population (human studies or experimental model systems)	
• Studies in adult humans (age ≥18 years old) utilizing a cohort, cross-sectional, case-control study design or case reports/series	• Non-human animals, including laboratory animal studies or pets • *In silico* studies or *in vitro* models utilizing organs, tissues, cell lines, or cellular components
Exposure	
• Occupational exposure to cancer chemotherapy agents (*e.g.*, workers in drug manufacturing;, preparation and administration of cancer chemotherapy; care of and housekeeping around patients receiving chemotherapy in the medical setting, veterinary practice, research laboratory, and home (*i.e.*, exposure occurring to medical personnel, family and friends caring for cancer patients at home) • Inhalation, dermal, or oral routes of exposure occurring via occupational exposure	• Non-cancer chemotherapy agents: Occupational exposure to other hazardous drugs or workplace exposures (*e.g.*, anesthetic gases, chlorine and other cleaning products, or viruses) • Transgenic attenuated viruses used as cancer chemotherapy agents • Non-occupational exposure to cancer chemotherapy agents (*i.e.*, cancer patient exposure to cancer chemotherapy)
Comparators	
• Humans exposed to lower levels (or no exposure/exposure below detection levels) of cancer chemotherapy agents in their workplace	• None
Outcomes	
• All non-reproductive health outcomes o Primary: cancer, immune system effects, acute effects, liver and kidney toxicity, *etc.* o Secondary: genetic biomarkers of effect (*i.e.*, chromosomal aberrations, micronuclei induction, or DNA damage measured by Comet Assay), immune function, kidney and liver markers, *etc.*	• Reproductive function (*e.g.*, fertility, effects on menstrual cycles) • Pregnancy outcomes and developmental effects (*e.g.*, fetal death, spontaneous preterm birth, teratogenicity)
Publications (e.g., language restrictions, use of conference abstracts)	
• Study must contain original data and must be peer-reviewed • English and non-English language studies	• Articles with no original data (*e.g.*, editorials, reviews[a]) • Non-peer reviewed articles (e.g., conference abstracts or other studies published in abstract form only, grant awards, and theses/dissertations) • Retracted articles
[a]Relevant reviews can be used as background and for reference scanning.	

Title/Abstract Review

Screeners will be trained using project-specific written instructions that reflect the criteria outlined in **Table 3** with an initial pilot phase undertaken to improve clarity of the inclusion and exclusion instructions and to improve accuracy and consistency among screeners. If changes to the inclusion criteria are made based on the pilot phase, they will be documented in a protocol amendment along with the date modifications were made and the logic for the changes. Trained screeners from the evaluation design team will then conduct a title and abstract screen of the search results to determine whether a reference meets the inclusion or exclusion criteria. All references will be independently screened by two screeners (one of which will be the project lead, who will screen all references). Studies

that are not excluded based on the title and abstract will be screened through a full-text review. In case of screening conflicts, screeners will independently review their screening results to confirm the inclusion/exclusion decision and, if needed, discuss discrepancies with the other screeners. If a true disagreement exists between screeners, the study passes to the full-text review.

Full-Text Review

After completion of the title/abstract screen, full-text articles are retrieved[3] for those studies that either clearly met the inclusion criteria or where eligibility to meet the inclusion criteria is unclear. Full-text review will be independently conducted by two screeners that participated in the title/abstract screening (again, one of which will be the project lead, who will screen all references). True disagreements will be resolved by discussion involving another member(s) of the team or, if necessary, through consultation with technical advisors.

Multiple publications of same data

Multiple publications with overlapping data for the same study (e.g., publications reporting subgroups, additional outcomes or exposures outside the scope of an evaluation, or longer follow-up) are identified by examining author affiliations, study designs, cohort name, enrollment criteria, and enrollment dates. If necessary, study authors will be contacted to clarify any uncertainty about the independence of two or more articles. OHAT will include all publications on the study, select one study to use as the primary, and consider the others as secondary publications with annotation as being related to the primary record during data extraction. The primary study will generally be the publication with the longest follow-up, or for studies with equivalent follow-up periods, the study with the largest number of cases or the most recent publication date. OHAT will include relevant data from all publications of the study, although if the same outcome is reported in more than one report, OHAT will include a single instance of the data (and avoid more than one, i.e. duplicate instances of the data.

Tracking study eligibility and reporting the flow of information

The reason for exclusion at the full-text-review stage will be annotated and reported in a study flow diagram in the final report**.** Commonly used categories for exclusion include: (1) is a review, commentary, or editorial with no original data; (2) lacks relevant exposure information; (3) lacks relevant health outcome information; and (4) is a conference abstract, thesis/dissertation.

Release of the list of included and excluded studies

The list of included and excluded studies will be posted on the OHAT website (http://ntp.niehs.nih.gov/go/evals) once screening has been completed and prior to completion of the draft OHAT monograph.

[3] OHAT will initially attempt to retrieve a full-text copy of the study using an automated program, such as QUOSA, when possible, and NIH library services (NIH subscriptions and interlibrary loans). For publications not available through NIH, OHAT will search the Internet and/or may attempt to contact the corresponding author. Studies not retrieved through these mechanisms are excluded and notated as “not available.”

Step 3. Data Extraction

Data Extraction Process and Data Warehousing

Data extraction will be managed with structured forms and stored in a database format using ICF International's proprietary Dose Response Analytical Generator and Organizational Network (DRAGON) software.[4] Data extraction elements for human studies are listed in Appendix 2. Study information collected during data extraction will be visualized and made publicly available in Excel format upon publication of the finalized report using Health Assessment Workspace Collaborative (HAWC), an open source, web-based interface.[5]

The extracted data will be used to summarize study designs and findings, facilitate assessment of risk of bias, and/or conduct statistical analyses during evidence synthesis in Step 7. The content of the data extraction may be revised following the identification of the studies included in the review. Data extraction will be performed by one member of the evaluation team and checked by a second member of the evaluation team for completeness and accuracy. Data extractors from the evaluation team will be trained using project-specific written instructions in an initial pilot phase using a subset of studies. Any discrepancies in data extraction will be resolved by discussion or consultation with a third member of the evaluation team. Information that is inferred, converted, or estimated during data extraction will be annotated and marked with brackets.

OHAT will attempt to contact authors of included studies to obtain missing data considered important for evaluating key study findings (e.g., level of data required to conduct a meta-analysis). The evaluation report will note that an attempt to contact study authors was unsuccessful if study researchers do not respond to an email or phone request within one month of the attempt to contact.

Step 4. Quality Assessment of Individual Studies

Internal validity or risk of bias will be assessed for individual studies using a tool developed by OHAT. Instructions for the risk-of-bias evaluation are provided in a guidance document tailored to the specific evidence stream and type of human study design in the detailed guide for using the tool (see "Risk-of-Bias Tool" at http://ntp.niehs.nih.gov/go/38673). The risk-of-bias tool is comprised of a common set of 11 questions that are answered based on the specific details of individual studies to develop risk-of-bias ratings (using the four response options in **Table 4** for each question. Study design determines the subset of questions that should be used to assess risk of bias for an individual study (**Table 5**). For example, the subset of risk-of-bias questions applicable to all of the experimental study designs (e.g., human case-control trials) includes a question on randomization of exposure that would not be applicable to human observational study designs.

Studies are independently assessed by two assessors who answer all applicable risk-of-bias questions with one of four options in **Table 4** based on answers from the CLARITY Research Group at McMaster University (CLARITY 2013) following pre-specified criteria detailed in Appendix 3. The criteria describe aspects of study design, conduct, and reporting required to reach risk-of-bias ratings for each question and specify factors that can distinguish among ratings (e.g., what separates "definitely low" from "probably low" risk of bias). The instructions and detailed criteria are tailored to the specific type of

[4] DRAGON (Dose Response Analytical Generator and Organizational Network) developed by ICF International (Fairfax, VA; http://www.icfi.com/insights/products-and-tools/dragon-online-tool-systematic-review).

[5] HAWC (Health Assessment Workspace Collaborative): A Modular Web-based Interface to Facilitate Development of Human Health Assessments of Chemicals (https://hawcproject.org/portal/).

human study designs. Risk of bias will be assessed at the outcome level because study design or method specifics may increase the risk of bias for some outcomes and not others within the same study.

Table 4: Answers to the Risk-of-Bias Questions Result in One of Four Risk-of-Bias Ratings	
++	***Definitely Low*** *risk of bias*: There is direct evidence of low risk of bias practices
+	***Probably Low*** *risk of bias*: There is indirect evidence of low risk of bias practices **OR** it is deemed that deviations from low risk of bias practices for these criteria during the study would not appreciably bias results, including consideration of direction and magnitude of bias
- NR	***Probably High*** *risk of bias*: There is indirect evidence of high risk of bias practices (indicated with "-") **OR** there is insufficient information provided about relevant risk of bias practices (indicated with "NR" for not reported). Both symbols indicate probably low risk of bias.
--	***Definitely High*** *risk of bias*: There is direct evidence of high risk of bias practices

Table 5: OHAT Risk-of-Bias Questions and Applicability by Study Design							
Risk-of-Bias Questions	**Experimental Animal***	***In Vitro* Exposure Studies**	**Human Controlled Trials****	**Cohort**	**Case-Control**	**Cross-Sectional*****	**Case Series**
1. Was administered dose or exposure level adequately randomized?	X	X	X				
2. Was allocation to study groups adequately concealed?	X	X	X				
3. Did selection of study participants result in the appropriate comparison groups?				X	X	X	
4. Did study design or analysis account for important confounding and modifying variables?				X	X	X	X
5. Were experimental conditions identical across study groups?	X	X					
6. Were research personnel blinded to the study group during the study?	X	X	X				
7. Were outcome data complete without attrition or exclusion from analysis?	X	X	X	X	X	X	
8. Can we be confident in the exposure characterization?	X	X	X	X	X	X	X
9. Can we be confident in the outcome assessment (including blinding of outcome assessors)?	X	X	X	X	X	X	X
10. Were all measured outcomes reported?	X	X	X	X	X	X	X
11. Were there no other potential threats to internal validity?	X	X	X	X	X	X	X

*Experimental animal studies are controlled exposure studies. Non-human animal observational studies can be evaluated using the design features of observational human studies such as cross-sectional study design. **Human Controlled Trials are studies in humans with controlled exposure (e.g., Randomized Controlled Trials, non-randomized experimental studies). ***Cross-sectional studies include population surveys with individual data (e.g., NHANES) and surveys with aggregate data (i.e., ecological studies).

Risk-of-Bias Assessment Process

Evaluation team members will be trained using project-specific instructions (Appendix 3) in an initial pilot-testing phase that is undertaken on a small subset of the included studies. All team members involved in the risk-of-bias assessment and asked to identify potential ambiguities in the criteria used to assign ratings for each question. Any ambiguities and rating conflicts will be discussed relative to opportunities to refine the criteria to more clearly distinguish between adjacent ratings. If major changes to the risk of bias criteria are made based on the pilot phase (i.e., those that would likely result in revision of response), they will be documented in a protocol amendment along with the date modifications were made and the logic for the changes. It is also expected that information about confounding, exposure characterization, outcome assessment, and other important issues may be identified during or after data extraction, which can lead to further refinement of the risk–of-bias criteria (Sterne *et al.* 2014).

Two members of the evaluation design team will independently make risk-of-bias determinations using structured forms for each study across all bias domains/question. Space is provided on the form for free-text response to justify each answer or provide context. Brief direct quotations from the text of the study should be used when appropriate. After completing their risk-of-bias assessment for a study, the two members of the evaluation team assessors will compare their results to identify discrepancies and attempt to resolve them. Any remaining discrepancies will be assessed by the project lead and, if needed, other members of the evaluation design team and/or technical advisors. If the evaluation team cannot reach agreement on a risk-of-bias determination for a particular domain, the more conservative judgment will be selected (e.g. if one reviewer makes a judgment of “Definitely low bias” and the other reviewer makes a judgment of “probably low risk of bias”, the judgement of ”probably low” will be used). The final risk-of-bias rating for each question will be recorded along with a statement of the basis for that rating. The risk-of-bias assessment of included studies will be part of the study summaries released in materials for the draft OHAT monograph that will be posted for public comment prior to peer review. Peer review will provide an opportunity for investigators and the public to comment on risk-of-bias assessment.

Missing Information for Risk-of-Bias Assessment

OHAT will attempt to contact authors of included studies by email to obtain missing information considered critical for evaluating risk of bias that cannot be inferred from the study. If additional data or information is received from study authors, risk-of-bias judgments will be modified to reflect the updated study information. If OHAT does not receive a response from the authors by one month of the contact attempt, a risk-of-bias response of “not reported; probably high risk of bias” will be used and a note made in the data extraction files that an attempt to contact the authors was unsuccessful.

Step 5. Organizing and Rating Confidence in the Bodies of Evidence

OHAT will consider the collection of studies on the same or closely related adverse health outcomes (e.g., cancer, immune effects, acute effects, biomarkers of effect) as bodies of evidence and develop overall confidence ratings in these bodies of evidence using a modification of the GRADE framework. Procedures for grouping the adverse health outcomes, considering quantitative or narrative synthesis, and developing confidence ratings for this evaluation are described below.

