Review

Health Effects of Naphthalene Exposure: A Systematic Evidence Map and Analysis of Potential Considerations for Dose–Response Evaluation

Erin E. Yost,[1] Audrey Galizia,[1] Dustin F. Kapraun,[1] Amanda S. Persad,[1] Suryanarayana V. Vulimiri,[1] Michelle Angrish,[1] Janice S. Lee,[1] and Ingrid L. Druwe[1]

[1]Center for Public Health and Environmental Assessment, U.S. Environmental Protection Agency, Washington, District of Columbia, USA

BACKGROUND: Naphthalene is a polycyclic aromatic hydrocarbon that has been associated with health effects, including cancer. As the state of the science on naphthalene toxicity continues to evolve, updated toxicity reference value(s) may be required to support human health risk assessment.

OBJECTIVES: We present a systematic evidence map of studies that could be used to derive toxicity reference value(s) for naphthalene.

METHODS: Human and animal health effect studies and physiologically based pharmacokinetic (PBPK) models were identified from a literature search based on populations, exposures, comparators, and outcomes (PECO) criteria. Human and animal studies meeting PECO criteria were refined to a smaller subset considered most informative for deriving chronic reference value(s), which are preferred for assessing risk to the general public. This subset was evaluated for risk of bias and sensitivity, and the suitability of each study for dose–response analysis was qualitatively assessed. Lowest observed adverse effect levels (LOAELs) were extracted and summarized. Other potentially relevant studies (e.g., mechanistic and toxicokinetic studies) were tracked as supplemental information but not evaluated further. Existing reference values for naphthalene are also summarized.

RESULTS: We identified 26 epidemiology studies and 16 animal studies that were considered most informative for further analysis. Eleven PBPK models were identified. The available epidemiology studies generally had significant risk of bias and/or sensitivity concerns and were mostly found to have low suitability for dose–response analysis due to the nature of the exposure measurements. The animal studies had fewer risk of bias and sensitivity concerns and were mostly found to be suitable for dose–response analysis.

CONCLUSION: Although both epidemiological and animal studies of naphthalene provide weight of evidence for hazard identification, the available animal studies appear more suitable for reference value derivation. PBPK models and mechanistic and toxicokinetic data can be applied to extrapolate these animal data to humans, considering mode of action and interspecies metabolic differences. https://doi.org/10.1289/EHP7381

Introduction

Naphthalene is a polycyclic aromatic hydrocarbon that is found naturally in fossil fuels (e.g., coal tar, petroleum) and biomass combustion (ATSDR 2005). It is also produced commercially and is considered a high production–volume chemical (U.S. EPA 2016). Major commercial uses of naphthalene are in the manufacture of chemical intermediates (e.g., phthalic anhydrides), dyes, surfactants, leather tanning agents, dispersants, pesticides, resins, and solvents, and the major consumer products containing naphthalene are moth repellents and toilet deodorant blocks (ATSDR 2005). The public can be exposed to naphthalene primarily through airborne emissions from industrial sources, open burning, traffic exhaust, cigarettes, and off-gassing of naphthalene-containing products (Jia and Batterman 2010; ATSDR 2005). Children may have additional susceptibility to naphthalene exposure though the ingestion of mothballs or contaminated soil (ATSDR 2005).

Naphthalene exposure has been associated with cancer and noncancer health effects, with most of the evidence coming from animal studies and human case studies (ATSDR 2005; U.S. EPA 1998). The current weight of evidence for naphthalene carcinogenicity is derived primarily from 2-y inhalation bioassays conducted in rodents by the National Toxicology Program (NTP) (NTP 1992b, 2000). These bioassays found that naphthalene exposure increased the incidence of nasal respiratory epithelial adenomas (a benign tumor) and olfactory epithelial neuroblastomas of the nose (a malignant tumor that has rarely been observed in NTP bioassays) in both male and female rats (NTP 2000) and pulmonary alveolar/bronchiolar adenomas (a benign tumor) in female mice (NTP 1992b). In humans, cancer following naphthalene exposure has been documented in two case series reports: increased laryngeal cancer incidence in a German naphthalene purification plant (Wolf 1976, 1978) and increased colorectal cancer among Nigerian patients with a history of taking a naphthalene-containing indigenous treatment for anorectal problems (Ajao et al. 1988). Based on these findings, the 14th NTP Report on Carcinogens classified naphthalene as reasonably anticipated to be a human carcinogen (NTP 2016), and the International Agency for Research on Cancer (IARC) classified naphthalene to be possibly carcinogenic to humans (Group 2B) (IARC 2002); both reports cited sufficient evidence in experimental animals and inadequate evidence in humans. The assessment of naphthalene by the U.S. Environmental Protection Agency (U.S. EPA) Integrated Risk Information System (IRIS) was released prior to the publication of NTP's 2-y study in rats (NTP 2000) and classified naphthalene as a possible human carcinogen (Group C), based on limited evidence of carcinogenicity in animals and inadequate data in humans (U.S. EPA 1998).

Reference values for noncancer effects have also been developed for naphthalene by multiple government entities. The term "reference value" refers to a value designed to provide an exposure limit at which some protection to human health can be inferred. Reference values are the most common final output from the dose–response assessment component of the risk assessment paradigm set forth by the National Research Council (NRC 1983, 2009). Examples of reference values for naphthalene include a reference dose (RfD) for oral exposure and reference concentration (RfC) for inhalation exposure from IRIS (U.S. EPA 1998); oral and inhalation minimal risk levels (MRLs) from the Agency of Toxic Substances and Disease Registry (ATSDR) (ATSDR 2005); and a reference exposure level (REL) from the California Environmental Protection Agency's Office of Environmental Health Hazard Assessment (OEHHA) (OEHHA 2000).

The scientific literature characterizing naphthalene toxicity continues to evolve, with hundreds of studies relevant to the health effects and mechanisms of naphthalene exposure published over the past decade [see U.S. EPA Health & Environment Research Online (HERO) database for naphthalene (U.S. EPA 2021)]. These include health effect studies in humans and animals, which provide information on the dose-related effects of naphthalene exposure; physiologically based pharmacokinetic (PBPK) models, which can be used to develop extrapolations between species and routes of exposure; and mechanistic and toxicokinetic studies that can help inform biological plausibility, human relevance of animal data, and mode of action (MOA) for dose–response analysis. These data could be applied by entities such as government agencies to develop updated reference value(s) for naphthalene that reflect the state of the science on this chemical.

This study identifies the reference values that are currently available for naphthalene and presents a systematic evidence map of the available data that could be used to develop or revise toxicity reference value(s) for exposure to naphthalene in the general public and assess cancer and noncancer outcomes, based on a survey of the literature through January 2021. To our knowledge, this approach involves the first use of systematic review methodologies to characterize the available literature on naphthalene toxicity. The focus of our evaluation is on the quality and availability of health effect studies in humans and animals that could serve as the basis for hazard identification and dose–response analysis, and on the availability of PBPK models that can be used for dose extrapolation. We first provide a broad summary of the range of health effect studies that are available for naphthalene and then perform a more in-depth analysis on a smaller subset of studies we considered to be most relevant for deriving reference value(s) for chronic exposure. Chronic reference values generally are preferred over reference values of shorter duration (e.g., acute, subchronic) for risk assessment of environmental exposures in the general public. The availability of mechanistic and toxicokinetic/absorption, distribution, metabolism, excretion (ADME) studies were also tracked as part of this evidence mapping exercise because they may serve as supporting information in the context of a chemical assessment, although an analysis of mechanistic and toxicokinetic data is not performed here.

Characterizing the evidence map for naphthalene not only demonstrates that availability of data that can be used to inform naphthalene risk assessment but also highlights data gaps and research needs that can be considered by researchers to design studies that are informative for chemical assessment purposes.

Methods

Scoping and problem formulation for this evidence map was conducted by considering the regulatory needs for chemical assessment in the U.S. EPA context, summarized most recently in an IRIS Assessment Plan (IAP) for naphthalene that underwent public comment in July 2018 (U.S. EPA 2018b). Literature search and screening were performed to support the IRIS assessment of naphthalene and were conducted in several rounds between 2013 and 2021, with methods evolving somewhat over time to reflect assessment needs and available technology. Results are documented in online databases that may allow readers to interact with the data more easily [HERO, the U.S. EPA's Health Assessment Workspace Collaborative (HAWC), and Tableau Public], with links provided throughout this text, but hard copies of the data are also in the Excel file provided as supplemental material with this publication. The methods used to conduct the evidence map are based on those described in the public comment

draft of the U.S. EPA's *Office of Research and Development Staff Handbook for Developing IRIS Assessments* (version 1.0, November 2020) (U.S. EPA 2020) and have previously been reviewed as part of the National Academies of Sciences, Engineering, and Medicine report on the implementation of systematic review by the IRIS Program (NASEM 2018). An overview of the key elements of the process is provided here.