Health Outcome and Endpoint Grouping

Health outcomes will be grouped into 4 primary outcome categories that have greater predictability for adverse health outcomes often assessed in occupational exposure to cancer chemotherapy agents: cancer, immune-related diseases and measures of immune function, liver and renal toxicity and acute effects. Other health outcomes may be considered as primary outcomes depending on the availability of data reported in the literature (e.g., cardiotoxicity). Secondary outcomes are considered those outcomes with less predictive value for adverse health outcomes and will be considered with the corresponding primary outcomes. Secondary outcomes in this evaluation include: genetic toxicity as a biomarker of effect for cancer (e.g., chromosomal aberrations, micronuclei induction, DNA damage by Comet assay); observational immune endpoints and immunosuppression (e.g., decreased levels of lymphocytes); and biomarkers of liver and kidney function (e.g., elevated liver enzymes). For further explanation of primary and secondary outcomes, see **Table 6**.

Considerations for Pursuing a Narrative or Quantitative Evidence Synthesis

Heterogeneity within the available evidence will determine the type of evidence integration that is appropriate: either a quantitative synthesis (meta-analysis) or narrative approach for evidence integration. When appropriate, we will perform a meta-analysis. Summaries of main characteristics for each included study will be compiled and reviewed by two evaluation team members to determine comparability between studies, identify data transformations necessary to ensure comparability, and determine whether heterogeneity is a concern. The main characteristics evaluated across all eligible studies include the following:

Human Studies

- Study design (e.g., cross-sectional, cohort)
- Details on how participants were classified into exposure groups, if any (e.g., quartiles of exposure concentration)
- Details on source of exposure data (e.g., questionnaire, area monitoring, biomonitoring)
- Concentrations of the chemical(s) for each exposure group
- Health outcome(s) reported
- Conditioning variables in the analysis (e.g., variables considered confounders)
- Type of data (e.g., continuous or dichotomous), statistics presented in paper, ability to access raw data
- Variation in degree of risk of bias at individual study level

Table 6. Identification of Primary and Secondary Non-Reproductive Health Outcomes				
Health outcome	**Cancer**	**Immune**	**Liver and kidney toxicity**	**Other health effects**
Primary	***Incidence of cancer*** (e.g., any type of cancer combined, or incidence of specific cancer types)	***Immune-related diseases and measures of immune function*** (e.g., infection, atopic dermatitis, asthma)	***Liver and kidney disease***	**Acute non-specific effects**
Secondary	***Genetic toxicity*** (e.g., structural chromosomal aberrations, micronuclei induction, DNA damage as measured by Comet assay)	***Observational immune endpoints and immunosuppression*** (e.g., lymphocyte counts, cytokine levels)	***Liver and kidney function*** (e.g., elevated liver enzymes, elevated urine creatinine and blood urea nitrogen (BUN) levels)	***Acute health effects*** (e.g., alopecia, eye-watering, nausea, skin irritation)

More detailed guidance on evaluating heterogeneity, transforming or normalizing data to ensure comparability, and the process for determining whether a meta-analysis will be pursued is provided in the OHAT Handbook for Conducting a Literature-Based Health Assessment (http://ntp.niehs.nih.gov/go/38673, see STEP 5). We expect to require input from topic-specific experts to help assess whether studies are too heterogeneous for meta-analysis to be appropriate. Situations where it may not be appropriate to include a study are (1) data on exposure or outcome are too different to be combined, (2) there are concerns about high risk of bias, or (3) other circumstances may indicate that averaging study results would not produce meaningful results. When it is inappropriate or not feasible to quantitatively combine results, OHAT will narratively describe or visually present findings.

Stratified Analyses, Meta-Regression, and Publication Bias

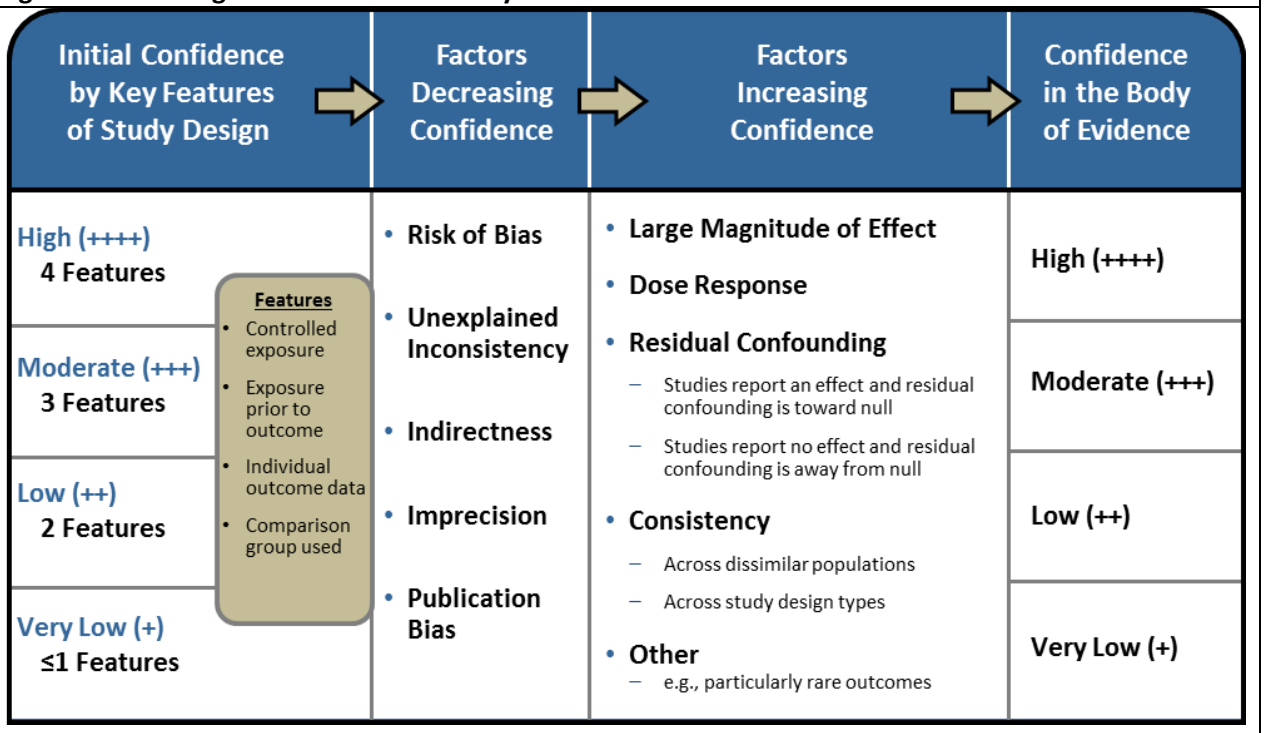
If there is significant study-level heterogeneity, then OHAT may conduct stratified analyses or multivariate meta-regression in an attempt to determine how much heterogeneity can be explained by taking into account both within- and between-study variance (Vesterinen *et al.* 2014). Multivariate meta-regression approaches are especially useful for assessing the significance of associations between study design characteristics. These approaches are considered most suitable if there are at least six to ten studies for a continuous variable and at least four studies for a categorical variable (Fu *et al.* 2011).
If possible, i.e., if there are enough studies; we will assess potential publication bias by developing funnels and performing Egger regression on the estimates of effect size. In addition, if these methods suggest that publication bias is present, we will use trim and fill methods to predict the impact of the hypothetical "missing" studies (Vesterinen *et al.* 2014).

Confidence Rating: Assessment of Body of Evidence

The quality of evidence for each outcome will be graded using the GRADE system for rating the confidence in the body of evidence (Guyatt et al 2011) as adapted by OHAT for human observational studies (NTP 2014, Rooney *et al.* 2014) (**Figure 1**). More detailed guidance on reaching confidence ratings in the body of evidence as "high", "moderate", "low" or "very low" is provided in the OHAT Handbook for Conducting a Literature-Based Health Assessment (http://ntp.niehs.nih.gov/go/38673, see Step 5). In brief, available studies on a particular outcome are initially grouped by key study design features, and each grouping of studies is given an initial confidence rating by those features. This initial

rating (column 1 of Figure 1) is downgraded for factors that decrease confidence in the results (column 2 of Figure 1 [risk of bias, unexplained inconsistency, indirectness or lack of applicability, imprecision, and publication bias]) and upgraded for factors that increase confidence in the results (column 3 of Figure 1 [large magnitude of effect, dose response, consistency across study designs/populations/animal models or species, consideration of residual confounding, and other factors that increase our confidence in the association or effect]).

Figure 1. Assessing Confidence in the Body of Evidence



The reasons for downgrading (or upgrading) confidence may not be due to a single domain of the body of evidence. If a decision to downgrade is on the borderline for two domains, the body of evidence is downgraded once in a single domain to account for both partial concerns based on considering the key drivers of the strengths or weaknesses. Similarly, the body of evidence is not downgraded twice for what is essentially the same limitation (or upgraded twice for the same asset) that could be considered applicable to more than one domain of the body of evidence. Confidence ratings are independently assessed by members of the evaluation team, and discrepancies are resolved by consensus and consultation with technical advisors as needed. Confidence ratings are summarized in evidence profile tables (see Table 7 for general format).

Table 7. Evidence Profile Table Format

Example of the type of information that will be in an evidence profile for immune health outcomes

Body of Evidence	Risk of Bias	Unexplained Inconsistency	Indirectness	Imprecision	Publication Bias	Magnitude	Dose Response	Residual Confounding	Consistency Across Study Designs	FINAL RATING
Evidence stream (human or animal)	Serious or not serious	Serious or not serious	Serious or not serious	Serious or not serious	Detected or undetected	Large or not large	Yes or no	Yes or no	Yes or no	**Final Rating**
(# Studies) Initial Rating	• Describe trend • Describe key questions • Describe issues	• Describe results in terms of consistency • Explain apparent inconsistency (if it can be explained)	• Discuss use of upstream indicators or populations with less relevance	• Discuss ability to distinguish treatment from control • Describe confidence intervals	• Discuss factors that might indicate publication bias (e.g., funding, lag)	• Describe magnitude of response	• Outline evidence for or against dose response	• Address whether there is evidence that confounding would bias toward null	• Describe study design consistency	**High, Moderate, or Low**

Adverse Health Outcomes

For the evaluation of adverse health effects associated with cancer chemotherapy, primary outcomes are considered to be the most direct, or applicable, to the evaluation (**Table 6**). Secondary outcomes are relevant, but less direct and can include upstream indicators or intermediate outcomes.

- Primary health outcomes: The primary outcomes are most predictive of adverse health outcomes and therefore there will be no downgrades for indirectness for these outcomes.
- Secondary health outcomes: The secondary outcomes are considered less predictive of adverse health effects and therefore will be downgraded one level for indirectness.

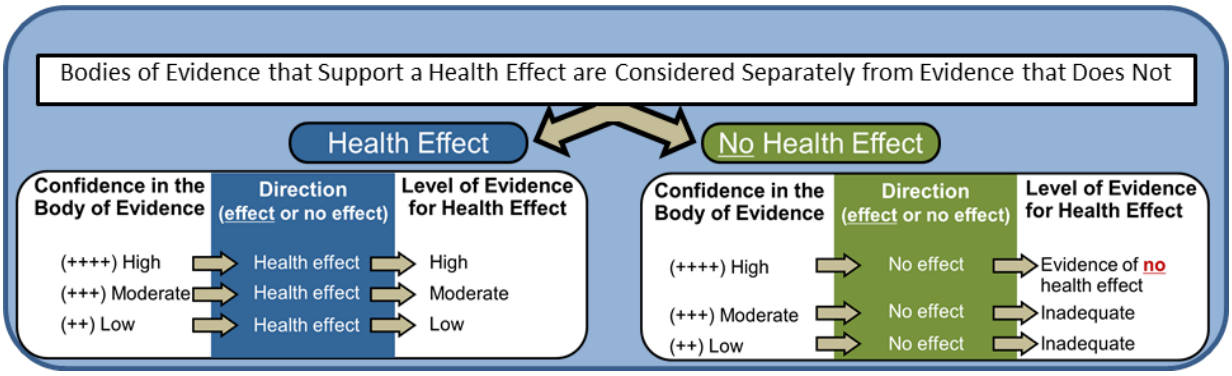
Exposure to Cancer Chemotherapy in the Workplace

All exposure levels and scenarios encountered in the human studies (e.g., general population, occupational settings, etc.) will be considered direct and not downgraded.

Step 6: Preparation of Draft Level of Evidence Statement

The confidence ratings will be translated into draft level of evidence of health effects for each type of health outcome, separately, according to one of four statements: (1) High, (2) Moderate, (3) Low, or (4) Inadequate (**Figure 2**). The descriptor “evidence of no health effect” is used to indicate confidence that the substance is not associated with a health effect. Because of the inherent difficulty in proving a negative, the conclusion “evidence of no health effect " is only reached when there is high confidence in the body of evidence.