Survey of Existing Reference Values for Naphthalene

A survey of existing reference values can be an excellent secondary resource when evaluating the health effects of a chemical because it provides insight into the types of assessments that have already been conducted, the age of those assessments, and the key studies used for reference value derivation. Existing reference values for naphthalene were surveyed in March 2021 by searching the list of national, state, and international agencies shown in Table S1 of the supplemental materials. Additionally, U.S. EPA's ToxVal database (a compilation of publicly available reference values from across the Internet) (U.S. EPA 2018a) was searched for any additional reference values from other sources. Information on derivation of each value (health effect, point of departure, uncertainty factors applied) was extracted whenever available.

This survey of reference values for naphthalene strives to be comprehensive, but it may miss values that are not publicly available. The list of sources used in this search was compiled by the U.S. EPA and consists of government agencies and recognized expert groups that develop reference values for the purpose of supporting regulatory decision-making. The list of sources for inhalation reference values was originally documented in a 2009 U.S. EPA report (U.S. EPA 2009) and has since been expanded to include oral reference values.

Populations, Exposures, Comparators, and Outcomes (PECO) Criteria

A set of PECO criteria (Table 1) were developed to serve as a guide for screening relevant health effect studies for naphthalene. The PECO identified human and mammalian animal health effect studies of naphthalene with appropriate control groups as the focus of the systematic evidence map, as well as any published PBPK models for naphthalene.

Literature Searches

Literature searches were conducted in three online scientific databases [PubMed, Web of Science (WOS), TOXLINE] in February 2013, December 2014, November 2015, January 2017, September 2017, February 2019, and January 2021. (The TOXLINE database was migrated to PubMed after the 2019 literature search update, so was not included in the 2021 literature search.) The initial search in 2013 was conducted without date limitations, and all subsequent searches were date-limited to the previous search. The search strategy included key terms related to PECO criteria and terms for specific experimental animal species. The January 2017 search added terms to the PubMed query looking for information on naphthalene metabolites (1,4-naphthoquinone; 1,2-naphthoquinone; naphthalene 1,2-oxide; and 1,2-dihydroxy-1,2-dihydronaphthalene). Additionally, Toxic Substances Control Act Test Submissions (TSCATS) were identified by searching TSCATS 2, TSCATS 1, the U.S. EPA's Chemical Data Access Tool, and Google searches for TSCA recent submissions. These search strategies are summarized in Table S2 of the supplemental materials.

The results of this literature search were supplemented by the following: *a*) manually searching citations from published review articles and national and international health agency documents;

Table 1. Populations, exposures, comparators, outcomes (PECO) criteria.

PECO element	Evidence
Populations	Human: Any population and life stage (occupational or general population, including children and other sensitive populations). The following study designs will be considered most informative: controlled exposure, cohort, case–control, cross-sectional, and ecological. Case reports and case series will be tracked as "supplemental information." Animal: Nonhuman mammalian animal species (whole organism) of any life stage (including preconception, *in utero*, lactation, peripubertal, and adult stages). Studies of transgenic animals will be tracked as mechanistic studies under "supplemental information."
Exposures	Human: Any exposure to naphthalene (CASRN 91-20-3), including occupational exposures, via oral, inhalation, or dermal route[s]. Animal: Any exposure to naphthalene (CASRN 91-20-3) via oral, inhalation, or dermal route[s]. Studies involving exposures to mixtures will be included only if they include an arm with exposure to naphthalene alone. Other exposure routes, including injection, will be tracked during title and abstract screening and tagged as "supplemental information."
Comparators	Human: A comparison or referent population exposed to lower levels (or no exposure/exposure below detection limits) of naphthalene, or exposure to naphthalene for shorter periods of time. Animal: A concurrent control group exposed to vehicle-only treatment.
Outcomes	All health outcomes (both cancer and noncancer). In general, end points related to clinical diagnostic criteria, disease outcomes, histopathological examination, or other apical/phenotypic outcomes will be prioritized for evidence synthesis over outcomes such as biochemical measures.
PBPK	Studies describing PBPK models for naphthalene will be included.

Note: PBPK, physiologically based pharmacokinetic.

b) "backward" searches (to identify articles cited by included studies, reviews, or prior assessments by other agencies); *c*) "forward" searches (to identify articles that cite those studies); *d*) searching a combination of Chemical Abstracts Service Registry Numbers (CASRNs) and synonyms for chemical assessment-related websites; and *e*) addition of references that had been previously added to the HERO database as part of an earlier U.S. EPA naphthalene review effort. A description of these additional search strategies is provided in Table S3 of the supplemental materials.

The results of these literature search strategies are compiled in the U.S. EPA's HERO database (https://hero.epa.gov/hero/index.cfm/project/page/project_id/3064). The complete list of studies identified in the literature search is also provided in the Excel file in the supplemental material.

Literature Screening

PECO criteria were used to determine inclusion or exclusion of human and animal health effect studies and published PBPK models. Screening methods evolved over time, reflecting the technology that was available at the time of each literature search. The results of this screening process, including documentation of studies included and excluded at each screening stage, are publicly available in the HERO database (https://hero.epa.gov/hero/index.cfm/project/page/project_id/3064). Tagging information for the complete list of studies identified in the literature search is also provided in the Excel file in the supplemental material.

For literature searches conducted through November 2015, all records were first electronically screened in EndNote with a set of terms intended to prioritize "on-topic" references for title and abstract review (inclusion terms are listed in Table S4 of the supplemental materials). Some of the electronic inclusion terms were generic (i.e., not chemical-specific) and were intended to capture health effect studies of any type. Other terms were specific to naphthalene and were based on previous knowledge of health effects and possible mechanisms of toxicity. Citations containing (in title, abstract, or keywords) at least one inclusion term related to health outcomes, epidemiological or toxicological study design, toxicokinetics, or mechanistic information proceeded to title/abstract screening. Citations that did not contain at least one inclusion term were subjected to a quality control check to verify that relevant references are not missed. Specifically, a random sample ($\sim 10\%$) of the electronically excluded citations were subjected to title/abstract review to confirm that the electronic screening process produced acceptable results (i.e., no relevant citations were inadvertently missed). If the random sample contained at least one potentially relevant citation, the list of electronic screening terms was revised to add terms pertaining to the missing citation, and the electronic screening process was repeated. This quality control and revision process was repeated as many times as necessary to ensure that relevant studies were retained for title/abstract screening. Citations that did not contain at least one term inclusion term were excluded from further review. For literature searches conducted after November 2015, no electronic screen was performed due to the smaller number of new records identified, and all studies underwent title/abstract screening.

Title/abstract screening was conducted by two independent reviewers using PECO criteria to guide the inclusion or exclusion of human and animal health effect studies and PBPK models. Title/abstract screening was conducted using EndNote (for literature searches conducted between 2013 and 2017), SWIFT-Active Screener software (for literature search conducted in 2019) (https://swift.sciome.com/activescreener), or DistillerSR (for literature search conducted in 2021) (https://www.evidencepartners.com/products/distillersr-systematic-review-software/). Screening was performed manually on all platforms (the machine learning functionality of SWIFT-Active Screener was not used), and therefore we do not anticipate that the screening results were affected by the type of software used. For citations with no abstract, articles were screened based on all or some of the following: title relevance, page numbers (articles two pages or fewer in length may be assumed to be conference reports, editorials, or letters), and PubMed Medical Subject Headings (MeSH; e.g., a study might not have been considered further if there were no human health or biology-related MeSH terms). In addition to PECO criteria, the following exclusion criteria were applied: *a*) study materials that were not peer reviewed; and *b*) records that did not contain original data, such as assessments by government agencies, review articles, editorials, or commentaries. Non-English studies were tracked during screening and tagged for possible further evaluation but were not translated or reviewed further for this evidence map.

An attempt was made to retrieve full texts for all studies not excluded during title/abstract screening, and full text review was performed to identify the final list of studies meeting PECO criteria. Conference abstracts and studies for which the full text was found to be unavailable following title/abstract screening were tracked but not reviewed further. At both the title/abstract and full text screening levels, screening conflicts were resolved by discussion among the primary screeners, with consultation of a third reviewer or technical advisor (if needed) to resolve any disagreements. When there were multiple publications using the

same or overlapping data, all publications on the research were included and one (generally the publication with the most complete reporting of results) was selected as the primary record. For instance, NTP's 2-y study in rats is described in a study report (NTP 2000) as well as in two publications (Long et al. 2003; Abdo et al. 2001) and a pathology report; we considered those to be one distinct reference and cited it as the NTP report (NTP 2000), which had the most complete description of results.