Figure 2. Translation of Confidence Ratings into Evidence of Health Effect Conclusions

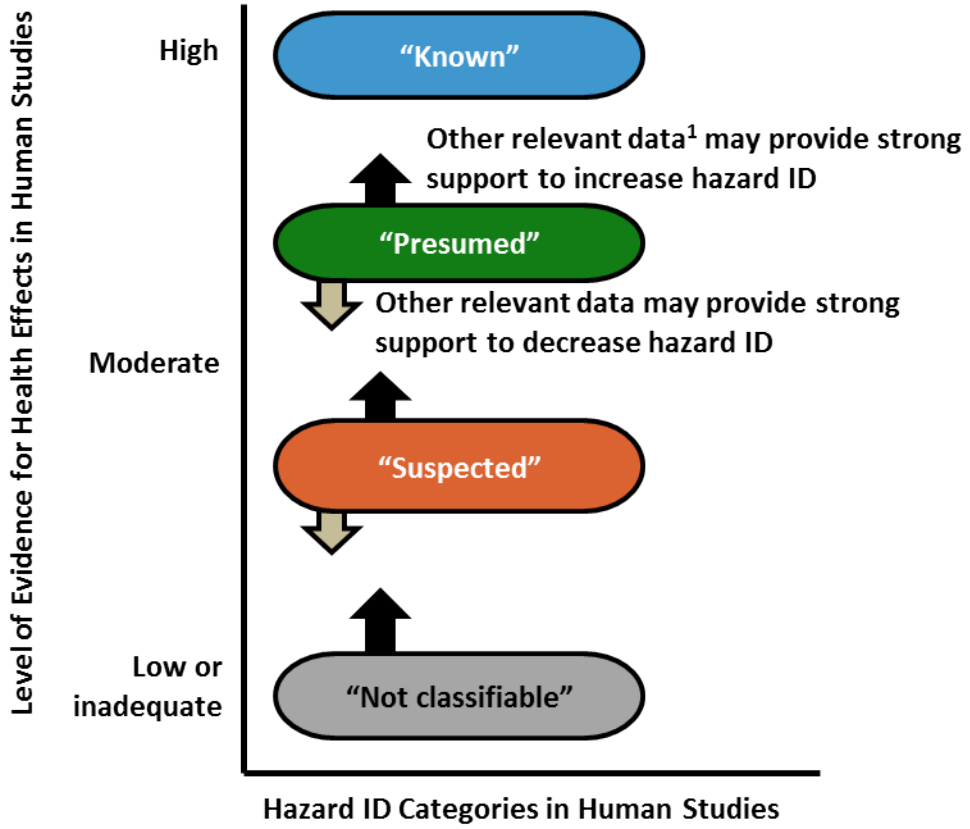


Evidence Descriptors	Definition
High Level of Evidence	There is high confidence in the body of evidence for an association between occupational exposure to cancer chemotherapy agents and the health outcome(s).
Moderate Level of Evidence	There is moderate confidence in the body of evidence for an association between occupational exposure to cancer chemotherapy agents and the health outcome(s).
Low Level of Evidence	There is low confidence in the body of evidence for an association between occupational exposure to cancer chemotherapy agents and the health outcome(s), or no data are available.
Inadequate Evidence	There is insufficient evidence available to assess if occupational exposure to cancer chemotherapy agents is associated with the health outcome(s).
Evidence of No Health Effect	There is high confidence in the body of evidence that occupational exposure to cancer chemotherapy agents is not associated with the health outcome(s).

Step 7: Integrate Evidence to Develop Hazard Identification Conclusions

Finally, the levels of evidence ratings from human studies for each type of health outcome will be used to reach one of five possible hazard identification categories: (1) Known, (2) Presumed, (3) Suspected, (4) Not classifiable, or (5) Not identified to be a hazard to humans (**Figure 3**). OHAT will consider the conclusions of the IARC and RoC for the agents based on the animal data and non-occupationally-exposed human studies (e.g., cancer patients administered cancer chemotherapy agents) as other relevant data that may up or downgrade the hazard identification of the relevant human literature on adverse health outcomes associated with occupational exposure to cancer chemotherapy agents.

Figure 3. Hazard Identification Scheme Based on Confidence in the Body of Evidence for Health Effects



[1]Relevant data for the hazard identification of cancer associated with occupational exposure to cancer chemotherapy include the IARC and NTP Report on Carcinogens evaluations of cancer chemotherapy agents.

NTP MONOGRAPH FORMAT

The NTP Monograph will include the following information:

Introduction

This section will provide a brief background on the topic.

Methodology

This section will provide a brief overview of the methodologies used in the review process, including:

- the research question;
- the search strategy used to identify and retrieve studies;
- the process for selecting the included studies;
- the methods of data extraction;
- the methods of quality assessment of included studies;
- the methods used to synthesize the data of included studies;
- the methods used to evaluate confidence in the bodies of evidence;
- the methods used to reach hazard identification conclusions

Results

This section will include the results from the systematic review. Results will be presented in tables or figures as appropriate using HAWC. The results from the included studies will be discussed by outcome. This will include a description of:

- the number of studies identified that reported the outcome;
- a full list of excluded studies, with the reasons for exclusion;
- a summary of the results and quality assessment for each individual included study (including files in downloadable format);
- a description of results across studies and analysis of confidence in the body of evidence using GRADE
- a GRADE evidence profile for each health outcome; and
- the level of evidence and draft hazard identification conclusions.

Discussion

The discussion will provide a summary of the review findings, including a discussion of any gaps identified in the evidence and any suggestions of areas for further research. Any important limitations of the review will be described and their impact on the available evidence will be discussed.

Conclusion

This will present the conclusion of the review.

REFERENCES

AHRQ (Agency for Healthcare Research and Quality). 2014. *Agency for Healthcare Research and Quality (AHRQ) Training Modules for the Systematic Review Methods Guide*. Rockville, MD: Service UDoHaH. Agency for Healthcare Research and Quality. Available: http://www.effectivehealthcare.ahrq.gov/index.cfm/tools-and-resources/slide-library/ [accessed 15 October 2015].

Bonassi S, Znaor A, Norppa H, Hagmar L. 2004. Chromosomal aberrations and risk of cancer in humans: an epidemiologic perspective. *Cytogenet Genome Res* 104(1-4): 376-382.

Bonassi S, Znaor A, Ceppi M, Lando C, Chang WP, Holland N, Kirsch-Volders M, Zeiger E, Ban S, Barale R, Bigatti MP, Bolognesi C, Cebulska-Wasilewska A, Fabianova E, Fucic A, Hagmar L, Joksic G, Martelli A, Migliore L, Mirkova E, Scarfi MR, Zijno A, Norppa H, Fenech M. 2007. An increased micronucleus frequency in peripheral blood lymphocytes predicts the risk of cancer in humans. *Carcinogenesis* 28(3): 625-631.

CLARITY (Clinical Advances through Research and Information Translation Research Group). 2013. *Tools to assess risk of bias in cohort studies, case control studies, randomized controlled trials, and longitudinal symptom research studies aimed at the general population*. Hamilton, Ontario, Canada: University M. Available: http://www.evidencepartners.com/resources/ [accessed 15 January 2013].

Collins A, Koppen G, Valdiglesias V, Dusinska M, Kruszewski M, Moller P, Rojas E, Dhawan A, Benzie I, Coskun E, Moretti M, Speit G, Bonassi S, ComNet p. 2014. The comet assay as a tool for human biomonitoring studies: the ComNet project. *Mutat Res Rev Mutat Res* 759: 27-39.

Connor TH, McDiarmid MA. 2006. Preventing occupational exposures to antineoplastic drugs in health care settings. *CA Cancer J Clin* 56(6): 354-365.

Connor TH, MacKenzie BA, Debord DG, Trout DB. 2012. NIOSH List of Antioneoplastic and Other Hazardous Drugs in Healthcare Settings 2012. *Govt Reports Announcements & Index*(20): 20.

Connor TH, Lawson CC, Polovich M, McDiarmid MA. 2014. Reproductive health risks associated with occupational exposures to antineoplastic drugs in health care settings: a review of the evidence. *J Occup Environ Med* 56(9): 901-910.

Couch J, West C. 2012. Health Hazard Evaluation Report: HETA-2009-0148-3158, June 2012. Chemotherapy Drug Exposure at an Oncology Clinic - Florida. *Govt Reports Announcements & Index*(20): 36.

Couch J, Gibbins J, Connor TH. 2013. Evaluation of chemotherapy drug exposure at a veterinary teaching hospital in Michigan. *J Occup Environ Hyg* 10(4): D45-51.

Dranitsaris G, Johnston M, Poirier S, Schueller T, Milliken D, Green E, Zanke B. 2005. Are health care providers who work with cancer drugs at an increased risk for toxic events? A systematic review and meta-analysis of the literature. *J Oncol Pharm Pract* 11(2): 69-78.

Fu R, Gartlehner G, Grant M, Shamliyan T, Sedrakyan A, Wilt TJ, Griffith L, Oremus M, Raina P, Ismaila A, Santaguida P, Lau J, Trikalinos TA. 2011. Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program. *J Clin Epidemiol* 64(11): 1187-1197.

Higgins J, Green S. 2011. *Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 (updated March 2011)*. Collaboration TC. Available: http://handbook.cochrane.org/ [accessed 3 February 2013].

Hon CY, Teschke K, Demers PA, Venners S. 2014. Antineoplastic drug contamination on the hands of employees working troughout the hospital medication system. *Ann Occup Hyg* 58(6): 761-770.

IARC (International Agency for Research on Cancer). 1966. *IARC Monograph on the Evaluation of the Carcinogenic Risk of Chemicals to Humans: Some Pharmaceutical Drugs*. Lyon, France. World Health Organization.

IARC (International Agency for Research on Cancer). 1981. *IARC Monograph on the Evaluation of the Carcinogenic Risk of Chemicals to Humans: Some Antineoplastics & Immunosuppressive Agents*. Lyon, France. World Health Organization.

IARC (International Agency for Research on Cancer). 1990. *IARC Monograph on the Evaluation of the Carcinogenic Risk of Chemicals to Humans: Pharmaceutical Drugs*. Lyon, France. World Health Organization.

IARC (International Agency for Research on Cancer). 2000. *IARC Monograph on the Evaluation of the Carcinogenic Risk of Chemicals to Humans: Some Antineoplastics, Antivirals, and Other Pharmaceutical Drugs*. Lyon, France. World Health Organization.

Kiffmeyer TK, Tuerk J, Hahn M, Stuetzer H, Hadtstein C, Heinemann A, Eickmann U. 2013. Application and assessment of a regular environmental monitoring of the antineoplastic drug contamination level in pharmacies - the MEWIP project. *Ann Occup Hyg* 57(4): 444-455.

Kopp B, Schierl R, Nowak D. 2013. Evaluation of working practices and surface contamination with antineoplastic drugs in outpatient oncology health care settings. *Int Arch Occup Environ Health* 86(1): 47-55.

McDiarmid MA, Rogers B, Oliver MS. 2014. Chromosomal effects of non-alkylating drug exposure in oncology personnel. *Environ Mol Mutagen* 55(4): 369-374.

Norppa H, Bonassi S, Hansteen IL, Hagmar L, Stromberg U, Rossner P, Boffetta P, Lindholm C, Gundy S, Lazutka J, Cebulska-Wasilewska A, Fabianova E, Sram RJ, Knudsen LE, Barale R, Fucic A. 2006. Chromosomal aberrations and SCEs as biomarkers of cancer risk. *Mutat Res* 600(1-2): 37-45.

NTP (National Toxicology Program). 2013. *NTP Monograph: Developmental Effects and Pregnancy Outcomes Associated With Cancer Chemotherapy Use During Pregnancy*. NIH 13-5956. Research Triangle Park, NC. National Institute of Environmental Health Sciences. Available: http://www.ncbi.nlm.nih.gov/pubmed/24736875 [accessed 28 May 2015].

NTP (National Toxicology Program). 2014. *Report on Carcinogens, 13th Edition*. Research Triangle Park, NC. US Department of Health and Human Services, Public Health Service. Available: http://ntp.niehs.nih.gov/pubhealth/roc/roc13/ [accessed 28 May 2015].

OSHA. 1999. Section VI: Chapter 2 Controlling Occupational Exposure to Hazardous Drugs. Directive number: TED 01-00-015 [TED 1-0.15A]. In: *Occupational Safety and Health Administration (OSHA) Technical Manual (OTM)*. Washington, DC: United States Department of Labor. Available: http://www.osha.gov/dts/osta/otm/otm_vi/otm_vi_2.html#2 [accessed 28 May 2015].

Pedersen-Bjergaard J, Andersen MK, Christiansen DH, Nerlov C. 2002. Genetic pathways in therapy-related myelodysplasia and acute myeloid leukemia. *Blood* 99(6): 1909-1912.

Quansah R, Jaakkola JJ. 2010. Occupational exposures and adverse pregnancy outcomes among nurses: a systematic review and meta-analysis. *J Womens Health (Larchmt)* 19(10): 1851-1862.

Rooney AA, Boyles AL, Wolfe MS, Bucher JR, Thayer KA. 2014. Systematic review and evidence integration for literature-based environmental health science assessments. *Environ Health Perspect* 122(7): 711-718.

Shamliyan T, Kane RL, Dickinson S. 2010. A systematic review of tools used to assess the quality of observational studies that examine incidence or prevalence and risk factors for diseases. *J Clin Epidemiol* 63(10): 1061-1070.

Shepard TH, Lemire RJ. 2004. *Catalog of Teratogenic Agents*. Eleventh ed., Baltimore, MD: The Johns Hopkins University Press.

Sterne J, Higgins J, Reeves B, on behalf of the development group for ACROBAT-NRSI. 2014. *A Cochrane Risk of Bias Assessment Tool: for Non-Randomized Studies of Interventions (ACROBAT-NRSI). Version 1.0.0*. Available: https://sites.google.com/site/riskofbiastool/ [accessed 28 September 2014].

USPSTF (US Preventative Task Force). 2011. *Procedure Manual*. AHRQ Publication No. 08-05118-EF. Washington, DC: USPTF. Available: http://www.uspreventiveservicestaskforce.org/Page/Name/procedure-manual [accessed 16 October 2015].

Vesterinen HM, Sena ES, Egan KJ, Hirst TC, Churolov L, Currie GL, Antonic A, Howells DW, Macleod MR. 2014. Meta-analysis of data from animal studies: a practical guide. *J Neurosci Methods* 221: 92-102.