In addition to studies meeting PECO criteria, the following types of studies containing potentially relevant supplemental information were also tracked during the screening process and are documented in HERO: mechanistic studies (including *in vitro* and *in silico* models); ADME/toxicokinetic studies; human case reports or case series; and animal studies with exposure routes other than oral, inhalation, and dermal (e.g., intraperitoneal injection). These studies were tracked based on title/abstract only (full-text screening not performed) and are not evaluated further in this evidence map.

Survey of Studies Meeting PECO Criteria

The human and animal studies that met PECO criteria were extracted and briefly summarized in an interactive dashboard in Tableau Public, which can be viewed at the following website: https://public.tableau.com/views/NaphthaleneEvidenceMap/ReadMe?:language=en&:display_count=y&publish=yes&:origin=viz_share_link. The Excel file used to create the Tableau Public database is provided as supplemental material. For human studies, information was captured on population type (e.g., general population, occupational), study type (e.g., cross-sectional, cohort, case–control), description of study population, major route of exposure (if known), description of how exposure was assessed, health outcome(s) assessed, and health outcome(s) observed. For animal studies, information was captured on species, strain, sex, dose or concentration levels tested, duration of treatment, route of exposure, health outcome(s) assessed, and health outcome(s) observed. Doses in animal studies reported as parts per million (ppm; inhalation exposures) were converted to milligrams per cubic meter, based on a molecular weight of 128.1705 g/mol.

All PBPK models for naphthalene identified in the literature search were summarized and qualitatively assessed for scientific and technical suitability for use in a human health risk assessment. We compared the models and identified relationships between them, such as cases in which one model was a revision of a previous model or a synthesis of multiple previous models.

Selection of Human and Animal Studies for Further Evaluation

Study evaluation can be a time-consuming process, so it is pragmatic to prioritize the evaluation of studies that are most relevant to assessment needs—in this case, studies that can be used to derive chronic reference value(s) for naphthalene. We therefore further screened the human and animal studies that met PECO criteria to identify a subset of studies that are more likely to be relevant for chronic reference value derivation based on the criteria outlined below and focused all further evaluation on that subset of studies:

- Animal studies with chronic and subchronic exposure durations were prioritized for further evaluation. Studies with exposures less than 30 d in duration were only included if they could contribute critical information to the weight of evidence. For instance, if both <30-d and >30-d exposures were reported as part of a study by the same laboratory group, all exposure durations were included because they can help inform the dose- and time-related development of health effects. All reproductive and developmental exposure

studies were included regardless of exposure duration because short exposures may coincide with windows of susceptibility.
- Animal studies evaluating multiple dose levels were prioritized over studies evaluating single dose levels.
- For both human and animal studies, health systems (e.g., respiratory, reproductive, etc.) were prioritized for further evaluation when the available evidence from multiple studies suggested an association with naphthalene exposure.

Study Evaluation of Human and Animal Studies

The subset of human and animal studies identified in the "Selection of Human and Animal Studies for Further Evaluation" section above were evaluated for their validity and utility in assessing health effects of chemical exposure by applying the U.S. EPA IRIS study evaluation method. This method is described in the *Staff Handbook for Developing IRIS Assessments* (U.S. EPA 2020) and in several previous publications; for instance, development and validation of this study evaluation method are described by Dishaw et al. (2020) and Radke et al. (2019), and this study evaluation method has been used in the systematic evidence map by Keshava et al. (2020) and systematic reviews including Radke et al. (2018) and Yost et al. (2019). The key concerns addressed by the study evaluations were risk of bias (factors that may affect the magnitude or direction of an effect) and insensitivity (factors that limit the ability to detect a true effect; low sensitivity is a bias toward the null when an effect exists). These evaluations addressed the study's utility for identification of individual hazards but did not address the usability of a study for dose–response analysis, which was considered separately (see "Analysis of Dose–Response Considerations" section below).

Human studies were evaluated by consideration of the following domains: participant selection, exposure methods sensitivity, outcome measures, confounding, analysis, selective reporting, and sensitivity. For animal studies, the following domains were considered: reporting quality; allocation; observational bias/blinding; confounding; selective reporting and attrition; chemical administration and characterization; exposure timing, frequency, and duration; end point sensitivity and specificity; and results presentation. These study evaluation domains were designed to be parallel between human and animal studies but not exactly matching. A description of each domain is provided in Tables S5–S7, including core questions and basic considerations used to guide reviewers in the evaluation of each domain.

Two reviewers evaluated each human and animal health effect study to identify characteristics that would bear on the informativeness (i.e., validity and sensitivity) of the results and provide additional chemical- or outcome-specific knowledge or methodological concerns. For studies that examined more than one outcome, the evaluation process was performed separately for each outcome because the utility of a study can vary for different outcomes. For each outcome in a study, reviewers reached a consensus judgment of *good*, *adequate*, *deficient*, *not reported*, or *critically deficient* for each evaluation domain. If a consensus was not reached between two reviewers, a third reviewer performed conflict resolution. The judgments were defined as follows:

- *Good* represented a judgment that the study was conducted appropriately in relation to the evaluation domain, and any deficiencies, if present, were minor and would not be expected to influence the study results.
- *Adequate* indicated a judgment that there were methodological limitations relating to the evaluation domain but that those limitations were not likely to be severe or to have a notable impact on the results.

- *Deficient* denoted identified biases or deficiencies that were interpreted as likely to have had a notable impact on the results or that prevented interpretation of the study findings.
- *Not reported* indicated that the information necessary to evaluate the domain question was not available in the study. Generally, this term carried the same functional interpretation as *deficient* for the purposes of the study confidence classification (described below).
- *Critically deficient* reflected a judgment that the study conduct introduced a serious flaw that made the study uninterpretable. Studies with a determination of *critically deficient* in an evaluation domain were considered overall *uninformative*, as described below. Given this potential for exclusion, this classification was used infrequently and with extreme care. Serious flaws that did not warrant study exclusion were classified as *deficient*.

After each domain was rated, the identified strengths and limitations across the domains were considered as a whole to reach a study confidence classification of *high*, *medium*, or *low* confidence, or *uninformative* for each health outcome reported by the study. This classification included consideration by the reviewers of the likely impact that the identified bias, insensitivity, or inadequate reporting could have on the results. The classifications were defined as follows:
- *High confidence*: A well-conducted study with no notable deficiencies or concerns identified; the potential for bias was unlikely or minimal, and the study used sensitive methodology. *High* confidence studies generally reflected judgments of *good* across all or most evaluation domains.
- *Medium confidence*: A satisfactory (acceptable) study where deficiencies or concerns were noted, but the limitations were unlikely to be of a notable degree. Generally, *medium* confidence studies included *adequate* or *good* judgments across most domains, with the impact of any identified limitation not being judged as severe.
- *Low confidence*: A substandard study where deficiencies or concerns were noted, and the potential for bias or inadequate sensitivity could have a significant impact on the study results or their interpretation. Typically, *low* confidence studies had a *deficient* evaluation for one or more domains, although some *medium* confidence studies had a *deficient* rating in domain(s) considered to have less influence on the magnitude or direction of effect estimates.
- *Uninformative*: An unacceptable study where serious flaw(s) made the study results unusable for informing hazard identification. Studies with *critically deficient* judgments in any evaluation domain were classified as *uninformative* (see explanation above). *Uninformative* studies would not be considered for hazard identification or dose–response analysis but might be used to highlight possible research gaps.

Reporting quality (the extent to which a study reports sufficient details to allow for an evaluation of risk of bias and sensitivity) was considered within all study evaluation domains but was also included as a stand-alone domain for animal studies and served as a triage for determining whether a study reported sufficient experimental details to be evaluated. Animal studies failing to report information considered critical for study evaluation (species, test article name, levels and duration of exposure, route, qualitative or quantitative results for at least one end point of interest) would be rated *critically deficient* in the reporting quality domain and excluded from further consideration.

All study evaluation ratings are documented and publicly available in U.S. EPA's version of HAWC, a free and open-source, web-based software application (https://hawcprd.epa.gov/summary/assessment/566/visuals/). An export of the study

evaluations from HAWC is provided in the Excel file in the supplemental material.

Data Extraction

Quantitative outcome measurements from all *high* and *medium confidence* animal studies were extracted into HAWC (available at https://hawcprd.epa.gov/study/assessment/566/; at the site, click on a study to view extracted data), and the lowest observed adverse effect level (LOAEL) for each outcome measurement was identified based on author-reported statistical significance. For studies in which an exposure-related effect was observed but the authors did not perform a statistical analysis, LOAELs were determined based on expert judgment. These LOAELs were recorded in HAWC. Data extractions were carried out using a controlled vocabulary for health system, organ, and effect type to facilitate the organization and retrieval of information. Results from general histopathology evaluations were typically only extracted into HAWC if an exposure-related response was reported; body weight changes presented as growth curves were not extracted into HAWC. An export of the extracted dose–response information from HAWC is provided in the Excel file in the supplemental material.