Villa AF, El Balkhi S, Aboura R, Sageot H, Hasni-Pichard H, Pocard M, Elias D, Joly N, Payen D, Blot F, Poupon J, Garnier R. 2015. Evaluation of oxaliplatin exposure of healthcare workers during heated intraperitoneal perioperative chemotherapy (HIPEC). *Ind Health* 53(1): 28-37.

ABOUT THE PROTOCOL

Contributors

Evaluation Team

Evaluation teams are composed of federal staff and contractor staff. Contractor staff members are screened for potential conflicts of interest. Federal staff members should do a self-evaluation. Epidemiologists and toxicologists on OHAT evaluation teams should have at least three years' experience and/or training in reviewing studies, including summarizing studies and critical review (e.g., assessing study quality and interpreting findings). Experience in evaluating occupational or environmental studies is preferred. Team members should have at least a master's degree or equivalent experience in epidemiology, toxicology, environmental health sciences, or a related field.

Name	Affiliation
Kembra Howdeshell, PhD	NIEHS/NTP, Project Lead
Tom Connor, PhD	NIOSH
Grace Kissling, PhD	NIEHS/NTP
Margy Lambert, PhD	OSHA
Tina Lawson, PhD	NIOSH
John Leighton, PhD	FDA
Barry McIntyre, PhD	NIEHS/NTP
Stephanie Holmgren, MSLS, MBA	NIEHS/OSIM
Harold Seifried, PhD	NCI
Nonnekaye Shelburne, MS, CRNP	NCI
Kristina Thayer, PhD	NIEHS/NTP
Kristine Witt, MS	NIEHS/NTP

Contract support: Will assist in data extraction and risk of bias assessment

Name	Affiliation
Robyn Blain, MS, PhD	ICF Incorporated
Pam Ross, MS	ICF Incorporated

Technical Advisors

Technical advisors are outside experts retained on an as-needed basis to provide individual advice to the NTP for a specific topic. The technical advisors were selected for their experience with occupational exposure to cancer chemotherapy agents. Technical advisors were screened for conflict of interest prior to their service and did not report any conflicts of interest. Service as a technical advisor does not necessarily indicate that an advisor has read the entire protocol or endorses the final document.

Name	Affiliation
Melissa McDiarmid, MD, MPH	University of Maryland School of Medicine

Sources of Support

National Institute of Environmental Health Sciences/Division of the National Toxicology Program

Protocol History and Revisions

Date	Activity or revision
May 28, 2015	Draft evaluation protocol reviewed: sent to experts for comment/review
October 9, 2015	Evaluation protocol posted on OHAT website

APPENDICES

Appendix 1. Literature Search Strategy

The strategy for this search is broad for the consideration of adverse health outcomes, including biomarkers of effect, and comprehensive for occupational exposure to cancer chemotherapy agents as an exposure or treatment in order to ensure inclusion of relevant papers.

Database	Search Terms
Embase	('Antineoplastic agent'/exp OR 'antineoplastic agent') AND ('health care personnel'/exp OR 'health care personnel') AND ('occupational exposure'/de OR 'occupational exposure' OR 'occupational hazard'/de OR 'occupational hazard' OR 'occupational safety'/de OR 'occupational safety') AND ([embase]/lim OR [embase classic]/lim)
PubMed	("Neoplasms/drug therapy"[mh] OR "antineoplastic protocols"[mh] OR "chemotherapy, adjuvant"[mh] OR "antineoplastic agents"[mh] OR "antineoplastic agents"[pharmacological action] OR chemotherapy[tiab] OR chemotherapies[tiab] OR chemotherapeutic*[tiab] OR anticancer[tiab] OR antineoplastic*[tiab] OR anti- tumor*[tiab] OR anti-tumour*[tiab] OR cytostatic[tiab] OR hazardous-drug*[tiab] OR drug*-hazardous[tiab] OR category-D[tiab] OR category-x[tiab])) OR ("4 aminofolic acid"[tiab] OR "4 epidoxorubicin"[tiab] OR "5 fluorouracil"[tiab] OR "6 mercaptopurine"[tiab] OR "6 thioguanine"[tiab] OR Abraxane[tiab] OR adrucil[tiab] OR "all-trans retinoic acid"[tiab] OR ATRA[tiab] OR altretamine[tiab] OR adriamycin[tiab] OR "actinomycin D"[tiab] OR aminopterin[tiab] OR Anastrozole[tiab] OR "ARA-C"[tiab] OR arimidex[tiab] OR aromasin[tiab] OR "behenoyl cytosine arabinoside"[tiab] OR bevacizumab[tiab] OR BHAC[tiab] OR bleomycin[tiab] OR bortezomib[tiab] OR busulfan[tiab] OR busulfex[tiab] OR carboplatin[tiab] OR capecitabine[tiab] OR carmustine[tiab] OR Cerubidine[tiab] OR chlorambucil[tiab] OR cisplatin[tiab] OR cis-platinum[tiab] OR cyclophosphamide[tiab] OR cytarabine[tiab] OR cytosar[tiab] OR "cytosine arabinoside"[tiab] OR Cytoxan[tiab] OR dacarbazine[tiab] OR dasatinib[tiab] OR daunorubicin[tiab] OR daunoxome[tiab] OR deltasone[tiab] OR docetaxel[tiab] OR doxorubicin[tiab] OR efudex[tiab] OR eldisine[tiab] OR Ellence[tiab] OR Eloxatin[tiab] OR emcyt[tiab] OR enocitabine[tiab] OR epirubicin[tiab] OR erlotinib[tiab] OR etopophos[tiab] OR etoposide[tiab] OR estramustine[tiab] OR exemestane[tiab] OR fareston[tiab] OR femara[tiab] OR fludara[tiab] OR fludarabine[tiab] OR folex[tiab] OR fulvestrant[tiab] OR Faslodex[tiab] OR gefitinib[tiab] OR gemcitabine[tiab] OR gemtuzumab[tiab] OR gemzar[tiab] OR gleevec[tiab] OR glivec[tiab] OR herceptin[tiab] OR hexamethylmelamine[tiab] OR hydroxycarbamide[tiab] OR hydroxyurea[tiab] OR idarubicin[tiab] OR IFEX[tiab] OR ifosfamide[tiab] OR imatinib[tiab] OR "interferon alpha"[tiab] OR iressa[tiab] OR irinotecan[tiab] OR ixabepilone[tiab] OR ixempra[tiab] OR lapatinib[tiab] OR letrozole[tiab] OR lomustine[tiab] OR matulane[tiab] OR mechlorethamine[tiab] OR melphalan[tiab] OR methotrexate[tiab] OR "mitomycin c"[tiab] OR mitoxantrone[tiab] OR mustargen[tiab] OR "mustine Hcl"[tiab] OR mutamycin[tiab] OR myleran[tiab] OR mylotarg[tiab] OR navelbine[tiab] OR nilotinib[tiab] OR "nitrogen mustard HCl"[tiab] OR nolvadex[tiab] OR novantrone[tiab] OR oncovin[tiab] OR oxaliplatin[tiab] OR ozogamicin[tiab] OR paclitaxel[tiab] OR paraplatin[tiab] OR pemetrexed[tiab] OR pentostatin[tiab] OR platinol[tiab] OR prednisone[tiab] OR procarbazine[tiab] OR rituxan[tiab] OR rituximab[tiab] OR sorafenib[tiab] OR sprycel[tiab] OR streptozocin[tiab] OR sunitinib[tiab] OR sunrabin[tiab] OR sutent[tiab] OR tamoxifen[tiab] OR tarceva[tiab] OR tasigna[tiab] OR taxol[tiab] OR taxotere[tiab] OR temodar[tiab] OR temozolomide[tiab] OR teniposide[tiab] OR thioplex[tiab] OR thiotepa[tiab] OR toposar[tiab] OR topotecan[tiab] OR toremifene[tiab] OR trastuzumab[tiab] OR tretinoin[tiab] OR tykerb[tiab] OR velban[tiab] OR velcade[tiab] OR vepesid[tiab] OR vesanoid[tiab] OR vinblastine[tiab] OR vincasar[tiab] OR vincrex[tiab] OR vincristine[tiab] OR vindesine[tiab] OR vinorelbine[tiab] OR VM26[tiab] OR VP16[tiab] OR Vumon[tiab] OR Xeloda[tiab] OR zanosar[tiab]) AND (Environmental exposure[mh] OR Occupational exposure[mh] OR occupational diseases[mh] OR occupation*[tiab] OR workplace[tiab] OR work-related[tiab] OR

Database	Search Terms
	exposure*[tiab] OR exposed[tiab] OR contaminat*[tiab] OR handl*[tiab]) AND Health services[mh] OR health occupations[mh] OR health personnel[mh] OR pharmacist*[tiab] OR technician*[tiab] OR nurse*[tiab] OR nursing[tiab] OR physician*[tiab] OR clinician*[tiab] OR doctor*[tiab] OR veterinarian*[tiab])) OR ((hospital[tiab] OR clinic[tiab] OR medical[tiab] OR health[tiab] OR healthcare[tiab] OR health-care[tiab] OR pharmacy[tiab] OR pharmaceutical*[tiab] OR cancer[tiab] OR oncolog*[tiab] OR veterinary[tiab]) AND (staff[tiab] OR personnel[tiab] OR worker*[tiab] OR employee*[tiab] OR technician*[tiab] OR aide*[tiab] OR assistant*[tiab] OR professional*[tiab] OR setting*[tiab] AND (Neoplasms[mh] OR neoplas*[tiab] OR cancer[tiab] OR cancerous[tiab] OR carcinogenic*[tiab] OR tumor[tiab] OR tumors[tiab] OR tumorigen*[tiab] OR tumour*[tiab] OR leukemia*[tiab] OR lymphoma*[tiab] OR sarcoma*[tiab] OR carcinoma*[tiab] OR adenoma*[tiab] OR melanoma*[tiab] OR DNA damage[mh] OR dna-damage[tiab] OR genetic-damage[tiab] OR chromosom*[tiab] OR chromatid[tiab] OR mutagenicity tests[mh] OR mutagen*[tiab] OR genotox*[tiab] OR micronucle*[tiab]) OR (Reproduction[mh] OR reproduction[tiab] OR reproductive[tiab] OR sexual[tiab] OR intercourse[tiab] OR coitus[tiab] OR ejaculat*[tiab] OR orgasm*[tiab] OR fertiliz*[tiab] OR conception[tiab] OR gametogenesis[tiab] OR oogenesis[tiab] OR spermatogenesis[tiab] OR inseminat*[tiab] OR luteinization[tiab] OR ovulat*[tiab] OR anovulat*[tiab] OR superovulat*[tiab] OR erection*[tiab] OR Reproductive physiological phenomena[mh] OR andropause[tiab] OR menopause[tiab] OR estrous[tiab] OR fertility[tiab] OR fecund*[tiab] OR time-to-pregnancy[tiab] OR menstrua*[tiab] OR pubert*[tiab] OR adrenarche[tiab] OR menarche[tiab] OR Genital diseases, female[mh] OR endometriosis[tiab] OR infertil*[tiab] OR pelvic-inflamm*[tiab] OR genital diseases, male[mh] OR epididymitis[tiab] OR hypospadia*[tiab] OR priapism[tiab] OR prostatitis[tiab] OR cryptorchidism[tiab] OR varicocele[tiab] OR erectile-dysfunction[tiab] OR Pregnancy[mh] OR Pregnant[tiab] OR pregnancy[tiab] OR pregnancies[tiab] OR obstetric*[tiab] OR maternal*[tiab] OR mother*[tiab] OR embryonic and fetal development[mh] OR embryonic-development[tiab] OR fetal-development[tiab] OR organogenesis[tiab] OR body size[mh] OR birth-weight*[tiab] OR growth-retardation[tiab] OR small-for-gestation*[tiab] OR embryo[tiab] OR embryos[tiab] OR embryonic[tiab] OR fetus[tiab] OR foetus[tiab] OR fetal[tiab] OR foetal[tiab] OR gestation*[tiab] OR peripartum[tiab] OR postpartum[tiab] OR placent*[tiab] OR prenatal[tiab] OR perinat*[tiab] OR neonat*[tiab] OR postnat*[tiab] OR labor-onset[tiab] OR Parturition*[tiab] OR birth[tiab] OR births[tiab] OR childbirth*[tiab] OR gravidity[tiab] OR parity[tiab] OR Pregnancy complications[mh] OR pregnancy-complication*[tiab] OR complicated-pregnan*[tiab] OR labor-complication*[tiab] OR complicated-labor[tiab] OR gestational-diabetes[tiab] OR (pregnancy[tiab] AND hypertension[tiab]) OR maternal-death[tiab] OR morning-sickness[tiab] OR breech[tiab] OR ectopic[tiab] OR Pregnancy outcome[mh] OR pregnancy-outcome*[tiab] OR live-birth*[tiab] OR full-term[tiab] OR term-birth*[tiab] OR newborn*[tiab] OR baby[tiab] OR babies[tiab] OR infant*[tiab] OR premature[tiab] OR preterm[tiab] OR pre-term[tiab] OR stillbirth*[tiab] OR still-birth*[tiab] OR stillborn*[tiab] OR still-born*[tiab] OR spontaneous-abortion*[tiab] OR miscarriage*[tiab] OR infant mortality[mh] OR infant-mortality[tiab] OR embryo*-loss[tiab] OR congenital, hereditary, and neonatal diseases and

Database	Search Terms
	abnormalities[mh] OR congenital[tiab] OR abnormal*[tiab] OR malform*[tiab] OR retard*[tiab] OR prenatal exposure delayed effects[mh] OR teratogen*[tiab] OR embryotoxic*[tiab])
Scopus	TITLE-ABS-KEY(chemotherapy* OR anticancer OR antineoplastic* OR anti-tumor* OR anti-tumour* OR cytostatic OR hazardous-drug* OR drug*-hazardous OR category-d OR category-x OR "4 aminofolic acid" OR "4 epidoxorubicin" OR "5 fluorouracil" OR "6 mercaptopurine" OR "6 thioguanine" OR abraxane OR adrucil OR "all-trans retinoic acid" OR atra OR altretamine OR adriamycin OR "actinomycin D" OR aminopterin OR anastrozole OR "ARA-C" OR arimidex OR aromasin OR "behenoyl cytosine arabinoside" OR bevacizumab OR bhac OR bleomycin OR bortezomib OR busulfan OR busulfex OR carboplatin OR capecitabine OR carmustine OR cerubidine OR chlorambucil OR cisplatin OR cis-platinum OR cyclophosphamide OR cytarabine OR cytosar OR "cytosine arabinoside" OR cytoxan OR dacarbazine OR dasatinib OR daunorubicin OR daunoxome OR deltasone OR docetaxel OR doxorubicin OR efudex OR eldisine OR ellence OR eloxatin OR emcyt OR enocitabine OR epirubicin OR erlotinib OR etopophos OR etoposide OR estramustine OR exemestane OR fareston OR femara OR fludara OR fludarabine OR folex OR fulvestrant OR faslodex OR gefitinib OR gemcitabine OR gemtuzumab OR gemzar OR gleevec OR glivec OR herceptin OR hexamethylmelamine OR hydroxycarbamide OR hydroxyurea OR idarubicin OR ifex OR ifosfamide OR imatinib OR "interferon alpha" OR iressa OR irinotecan OR ixabepilone OR ixempra OR lapatinib OR letrozole OR lomustine OR matulane OR mechlorethamine OR melphalan OR methotrexate OR "mitomycin c" OR mitoxantrone OR mustargen OR "mustine Hcl" OR mutamycin OR myleran OR mylotarg OR navelbine OR nilotinib OR "nitrogen mustard HCl" OR nolvadex OR novantrone OR oncovin OR oxaliplatin OR ozogamicin OR paclitaxel OR paraplatin OR pemetrexed OR pentostatin OR platinol OR prednisone OR procarbazine OR rituxan OR rituximab OR sorafenib OR sprycel OR streptozocin OR sunitinib OR sunrabin OR sutent OR tamoxifen OR tarceva OR tasigna OR taxol OR taxotere OR temodar OR temozolomide OR teniposide OR thioplex OR thiotepa OR toposar OR topotecan OR toremifene OR trastuzumab OR tretinoin OR tykerb OR velban OR velcade OR vepesid OR vesanoid OR vinblastine OR vincasar OR vincrex OR vincristine OR vindesine OR vinorelbine OR vm26 OR vp16 OR vumon OR xeloda OR zanosar) AND TITLE-ABS-KEY(occupation* OR workplace OR work-related OR contaminat* OR handl*) AND TITLE-ABS-KEY((pharmacist* OR technician* OR nurse* OR nursing OR physician* OR clinician* OR doctor* OR veterinarian* OR hospital OR clinic OR medical OR health OR healthcare OR health-care OR pharmacy OR pharmaceutical* OR cancer OR oncolog* OR veterinary) AND (staff OR personnel OR worker* OR employee* OR technician* OR aide* OR assistant* OR professional* OR setting*))
TOXLINE Did not include PubMed articles (NOT PubMed [org] NOT pubdart [org])	(chemotherapy OR chemotherapies OR chemotherapeutic OR anticancer OR anti cancer OR antineoplastic* OR anti tumor OR anti tumorigenic OR anti tumour OR anti tumourigenic OR cytostatic) AND

Database	Search Terms
	(occupation* OR workplace OR work related OR contaminat* OR handl*) NOT AND (pharmacist OR pharmacists OR technician* OR nurse* OR nursing OR physician* OR clinician* OR doctor* OR veterinarian* OR hospital OR clinic OR medical OR health OR healthcare OR health care OR pharmacy OR pharmaceutical* OR cancer OR oncolog* OR veterinary) AND (staff OR personnel OR worker* OR employee* OR technician* OR aide* OR assistant* OR professional* OR setting*)
Web of Science	Topic=(Chemotherap* OR anticancer OR antineoplastic* OR anti- tumor* OR anti-tumour* OR cytostatic OR hazardous-drug* OR drug*-hazardous OR category-D OR category-x OR “4 aminofolic acid” OR “4 epidoxorubicin” OR "5 fluorouracil" OR "6 mercaptopurine" OR “6 thioguanine” OR Abraxane OR adrucil OR “all-trans retinoic acid” OR ATRA OR altretamine OR adriamycin OR “actinomycin D” OR aminopterin OR Anastrozole OR “ARA-C” OR arimidex OR aromasin OR “behenoyl cytosine arabinoside” OR bevacizumab OR BHAC OR bleomycin OR bortezomib OR busulfan OR busulfex OR carboplatin OR capecitabine OR carmustine OR Cerubidine OR chlorambucil OR cisplatin OR cis-platinum OR cyclophosphamide OR cytarabine OR cytosar OR “cytosine arabinoside” OR Cytoxan OR dacarbazine OR dasatinib OR daunorubicin OR daunoxome OR deltasone OR docetaxel OR doxorubicin OR efudex OR eldisine OR Ellence OR Eloxatin OR emcyt OR enocitabine OR epirubicin OR erlotinib OR etopophos OR etoposide OR estramustine OR exemestane OR fareston OR femara OR fludara OR fludarabine OR folex OR fulvestrant OR Faslodex OR gefitinib OR gemcitabine OR gemtuzumab OR gemzar OR gleevec OR glivec OR herceptin OR hexamethylmelamine OR hydroxycarbamide OR hydroxyurea OR idarubicin OR IFEX OR ifosfamide OR imatinib OR “interferon alpha” OR iressa OR irinotecan OR ixabepilone OR ixempra OR lapatinib OR letrozole OR lomustine OR matulane OR mechlorethamine OR melphalan OR methotrexate OR “mitomycin c” OR mitoxantrone OR mustargen OR “mustine Hcl” OR mutamycin OR myleran OR mylotarg OR navelbine OR nilotinib OR “nitrogen mustard HCl” OR nolvadex OR novantrone OR oncovin OR oxaliplatin OR ozogamicin OR paclitaxel OR paraplatin OR pemetrexed OR pentostatin OR platinol OR prednisone OR procarbazine OR rituxan OR rituximab OR sorafenib OR sprycel OR streptozocin OR sunitinib OR sunrabin OR sutent OR tamoxifen OR tarceva OR tasigna OR taxol OR taxotere OR temodar OR temozolomide OR teniposide OR thioplex OR thiotepa OR toposar OR topotecan OR toremifene OR trastuzumab OR tretinoin OR tykerb OR velban OR velcade OR vepesid OR vesanoid OR vinblastine OR vincasar OR vincrex OR vincristine OR vindesine OR vinorelbine OR VM26 OR VP16 OR Vumon OR Xeloda OR zanosar) AND Topic=((pharmacist* OR technician* OR nurse* OR nursing OR physician* OR clinician* OR doctor* OR veterinarian* OR hospital OR clinic OR medical OR health OR healthcare OR health-care OR pharmacy OR pharmaceutical* OR cancer OR oncolog* OR veterinary) AND (staff OR personnel OR worker* OR employee* OR technician* OR aide* OR assistant* OR professional* OR setting*)) AND Topic=(occupation* OR workplace OR work-related OR contaminat* OR handl*)

Database	Search Terms
	AND Topic=(neoplas* OR cancer OR cancerous OR carcinogenic* OR tumor OR tumors OR tumorigen* OR tumour* OR leukemia* OR lymphoma* OR sarcoma* OR carcinoma* OR adenoma* OR melanoma* OR dna-damage OR genetic-damage OR chromosom* OR chromatid OR mutagen* OR genotox* OR micronucle*) OR Topic=(reproduction OR reproductive OR sexual OR intercourse OR coitus OR ejaculat* OR orgasm* OR fertiliz* OR conception OR gametogenesis OR oogenesis OR spermatogenesis OR inseminat* OR luteinization OR ovulat* OR anovulat* OR superovulat* OR erection* OR andropause OR menopause OR estrous OR fertility OR fecund* OR time-to-pregnancy OR menstrua* OR pubert* OR adrenarche OR menarche OR endometriosis OR infertil* OR pelvic-inflamm* OR epididymitis OR hypospadia* OR priapism OR prostatitis OR cryptorchidism OR varicocele OR erectile-dysfunction OR pregnant OR pregnancy OR pregnancies OR obstetric* OR maternal* OR mother* OR embryonic-development OR fetal-development OR organogenesis OR "body size" OR birth-weight* OR growth-retardation OR small-for-gestation* OR embryo OR embryos OR embryonic OR fetus OR foetus OR fetal OR foetal OR gestation* OR peripartum OR postpartum OR placent* OR prenatal OR perinat* OR neonat* OR postnat* OR labor-onset OR Parturition* OR birth OR births OR childbirth* OR gravidity OR parity OR pregnancy-complication* OR complicated-pregnan* OR labor-complication* OR complicated-labor OR gestational-diabetes OR (pregnancy AND hypertension) OR maternal-death OR morning-sickness OR breech OR ectopic OR pregnancy-outcome* OR live-birth* OR full-term OR term-birth* OR newborn* OR baby OR babies OR infant* OR premature OR preterm OR pre-term OR stillbirth* OR still-birth* OR stillborn* OR still-born* OR spontaneous-abortion* OR miscarriage* OR infant-mortality OR embryo*-loss OR congenital OR abnormal* OR malform* OR retard* OR teratogen* OR embryotoxic*)

Appendix 2. Data Extraction Elements for Human Studies

Data Extraction Elements for Human Studies	
HUMAN	
Funding	Funding source(s)
	Reporting of conflict of interest (COI) by authors (*reporting bias)
Subjects	Study population name/description
	Dates of study and sampling time frame
	Geography (country, region, state, etc.)
	Demographics (sex, race/ethnicity, age or lifestage at exposure and at outcome assessment)
	Number of subjects (target, enrolled, n per group in analysis, and participation/follow-up rates) (*missing data bias)
	Inclusion/exclusion criteria/recruitment strategy (*selection bias)
	Description of reference group (*selection bias)
Methods	Study design (e.g., prospective or retrospective cohort, nested case-control study, cross-sectional, population-based case-control study, intervention, case report, etc.)
	Length of follow-up (*information bias)
	Health outcome category, e.g., cardiovascular
	Health outcome, e.g., blood pressure (*reporting bias)
	Diagnostic or methods used to measure health outcome (*information bias)
	Confounders or modifying factors and how considered in analysis (e.g., included in final model, considered for inclusion but determined not needed (*confounding bias)
	Substance name and CAS number
	Exposure assessment (e.g., blood, urine, hair, air, drinking water, job classification, residence, administered treatment in controlled study, etc.) (*information bias)
	Methodological details for exposure assessment (e.g., HPLC-MS/MS, limit of detection) (*information bias)
	Statistical methods (*information bias)
Results	Exposure levels (e.g., mean, median, measures of variance as presented in paper, such as SD, SEM, 75th/90th/95th percentile, minimum/maximum); range of exposure levels, number of exposed cases
	Statistical findings (e.g., adjusted β, standardized mean difference, adjusted odds ratio, standardized mortality ratio, relative risk, etc.) or description of qualitative results. When possible, OHAT will convert measures of effect to a common metric with associated 95% confidence intervals (CI). Most often, measures of effect for continuous data are expressed as mean difference, standardized mean difference, and percent control response. Categorical data are typically expressed as odds ratio, relative risk (RR, also called risk ratio), or β values, depending on what metric is most commonly reported in the included studies and on OHAT's ability to obtain information for effect conversions from the study or through author query.
	If not presented in the study, statistical power can be assessed during data extraction using an approach that can detect a 10% to 20% change from response by control or referent group for continuous data, or a relative risk or odds ratio of 1.5 to 2 for categorical data, using the prevalence of exposure or prevalence of outcome in the control or referent group to determine sample size. For categorical data where the sample sizes of exposed and control or referent groups differ, the sample size of the exposed group will be used to determine the relative power category. Recommended sample sizes to achieve 80% power for a given effect size, i.e., 10% or 20% change from control, will be compared to sample sizes used in the study to categorize statistical power as "appears to be adequately powered" (sample size for 80% power met), somewhat underpowered (sample size is 75%

Data Extraction Elements for Human Studies	
HUMAN	
	to < 100% of number required for 80% power), "underpowered" (sample size is 50% to < 75% of number required for 80% power), or "severely underpowered" (sample size is < 50% of number required for 80% power).
	Observations on dose response (e.g., trend analysis, description of whether dose-response shape appears to be monotonic, non-monotonic)
Other	Documentation of author queries, use of digital rulers to estimate data values from figures, exposure unit, and statistical result conversions, etc.

Appendix 3. Risk-of-Bias Criteria

The OHAT risk-of-bias tool for human and animal studies (version date January 2015 and available at http://ntp.niehs.nih.gov/go/38673) reflects OHAT's current best practices and provides the detailed discussion and instructions for the risk-of-bias practices used in this evaluation. The OHAT tool uses a single set of questions (also called "elements" or "domains") to assess risk of bias across various study types to facilitate consideration of conceptually similar potential sources of bias across the human and animal evidence streams with a common terminology. Individual risk-of-bias questions are designated as only applicable to certain study designs (e.g., cohort studies or experimental animal studies), and a subset of the questions apply to each study design (**Table 5**).