Analysis of Dose–Response Considerations

We finally evaluated whether each of the selected human or animal health effect studies would be suitable for dose–response analysis, based on the following considerations [similar to those used by Keshava et al. (2020)]:
- For both human and animal studies, quantitative exposure–response data was necessary to be considered suitable for dose–response analysis.
- Animal studies were considered more suitable for dose–response analysis if multiple dose levels were evaluated. Studies evaluating only a single dose level were considered less suitable.
- Epidemiology studies that used biomarker measurements in tissues or bodily fluids as the metric for exposure were considered suitable for dose–response analysis only if data or PBPK models were available to extrapolate between the reported biomarker measurements and the level of exposure.
- Epidemiology studies with limited ability to assess temporality between exposure and response were considered to have lower suitability for dose–response analysis. In particular, due to the short half-life of naphthalene in the body, cross-sectional study designs were considered to have limited ability to assess temporality unless the outcome was of an immediate nature.

In addition to these general considerations, the evaluation of suitability for dose–response analysis also considered any other study-specific concerns that could interfere with the interpretation of the exposure–response relationship. The identification of these study-specific concerns was based on expert judgment and discussion among reviewers, and the rationale for the judgment about each study is documented in tables herein.

Summary of the Evidence Base for Major Health Systems

The human and animal evidence base was qualitatively summarized for each of the health systems that were prioritized for further evaluation. The focus was on the availability of studies that could be potential candidates for reference value derivation, i.e., *high* or *medium confidence* studies that were considered suitable for dose–response analysis. *Low confidence* studies or studies with limited suitability for dose–response analysis would generally not be used for reference value derivation. Summaries also

highlighted the biological relevance of various outcome measurements for risk assessment.

Finally, to evaluate relative sensitivity to naphthalene exposure, the lowest LOAELs within each health system were summarized according to the route and duration of exposure. LOAELs reflect the exposure levels that caused an effect in each health system and therefore are an appropriate metric for evaluating sensitivity, but the NOAEL or (ideally) a benchmark dose would be the preferred values used for reference value derivation. Dose–response modeling for these studies to calculate a benchmark dose is outside the scope of this evidence map because our goal was to summarize only author-reported data.

Results

Survey of Existing Reference Values for Naphthalene

Reference values for naphthalene derived by the U.S. EPA and other national, state, and international agencies are depicted in Figure 1 (inhalation exposure) and Figure 2 (oral exposure). These figures indicate the exposure durations (acute, short-term,

subchronic, chronic) and population/exposure scenario that each reference value is designed to assess. Inhalation reference values for naphthalene include values used for risk assessment in emergency-response situations (designed to assess acute exposure to hazardous chemical releases), occupational exposure (8-h time-weighted averages designed to protect workers for a 5-d work week for 40 y), or exposure of the general public. All oral reference values for naphthalene are intended to assess risk in the general public. See the Supplemental Materials Appendix Table A1 (inhalation reference values) and Appendix Table A2 (oral reference values) for a tabular summary of these values, including information on how each value was derived. Values identified from sources that did not provide derivation details or were based on another agency's values are not shown in Figures 1–2 but are summarized in Appendix Table A3.