The current evaluation will only consider the human evidence stream. The specific criteria used to assess risk of bias for this evaluation are outlined below for human studies.

Observational Studies (Human studies)

Cohort studies

1. Was administered dose or exposure level adequately randomized? [NA]

2. Was allocation to study groups adequately concealed? [NA]

3. Did selection of study participants result in the appropriate comparison groups?

Risk-of-Bias Criteria for Appropriate Comparison Groups (Cohort Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that subjects (both exposed and non-exposed) were similar (e.g., recruited from the same eligible population, recruited with the same method of ascertainment using the same inclusion and exclusion criteria, and were of similar age and health status), recruited within the same time frame, and had the similar participation/response rates, • **Note:** A study will be considered low risk of bias if baseline characteristics of groups differed but these differences were considered as potential confounding or stratification variables (see question #4).
Probably Low Risk of Bias (+)
• Indirect evidence that subjects (both exposed and non-exposed) were similar (e.g., recruited from the same eligible population, recruited with the same method of ascertainment using the same inclusion and exclusion criteria, and were of similar age and health status), recruited within the same time frame, and had the similar participation/response rates, • **OR** differences between groups would not appreciably bias results.
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that subjects (both exposed and non-exposed) were not similar, recruited within very different time frames, or had the very different participation/response rates, • **OR** there is insufficient information provided about the comparison group including a different rate of non-response without an explanation (record "NR" as basis for answer).
Definitely High Risk of Bias (--)
• Direct evidence that subjects (both exposed and non-exposed) were not similar, recruited within very different time frames, or had the very different participation/response rates.

4. Did study design or analysis account for important confounding and modifying variables?

Risk-of-Bias Criteria for Confounding and Modifying Variables (Cohort Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that appropriate adjustments or explicit considerations were made for the variables listed below as potential confounders and/or effect measure modifiers in the final analyses through the use of statistical models to reduce research-specific bias including standardization, matching, adjustment in multivariate model, stratification, propensity scoring, or other methods that were appropriately justified. Acceptable consideration of appropriate adjustment factors includes cases when the factor is not included in the final adjustment model because the author conducted analyses that indicated it did not need to be included, • **AND** there is direct evidence that primary covariates and confounders were assessed using valid and reliable measurements, • **AND** there is direct evidence that other exposures anticipated to bias results were not present or were appropriately measured and adjusted for. In occupational studies or studies of contaminated sites, other chemical exposures known to be associated with those settings were appropriately considered. • **Note:** The following variables should be considered as potential confounders and/or effect measure modifiers for the relationship between occupational exposure to cancer chemotherapy and cancer are: age, sex, and tobacco smoking; and for genetic toxicity endpoints: age and smoking. **Note:** For occupational exposure to cancer chemotherapy, co-exposure to other known or suspected genetic toxicants will be considered as confounders (e.g., anesthetic gases, X-radiation and gamma radiation (i.e., radiation therapy)).
Probably Low Risk of Bias (+)
• Indirect evidence that appropriate adjustments were made, • **OR** it is deemed that not considering or only considering a partial list of covariates or confounders in the final analyses would not appreciably bias results, • **AND** there is evidence (direct or indirect) that covariates and confounders considered were assessed using valid and reliable measurements, • **OR** it is deemed that the measures used would not appreciably bias results (i.e., the authors justified the validity of the measures from previously published research), • **AND** there is evidence (direct or indirect) that other co-exposures anticipated to bias results were not present or were appropriately adjusted for, • **OR** it is deemed that co-exposures present would not appreciably bias results. • **Note:** this includes insufficient information provided on co-exposures in general population studies.
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that the distribution of important covariates and known confounders differed between the groups and was not appropriately adjusted for in the final analyses, • **OR** there is insufficient information provided about the distribution of known confounders (record "NR" as basis for answer), • **OR** there is indirect evidence that covariates and confounders considered were assessed using measurements of unknown validity, • **OR** there is insufficient information provided about the measurement techniques used to assess covariates and confounders considered (record "NR" as basis for answer), • **OR** there is indirect evidence that there was an unbalanced provision of additional co-exposures across the primary study groups, which were not appropriately adjusted for,

Risk-of-Bias Criteria for Confounding and Modifying Variables (Cohort Studies)
• **OR** there is insufficient information provided about co-exposures in occupational studies or studies of contaminated sites where high exposures to other chemical exposures would have been reasonably anticipated (record "NR" as basis for answer).
Definitely High Risk of Bias (--)
• Direct evidence that the distribution of important covariates and known confounders differed between the groups, confounding was demonstrated, and was not appropriately adjusted for in the final analyses, • **OR** there is direct evidence that covariates and confounders considered were assessed using non valid measurements, • **OR** there is direct evidence that there was an unbalanced provision of additional co-exposures across the primary study groups, which were not appropriately adjusted for.

5. Were experimental conditions identical across study groups? [NA]

6. Were the research personnel blinded to the study group during the study? [NA]

7. Were outcome data complete without attrition or exclusion from analysis?

Risk-of-Bias Criteria for Data Attrition or Exclusion (Cohort Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that loss of subjects (i.e., incomplete outcome data) was adequately addressed and reasons were documented when human subjects were removed from a study. • **Note**: Acceptable handling of subject attrition includes: very little missing outcome data; reasons for missing subjects unlikely to be related to outcome (for survival data, censoring unlikely to be introducing bias); missing outcome data balanced in numbers across study groups, with similar reasons for missing data across groups, • **OR** missing data have been imputed using appropriate methods and characteristics of subjects lost to follow up or with unavailable records are described in identical way and are not significantly different from those of the study participants.
Probably Low Risk of Bias (+)
• Indirect evidence that loss of subjects (i.e., incomplete outcome data) was adequately addressed and reasons were documented when human subjects were removed from a study, • **OR** it is deemed that the proportion lost to follow-up would not appreciably bias results. This would include reports of no statistical differences in characteristics of subjects lost to follow up or with unavailable records from those of the study participants. Generally, the higher the ratio of participants with missing data to participants with events, the greater potential there is for bias. For studies with a long duration of follow-up, some withdrawals for such reasons are inevitable.
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that loss of subjects (i.e., incomplete outcome data) was unacceptably large and not adequately addressed, • **OR** there is insufficient information provided about numbers of subjects lost to follow-up (record "NR" as basis for answer).
Definitely High Risk of Bias (--)
• Direct evidence that loss of subjects (i.e., incomplete outcome data) was unacceptably large and not adequately addressed. • **Note**: Unacceptable handling of subject attrition includes: reason for missing outcome data likely to be

Risk-of-Bias Criteria for Data Attrition or Exclusion (Cohort Studies)
related to true outcome, with either imbalance in numbers or reasons for missing data across study groups; or potentially inappropriate application of imputation.

8. Can we be confident in the exposure characterization?

Risk-of-Bias Criteria for Exposure Characterization (Cohort Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that exposure was consistently assessed (i.e., under the same method and time-frame) using well-established methods that directly measure exposure (e.g., measurement of parent or metabolites of cancer chemotherapy agents in blood, serum, or plasma), • **OR** exposure was assessed using less-established methods that directly measure exposure and are validated against well-established methods, • **AND** exposure was assessed in a relevant time-window for development of the outcome, • **AND** there is sufficient range or variation in exposure measurements across groups to potentially identify associations with health outcomes, • **AND** there is evidence that most of the exposure data measurements are above the limit of quantitation for the assay such that different exposure groups can be distinguished, • **AND** the study used spiked samples to confirm assay performance.
Probably Low Risk of Bias (+)
• Indirect evidence that the exposure was consistently assessed using well-established methods that directly measure exposure), • **OR** exposure was assessed using indirect measures (e.g., questionnaire or occupational exposure assessment by a certified industrial hygienist) that have been validated or empirically shown to be consistent with methods that directly measure exposure (i.e., inter-methods validation: one method vs. another), • **AND** exposure was assessed in a relevant time-window for development of the outcome, • **AND** there is sufficient range or variation in exposure measurements across groups to potentially identify associations with health outcomes (at a minimum from high exposure or ever exposed from low exposure or never exposed), • **AND** there is evidence that most of the exposure data measurements are above the limit of quantitation for the assay such that different exposure groups can be distinguished.
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that the exposure was assessed using poorly validated methods that directly measure exposure • **OR** there is evidence that the exposure was assessed using indirect measures that have not been validated or empirically shown to be consistent with methods that directly measure exposure (e.g., questionnaire, job-exposure matrix or self-report without validation) (record “NR” as basis for answer), • **OR** there is insufficient information provided about the exposure assessment, including validity and reliability, but no evidence for concern about the method used (record “NR” as basis for answer).
Definitely High Risk of Bias (--)
• Direct evidence that the exposure was assessed using methods with poor validity, • **OR** evidence of exposure misclassification (e.g., differential recall of self-reported exposure).

9. Can we be confident in the outcome assessment?

Risk-of-Bias Criteria for Outcome Assessment (Cohort Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that the adverse health outcomes or biomarkers of effect were assessed using well-established methods (e.g., gold standard) • **AND** subjects had been followed for the same length of time in all study groups, • **AND** there is direct evidence that the outcome assessors (including study subjects, if outcomes were self-reported) were adequately blinded to the study group or exposure level, and it is unlikely that they could have broken the blinding prior to reporting outcomes. • **Note:** Well-established methods will depend on the outcome, but examples of such methods may include: objectively measured chromosomal aberrations, DNA damage as measured by Comet assay, and micronuclei induction with standard assays with sufficiently low variation and limits of detection to allow discrimination between groups (or evidence that the assay could have detected a difference based on responses to a positive control); doctor diagnosis of cancer, immune or acute effects data obtained from medical records; obtained from registries (Shamliyan *et al.* 2010). • **Note:** Preferred reporting was used for genetic toxicity assays, which allowed for more direct comparison between studies; % of cells with structural chromosomal aberration (for chromosomal aberration studies); number of micronucleated cells/1000 cells scored or % of cells with micronuclei (micronuclei induction); and % tail DNA or % tail intensity (Comet assay). • **Note:** Because DNA repair enzymes begin to work quickly following exposure, studies reporting Comet assays results need to have: 1) collected the blood sample during or immediately after the work day and 2) chilled the sample until it was processed.
Probably Low Risk of Bias (+)
• Indirect evidence that the outcome was assessed using acceptable methods (i.e., deemed valid and reliable but not the gold standard), • **AND** subjects had been followed for the same length of time in all study groups • **OR** it is deemed that the outcome assessment methods used would not appreciably bias results, • **AND** there is indirect evidence that the outcome assessors (including study subjects, if outcomes were self-reported) were adequately blinded to the study group, and it is unlikely that they could have broken the blinding prior to reporting outcomes, • **OR** it is deemed that lack of adequate blinding of outcome assessors would not appreciably bias results, which is more likely to apply to objective outcome measures, • **Note:** Acceptable, but not ideal assessment methods will depend on the outcome, but examples of such methods may include proxy reporting of outcomes such as mining of data collected for other purposes. • **Note:** Use of common, but less than preferred, reporting for genetic toxicity assays reduces ability to directly compare across studies, such as: number of structural chromosomal aberrations (excluding gaps) per cells scored (chromosomal aberrations); number of micronucleated cells/1000 cells scored or frequency (%) of micronuclei (micronuclei induction); and tail length (μm), Olive tail moment, or tail length DNA damage index (Comet assay).
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that the outcome assessment method is an insensitive instrument (e.g., a questionnaire used to assess outcomes with no information on validation), • **OR** the length of follow up differed by study group, • **OR** there is indirect evidence that it was possible for outcome assessors (including study subjects if

Risk-of-Bias Criteria for Outcome Assessment (Cohort Studies)
outcomes were self-reported) to infer the study group prior to reporting outcomes, • **OR** there is insufficient information provided about blinding of outcome assessors (record "NR" as basis for answer). • **Note:** Comet assay studies that do not report data on when the blood sample was drawn or how it was stored (record "NR" as basis for answer) suggest samples may have been collected or handled in less than preferred methods. • **Note:** Chromosomal gaps occur naturally in unexposed populations and the implication of their presence is not known. Thus, the scoring of structural chromosomal aberrations (including gaps) per cells evaluated may introduce variability in the measurement.
Definitely High Risk of Bias (--)
• Direct evidence that the outcome assessment method is an insensitive instrument, • **OR** the length of follow up differed by study group, • **OR** there is direct evidence for lack of adequate blinding of outcome assessors (including study subjects if outcomes were self-reported), including no blinding or incomplete blinding. • **Note:** Comet assay studies that report the blood sample was drawn 24 hours after completion of the work day or state that the sample was stored at room temperature indicate that less than preferred methods were used.

10. Were all measured outcomes reported?

Risk-of-Bias Criteria for Outcome Reporting (Cohort Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that all of the study's measured outcomes (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have been reported. This would include outcomes reported with sufficient detail to be included in meta-analysis or fully tabulated during data extraction and analyses had been planned in advance.
Probably Low Risk of Bias (+)
• Indirect evidence that all of the study's measured outcomes (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have been reported, • **OR** analyses that had not been planned in advance (i.e., retrospective unplanned subgroup analyses) are clearly indicated as such and deemed that unplanned analyses were appropriate and selective reporting would not appreciably bias results (e.g., appropriate analyses of an unexpected effect). This would include outcomes reported with insufficient detail such as only reporting that results were statistically significant (or not).
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that all of the study's measured outcomes (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have not been reported, • **OR** and there is indirect evidence that unplanned analyses were included that may appreciably bias results, • **OR** there is insufficient information provided about selective outcome reporting (record "NR" as basis for answer).

Risk-of-Bias Criteria for Outcome Reporting (Cohort Studies)
Definitely High Risk of Bias (--)
• Direct evidence that all of the study's measured outcomes (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have not been reported. In addition to not reporting outcomes, this would include reporting outcomes based on composite score without individual outcome components or outcomes reported using measurements, analysis methods or subsets of the data (e.g., subscales) that were not pre-specified or reporting outcomes not pre-specified, or that unplanned analyses were included that would appreciably bias results.

11. Were there no other potential threats to internal validity?

This question will be used to examine individual studies for appropriate statistical methods (e.g., confirmation of homogeneity of variance for ANOVA and other statistical tests that require normally distributed data). It will also be used for risk-of-bias considerations that do not fit under the other questions.

Cross Sectional and Case Series Studies

1. Was administered dose or exposure level adequately randomized? [NA]

2. Was allocation to study groups adequately concealed? [NA]

3. Did selection of study participants result in the appropriate comparison groups? [NA to Case reports or case series]

Risk-of-Bias Criteria for Appropriate Comparison Groups (Cross Sectional Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that subjects (both exposed and non-exposed) were similar (e.g., recruited from the same eligible population, recruited with the same method of ascertainment using the same inclusion and exclusion criteria, and were of similar age and health status), recruited within the same time frame, and had the similar participation/response rates, • **Note:** A study will be considered low risk of bias if baseline characteristics of groups differed but these differences were considered as potential confounding or stratification variables (see question #4). •
Probably Low Risk of Bias (+)
• Indirect evidence that subjects (both exposed and non-exposed) were similar (e.g., recruited from the same eligible population, recruited with the same method of ascertainment using the same inclusion and exclusion criteria, and were of similar age and health status), recruited within the same time frame, and had the similar participation/response rates, • **OR** differences between groups would not appreciably bias results.
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that subjects (both exposed and non-exposed) were not similar, recruited within very different time frames, or had the very different participation/response rates, • **OR** there is insufficient information provided about the comparison group including a different rate of non-response without an explanation (record "NR" as basis for answer).
Definitely High Risk of Bias (--)
• Direct evidence that subjects (both exposed and non-exposed) were not similar, recruited within very different time frames, or had the very different participation/response rates.

4. Did study design or analysis account for important confounding and modifying variables?

Risk-of-Bias Criteria for Confounding and Modifying Variables (Cross Sectional and Case Series Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that appropriate adjustments or explicit considerations were made for the variables listed below as potential confounders and/or effect measure modifiers in the final analyses through the use of statistical models to reduce research-specific bias including standardization, matching, adjustment in multivariate model, stratification, propensity scoring, or other methods that were appropriately justified. Acceptable consideration of appropriate adjustment factors includes cases when the factor is not included in the final adjustment model because the author conducted analyses that indicated it did not need to be included, • **AND** there is direct evidence that primary covariates and confounders were assessed using valid and reliable measurements, • **AND** there is direct evidence that other exposures anticipated to bias results were not present or were appropriately measured and adjusted for. In occupational studies or studies of contaminated sites,

Risk-of-Bias Criteria for Confounding and Modifying Variables (Cross Sectional and Case Series Studies)
other chemical exposures known to be associated with those settings were appropriately considered. • **Note:** The following variables should be considered as potential confounders and/or effect measure modifiers for the relationship between occupational exposure to cancer chemotherapy and cancer are: age, sex, and tobacco smoking; and for genetic toxicity endpoints: age and smoking. • **Note:** For occupational exposure to cancer chemotherapy, co-exposure to other known or suspected genetic toxicants will be considered as confounders (e.g., anesthetic gases, X-radiation and gamma radiation (i.e., radiation therapy)).
Probably Low Risk of Bias (+)
• Indirect evidence that appropriate adjustments were made, • **OR** it is deemed that not considering or only considering a partial list of covariates or confounders in the final analyses would not appreciably bias results, • **AND** there is evidence (direct or indirect) that covariates and confounders considered were assessed using valid and reliable measurements, • **OR** it is deemed that the measures used would not appreciably bias results (i.e., the authors justified the validity of the measures from previously published research), • **AND** there is evidence (direct or indirect) that other co-exposures anticipated to bias results were not present or were appropriately adjusted for, • **OR** it is deemed that co-exposures present would not appreciably bias results. • **Note:** this includes insufficient information provided on co-exposures in general population studies.
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that the distribution of important covariates and known confounders differed between the groups and was not appropriately adjusted for in the final analyses, • **OR** there is insufficient information provided about the distribution of known confounders (record "NR" as basis for answer), • **OR** there is indirect evidence that covariates and confounders considered were assessed using measurements of unknown validity, • **OR** there is insufficient information provided about the measurement techniques used to assess covariates and confounders considered (record "NR" as basis for answer), • **OR** there is indirect evidence that there was an unbalanced provision of additional co-exposures across the primary study groups, which were not appropriately adjusted for, • **OR** there is insufficient information provided about co-exposures in occupational studies or studies of contaminated sites where high exposures to other chemical exposures would have been reasonably anticipated (record "NR" as basis for answer).
Definitely High Risk of Bias (--)
• Direct evidence that the distribution of important covariates and known confounders differed between the groups, confounding was demonstrated, and was not appropriately adjusted for in the final analyses, • **OR** there is direct evidence that covariates and confounders considered were assessed using non valid measurements, • **OR** there is direct evidence that there was an unbalanced provision of additional co-exposures across the primary study groups, which were not appropriately adjusted for.

5. Were experimental conditions identical across study groups? [NA]

6. Were the research personnel blinded to the study group during the study? [NA]

7. Were outcome data complete without attrition or exclusion from analysis?

Risk-of-Bias Criteria for Data Attrition or Exclusion (Cross-Sectional and Case Series Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
Probably Low Risk of Bias (+)
• Indirect evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that exclusion of subjects from analyses was not adequately addressed, • **OR** there is insufficient information provided about why subjects were removed from the study or excluded from analyses (record "NR" as basis for answer).
Definitely High Risk of Bias (--)
• Direct evidence that exclusion of subjects from analyses was not adequately addressed. • **Note**: Unacceptable handling of subject exclusion from analyses includes: reason for exclusion likely to be related to true outcome, with either imbalance in numbers or reasons for exclusion across study groups.

8. Can we be confident in the exposure characterization?

Risk-of-Bias Criteria for Exposure Characterization (Cross-Sectional and Case Series Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that exposure was consistently assessed (i.e., under the same method and time-frame) using well-established methods that directly measure exposure (e.g., measurement of parent or metabolites of cancer chemotherapy agents in blood, serum, or plasma), • **OR** exposure was assessed using less-established methods that directly measure exposure and are validated against well-established methods, • **AND** exposure was assessed in a relevant time-window for development of the outcome, • **AND** there is sufficient range or variation in exposure measurements across groups to potentially identify associations with health outcomes, • **AND** there is evidence that most of the exposure data measurements are above the limit of quantitation for the assay such that different exposure groups can be distinguished, • **AND** the study used spiked samples to confirm assay performance.
Probably Low Risk of Bias (+)
• Indirect evidence that the exposure was consistently assessed using well-established methods that directly measure exposure), • **OR** exposure was assessed using indirect measures (e.g., questionnaire or occupational exposure assessment by a certified industrial hygienist) that have been validated or empirically shown to be consistent with methods that directly measure exposure (i.e., inter-methods validation: one method vs. another), • **AND** exposure was assessed in a relevant time-window for development of the outcome., • **AND** there is sufficient range or variation in exposure measurements across groups to potentially identify associations with health outcomes (at a minimum from high exposure or ever exposed from low exposure or never exposed),

Risk-of-Bias Criteria for Exposure Characterization (Cross-Sectional and Case Series Studies)
• **AND** there is evidence that most of the exposure data measurements are above the limit of quantitation for the assay such that different exposure groups can be distinguished.
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that the exposure was assessed using poorly validated methods that directly measure exposure • **OR** there is evidence that the exposure was assessed using indirect measures that have not been validated or empirically shown to be consistent with methods that directly measure exposure (e.g., a job-exposure matrix or self-report without validation) (record "NR" as basis for answer), • **OR** there is insufficient information provided about the exposure assessment, including validity and reliability, but no evidence for concern about the method used (record "NR" as basis for answer).
Definitely High Risk of Bias (--)
• Direct evidence that the exposure was assessed using methods with poor validity, • **OR** evidence of exposure misclassification (e.g., differential recall of self-reported exposure).

9. Can we be confident in the outcome assessment?

Risk-of-Bias Criteria for Outcome Assessment (Cross Sectional and Case Series Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that the adverse health outcomes or biomarkers of effect were assessed using well-established methods (e.g., gold standard) • **AND** subjects had been followed for the same length of time in all study groups, • **AND** there is direct evidence that the outcome assessors (including study subjects, if outcomes were self-reported) were adequately blinded to the study group or exposure level, and it is unlikely that they could have broken the blinding prior to reporting outcomes. • **Note:** Well-established methods will depend on the outcome, but examples of such methods may include: objectively measured chromosomal aberrations, DNA damage as measured by Comet assay, and micronuclei induction with standard assays with sufficiently low variation and limits of detection to allow discrimination between groups (or evidence that the assay could have detected a difference based on responses to a positive control); doctor diagnosis of cancer, immune or acute effects data obtained from medical records; obtained from registries (Shamliyan *et al.* 2010). • **Note:** Preferred reporting was used for genetic toxicity assays, which allowed for more direct comparison between studies; % of cells with structural chromosomal aberration (for chromosomal aberration studies); number of micronucleated cells/1000 cells scored or % of cells with micronuclei (micronuclei induction); and % tail DNA or % tail intensity (Comet assay). • **Note:** Because DNA repair enzymes begin to work quickly following exposure, studies reporting Comet assays results need to have: 1) collected the blood sample during or immediately after the work day and 2) chilled the sample until it was processed.

Risk-of-Bias Criteria for Outcome Assessment (Cross Sectional and Case Series Studies)
Probably Low Risk of Bias (+)
• Indirect evidence that the outcome was assessed using acceptable methods, • **OR** it is deemed that the outcome assessment methods used would not appreciably bias results, • **AND** there is indirect evidence that the outcome assessors were adequately blinded to the exposure level, and it is unlikely that they could have broken the blinding prior to reporting outcomes, • **OR** it is deemed that lack of adequate blinding of outcome assessors would not appreciably bias results (including that subjects self-reporting outcomes were likely not aware of reported links between the exposure and outcome lack of blinding is unlikely to bias a particular outcome). • **Note:** Acceptable, but not ideal assessment methods will depend on the outcome, but examples of such methods may include proxy reporting of outcomes such as mining of data collected for other purposes. • **Note:** Use of common, but less than preferred, reporting for genetic toxicity assays reduces ability to directly compare across studies, such as: number of structural chromosomal aberrations (excluding gaps) per cells scored (chromosomal aberrations); number of micronuclei/1000 cells scored or frequency (%) of micronuclei (micronuclei induction); and tail length (μm), Olive tail moment, or tail length DNA damage index (Comet assay).
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that the outcome assessment method is an insensitive instrument, • **OR** there is indirect evidence that it was possible for outcome assessors to infer the exposure level prior to reporting outcomes (including that subjects self-reporting outcomes were likely aware of reported links between the exposure and outcome), • **OR** there is insufficient information provided about blinding of outcome assessors (record "NR" as basis for answer). • **Note:** Comet assay studies that do not report data on when the blood sample was drawn or how it was stored (record "NR" as basis for answer) suggest samples may have been collected or handled in less than preferred methods. • **Note:** Chromosomal gaps occur naturally in unexposed populations and the implication of their presence is not known. Thus, the scoring of structural chromosomal aberrations (including gaps) per cells evaluated may introduce variability in the measurement.
Definitely High Risk of Bias (--)
• Direct evidence that the outcome assessment method is an insensitive instrument, • **OR** there is direct evidence that outcome assessors were aware of the exposure level prior to reporting outcomes (including that subjects self-reporting outcomes were aware of reported links between the exposure and outcome). • **Note:** Comet assay studies that report the blood sample was drawn 24 hours after completion of the work day or state that the sample was stored at room temperature indicate that less than preferred methods were used.

10. Were all measured outcomes reported?

Risk-of-Bias Criteria for Outcome Reporting (Cross Sectional and Case Series Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that all of the study's measured outcomes (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have been reported. This would include outcomes reported with sufficient detail to be included in meta-analysis or fully tabulated during data extraction and analyses had been planned in advance.

Risk-of-Bias Criteria for Outcome Reporting (Cross Sectional and Case Series Studies)
Probably Low Risk of Bias (+)
• Indirect evidence that all of the study's measured outcomes (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have been reported, • **OR** analyses that had not been planned in advance (i.e., retrospective unplanned subgroup analyses) are clearly indicated as such and deemed that unplanned analyses were appropriate and selective reporting would not appreciably bias results (e.g., appropriate analyses of an unexpected effect). This would include outcomes reported with insufficient detail such as only reporting that results were statistically significant (or not).
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that all of the study's measured outcomes (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have not been reported, • **OR** and there is indirect evidence that unplanned analyses were included that may appreciably bias results, • **OR** there is insufficient information provided about selective outcome reporting (record "NR" as basis for answer).
Definitely High Risk of Bias (--)
• Direct evidence that all of the study's measured outcomes (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have not been reported. In addition to not reporting outcomes, this would include reporting outcomes based on composite score without individual outcome components or outcomes reported using measurements, analysis methods or subsets of the data (e.g., subscales) that were not pre-specified or reporting outcomes not pre-specified, or that unplanned analyses were included that would appreciably bias results.

11. Were there no other potential threats to internal validity?

This question will be used to examine individual studies for appropriate statistical methods (e.g., confirmation of homogeneity of variance for ANOVA and other statistical tests that require normally distributed data). It will also be used for risk-of-bias considerations that do not fit under the other questions.

Case Control Studies

1. Was administered dose or exposure level adequately randomized? [NA]

2. Was allocation to study groups adequately concealed? [NA]

3. Did selection of study participants result in the appropriate comparison groups?

Risk-of-Bias Criteria for Appropriate Comparison Groups (Case Control Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that cases and controls were similar (e.g., recruited from the same eligible population including being of similar age, gender, ethnicity, and eligibility criteria other than outcome of interest as appropriate), recruited within the same time frame, and controls are described as having no history of the outcome, • **Note:** A study will be considered low risk of bias if baseline characteristics of groups differed but these differences were considered as potential confounding or stratification variables (see question #4).
Probably Low Risk of Bias (+)
• Indirect evidence that cases and controls were similar (e.g., recruited from the same eligible population, recruited with the same method of ascertainment using the same inclusion and exclusion criteria, and were of similar age), recruited within the same time frame, and controls are described as having no history of the outcome, • **OR** it is deemed differences between cases and controls would not appreciably bias results.
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that controls were drawn from a very dissimilar population than cases or recruited within very different time frames, • **OR** there is insufficient information provided about the appropriateness of controls including rate of response reported for cases only (record "NR" as basis for answer).
Definitely High Risk of Bias (--)
• Direct evidence that controls were drawn from a very dissimilar population than cases or recruited within very different time frames.

4. Did study design or analysis account for important confounding and modifying variables?

Risk-of-Bias Criteria for Confounding and Modifying Variables (Case Control Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that appropriate adjustments were made for the variables listed below as potential confounders and/or effect measure modifiers in the final analyses through the use of statistical models to reduce research-specific bias including standardization, matching of cases and controls, adjustment in multivariate model, stratification, propensity scoring, or other methods were appropriately justified, • **AND** there is direct evidence that primary covariates and confounders were assessed using valid and reliable measurements, • **AND** there is direct evidence that other exposures anticipated to bias results were not present or were appropriately measured and adjusted for. • **Note:** The following variables should be considered as potential confounders and/or effect measure modifiers for the relationship between occupational exposure to cancer chemotherapy and cancer are: age, sex, and tobacco smoking; and for genetic toxicity endpoints: age and smoking. • **Note:** For occupational exposure to cancer chemotherapy, co-exposure to other known or suspected

Risk-of-Bias Criteria for Confounding and Modifying Variables (Case Control Studies)
genetic toxicants will be considered as confounders (e.g., anesthetic gases, X-radiation and gamma radiation (i.e., radiation therapy)).
Probably Low Risk of Bias (+)
• Indirect evidence that appropriate adjustments were made, • **OR** it is deemed that not considering or only considering a partial list of covariates or confounders in the final analyses would not appreciably bias results, • **AND** there is evidence (direct or indirect) that covariates and confounders considered were assessed using valid and reliable measurements, • **OR** it is deemed that the measures used would not appreciably bias results (i.e., the authors justified the validity of the measures from previously published research), • **AND** there is evidence (direct or indirect) that other co-exposures anticipated to bias results were not present or were appropriately adjusted for, • **OR** it is deemed that co-exposures present would not appreciably bias results. • **Note:** this includes insufficient information provided on co-exposures in general population studies.
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that the distribution of important covariates and known confounders differed between cases and controls and was not investigated further, • **OR** there is insufficient information provided about the distribution of known confounders in cases and controls (record "NR" as basis for answer), • **OR** there is indirect evidence that covariates and confounders considered were assessed using measurements of unknown validity, • **OR** there is insufficient information provided about the measurement techniques used to assess covariates and confounders considered (record "NR" as basis for answer), • **OR** there is indirect evidence that there was an unbalanced provision of additional co-exposures across cases and controls, which were not appropriately adjusted for, • **OR** there is insufficient information provided about co-exposures in occupational studies or studies of contaminated sites where high exposures to other chemical exposures would have been reasonably anticipated (record "NR" as basis for answer).
Definitely High Risk of Bias (--)
• Direct evidence that the distribution of important covariates and known confounders differed between cases and controls, confounding was demonstrated, but was not appropriately adjusted for in the final analyses, • **OR** there is direct evidence that covariates and confounders considered were assessed using non valid measurements, • **OR** there is direct evidence that there was an unbalanced provision of additional co-exposures across cases and controls, which were not appropriately adjusted for.

5. Were experimental conditions identical across study groups? [NA]

6. Were the research personnel blinded to the study group during the study? [NA]

7. Were outcome data complete without attrition or exclusion from analysis?

Risk-of-Bias Criteria for Data Attrition or Exclusion (Case Control Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that exclusion of subjects from analyses was adequately addressed, and reasons were

Risk-of-Bias Criteria for Data Attrition or Exclusion (Case Control Studies)
documented when subjects were removed from the study or excluded from analyses.
Probably Low Risk of Bias (+)
• Indirect evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that exclusion of subjects from analyses was not adequately addressed, • **OR** there is insufficient information provided about why subjects were removed from the study or excluded from analyses (record "NR" as basis for answer).
Definitely High Risk of Bias (--)
• Direct evidence that exclusion of subjects from analyses was not adequately addressed. • **Note**: Unacceptable handling of subject exclusion from analyses includes: reason for exclusion likely to be related to true outcome, with either imbalance in numbers or reasons for exclusion across study groups.

8. Can we be confident in the exposure characterization?

Risk-of-Bias Criteria for Exposure Characterization (Case Control Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that exposure was consistently assessed (i.e., under the same method and time-frame) using well-established methods that directly measure exposure (e.g., measurement of parent or metabolites of cancer chemotherapy agents in blood, serum, or plasma), • **OR** exposure was assessed using less-established methods that directly measure exposure and are validated against well-established methods, • **AND** exposure was assessed in a relevant time-window for development of the outcome, • **AND** there is sufficient range or variation in exposure measurements across groups to potentially identify associations with health outcomes, • **AND** there is evidence that most of the exposure data measurements are above the limit of quantitation for the assay such that different exposure groups can be distinguished, • **AND** the study used spiked samples to confirm assay performance.
Probably Low Risk of Bias (+)
• Indirect evidence that the exposure was consistently assessed using well-established methods that directly measure exposure), • **OR** exposure was assessed using indirect measures (e.g., questionnaire or occupational exposure assessment by a certified industrial hygienist) that have been validated or empirically shown to be consistent with methods that directly measure exposure (i.e., inter-methods validation: one method vs. another), • **AND** exposure was assessed in a relevant time-window for development of the outcome., • **AND** there is sufficient range or variation in exposure measurements across groups to potentially identify associations with health outcomes (at a minimum from high exposure or ever exposed from low exposure or never exposed), • **AND** there is evidence that most of the exposure data measurements are above the limit of quantitation for the assay such that different exposure groups can be distinguished.
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that the exposure was assessed using poorly validated methods that directly measure exposure,

Risk-of-Bias Criteria for Exposure Characterization (Case Control Studies)
• **OR** there is direct evidence that the exposure was assessed using indirect measures that have not been validated or empirically shown to be consistent with methods that directly measure exposure (e.g., a job-exposure matrix or self-report without validation) (record "NR" as basis for answer), • **OR** there is insufficient information provided about the exposure assessment, including validity and reliability, but no evidence for concern about the method used (record "NR" as basis for answer).
Definitely High Risk of Bias (--)
• Direct evidence that the exposure was assessed using methods with poor validity, • **OR** evidence of exposure misclassification (e.g., differential recall of self-reported exposure).

9. Can we be confident in the outcome assessment?

Risk-of-Bias Criteria for Outcome Assessment (Case Control Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that the adverse health outcomes or biomarkers of effect were assessed using well-established methods (e.g., gold standard) • **AND** subjects had been followed for the same length of time in all study groups, • **AND** there is direct evidence that the outcome assessors (including study subjects, if outcomes were self-reported) were adequately blinded to the study group or exposure level, and it is unlikely that they could have broken the blinding prior to reporting outcomes. • **Note:** Well-established methods will depend on the outcome, but examples of such methods may include: objectively measured chromosomal aberrations, DNA damage as measured by Comet assay, and micronuclei induction with standard assays with sufficiently low variation and limits of detection to allow discrimination between groups (or evidence that the assay could have detected a difference based on responses to a positive control); doctor diagnosis of cancer, immune or acute effects data obtained from medical records; obtained from registries (Shamliyan *et al.* 2010). • **Note:** Preferred reporting was used for genetic toxicity assays, which allowed for more direct comparison between studies; % of cells with structural chromosomal aberration (for chromosomal aberration studies); number of micronucleated cells/1000 cells scored or % of cells with micronuclei (micronuclei induction); and % tail DNA or % tail intensity (Comet assay). • **Note:** Because DNA repair enzymes begin to work quickly following exposure, studies reporting Comet assays results need to have: 1) collected the blood sample during or immediately after the work day and 2) chilled the sample until it was processed.

Risk-of-Bias Criteria for Outcome Assessment (Case Control Studies)
Probably Low Risk of Bias (+)
• Indirect evidence that the outcome was assessed in cases (i.e., case definition) and controls using acceptable methods), • **AND** subjects had been followed for the same length of time in all study groups, • **OR** it is deemed that the outcome assessment methods used would not appreciably bias results, • **AND** there is indirect evidence that the outcome assessors were adequately blinded to the exposure level when reporting outcomes, • **OR** it is deemed that lack of adequate blinding of outcome assessors would not appreciably bias results (including that subjects self-reporting outcomes were likely not aware of reported links between the exposure and outcome or lack of blinding is unlikely to bias a particular outcome). • **Note:** Acceptable, but not ideal assessment methods will depend on the outcome, but examples of such methods may include proxy reporting of outcomes such as mining of data collected for other purposes. • **Note:** Use of common, but less than preferred, reporting for genetic toxicity assays reduces ability to directly compare across studies, such as: number of structural chromosomal aberrations (excluding gaps) per cells scored (chromosomal aberrations); number of micronuclei/1000 cells scored or frequency (%) of micronuclei (micronuclei induction); and tail length (μm), Olive tail moment, or tail length DNA damage index (Comet assay).
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that the outcome was assessed in cases (i.e., case definition) using an insensitive instrument, • **OR** there is insufficient information provided about how cases were identified (record "NR" as basis for answer). • **OR** there is indirect evidence that it was possible for outcome assessors to infer the exposure level prior to reporting outcomes (including that subjects self-reporting outcomes were likely aware of reported links between the exposure and outcome), • **OR** there is insufficient information provided about blinding of outcome assessors (record "NR" as basis for answer). • **Note:** Comet assay studies that do not report data on when the blood sample was drawn or how it was stored (record "NR" as basis for answer) suggest samples may have been collected or handled in less than preferred methods. • **Note:** Chromosomal gaps occur naturally in unexposed populations and the implication of their presence is not known. Thus, the scoring of structural chromosomal aberrations (including gaps) per cells evaluated may introduce variability in the measurement.
Definitely High Risk of Bias (--)
• Direct evidence that the outcome was assessed in cases (i.e., case definition) using an insensitive instrument, • **OR** there is direct evidence that outcome assessors were aware of the exposure level prior to reporting outcomes (including that subjects self-reporting outcomes were aware of reported links between the exposure and outcome). • **Note:** Comet assay studies that report the blood sample was drawn 24 hours after completion of the work day or state that the sample was stored at room temperature indicate that less than preferred methods were used.

10. Were all measured outcomes reported?

Risk-of-Bias Criteria for Outcome Reporting (Case Control Studies)
Definitely Low Risk of Bias (++)
• Direct evidence that all of the study's measured outcomes (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have been reported. This would include outcomes reported with sufficient detail to be included in meta-analysis or fully tabulated during data extraction and analyses had been planned in advance.
Probably Low Risk of Bias (+)
• Indirect evidence that all of the study's measured outcomes (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have been reported, • **OR** analyses that had not been planned in advance (i.e., retrospective unplanned subgroup analyses) are clearly indicated as such and deemed that unplanned analyses were appropriate and selective reporting would not appreciably bias results (e.g., appropriate analyses of an unexpected effect). This would include outcomes reported with insufficient detail such as only reporting that results were statistically significant (or not).
Probably High Risk of Bias (-) or (NR)
• Indirect evidence that all of the study's measured outcomes (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have not been reported, • **OR** and there is indirect evidence that unplanned analyses were included that may appreciably bias results, • **OR** there is insufficient information provided about selective outcome reporting (record "NR" as basis for answer).
Definitely High Risk of Bias (--)
• Direct evidence that all of the study's measured outcomes (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have not been reported. In addition to not reporting outcomes, this would include reporting outcomes based on composite score without individual outcome components or outcomes reported using measurements, analysis methods or subsets of the data (e.g., subscales) that were not pre-specified or reporting outcomes not pre-specified, or that unplanned analyses were included that would appreciably bias results.

11. Were there no other potential threats to internal validity?

This question will be used to examine individual studies for appropriate statistical methods (e.g., confirmation of homogeneity of variance for ANOVA and other statistical tests that require normally distributed data). It will also be used for risk-of-bias considerations that do not fit under the other questions.