Focusing on the values intended for risk assessment in the general public, all chronic or intermediate duration inhalation and oral reference values for naphthalene from U.S. federal agencies, state agencies, and Health Canada were derived based on studies that are evaluated in this evidence map. The ATSDR chronic inhalation MRL and the Minnesota Department of Health's Chronic Health-
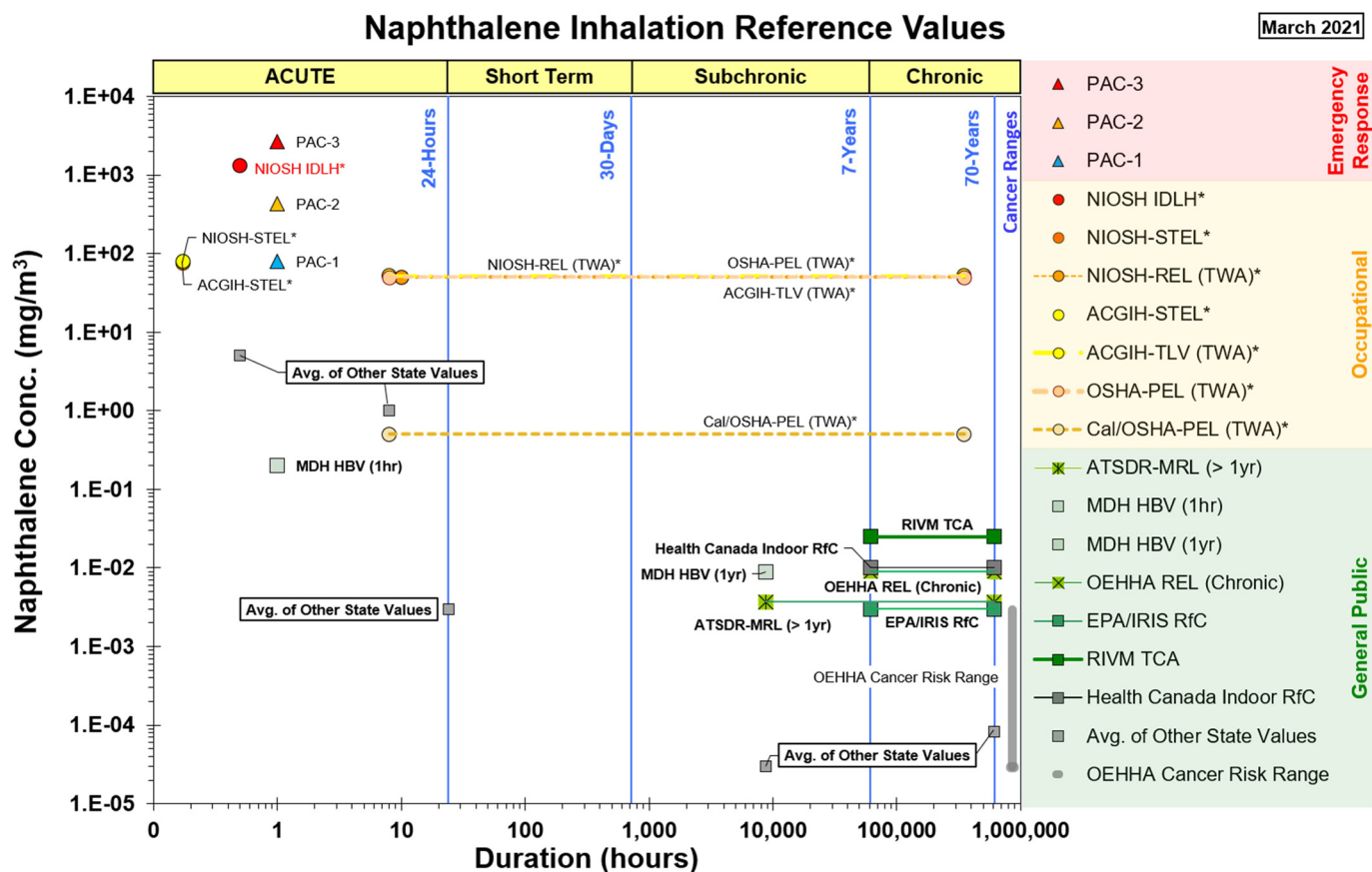
Figure 1. Available health effect reference values for inhalation exposure to naphthalene. See Supplemental Materials Appendix Table A1 for a tabular summary, including information on how each value was derived. Categories for the reference values based on their intended purpose are shown in the legend: red for Emergency Response, gold for Occupational, and green for values applicable to the General Public. Note: ACGIH, American Conference of Government Industrial Hygienists; ATSDR, Agency for Toxic Substances and Disease Registry; HBV, Health-Based Value; IDLH, immediately dangerous to life and health; IRIS, Integrated Risk Information System; MDH, Minnesota Department of Health; MRL, minimal risk level; NIOSH, National Institute for Occupational Safety and Health; OEHHA, California Environmental Protection Agency's Office of Environmental Health Hazard Assessment; OSHA, Occupational Safety and Health Administration; PAC, protective action criteria; PEL, Permissible Exposure Limit; REL, recommended exposure limit (NIOSH) or reference exposure level; RfC, reference concentration; RIVM, *Rijksinstituut voor Volksgezondheid en Milieu* (Netherlands Institute for Public Health and the Environment); STEL, short-term exposure limit; TCA, tolerable concentration; TLV, threshold limit value; TWA, time-weighted average.
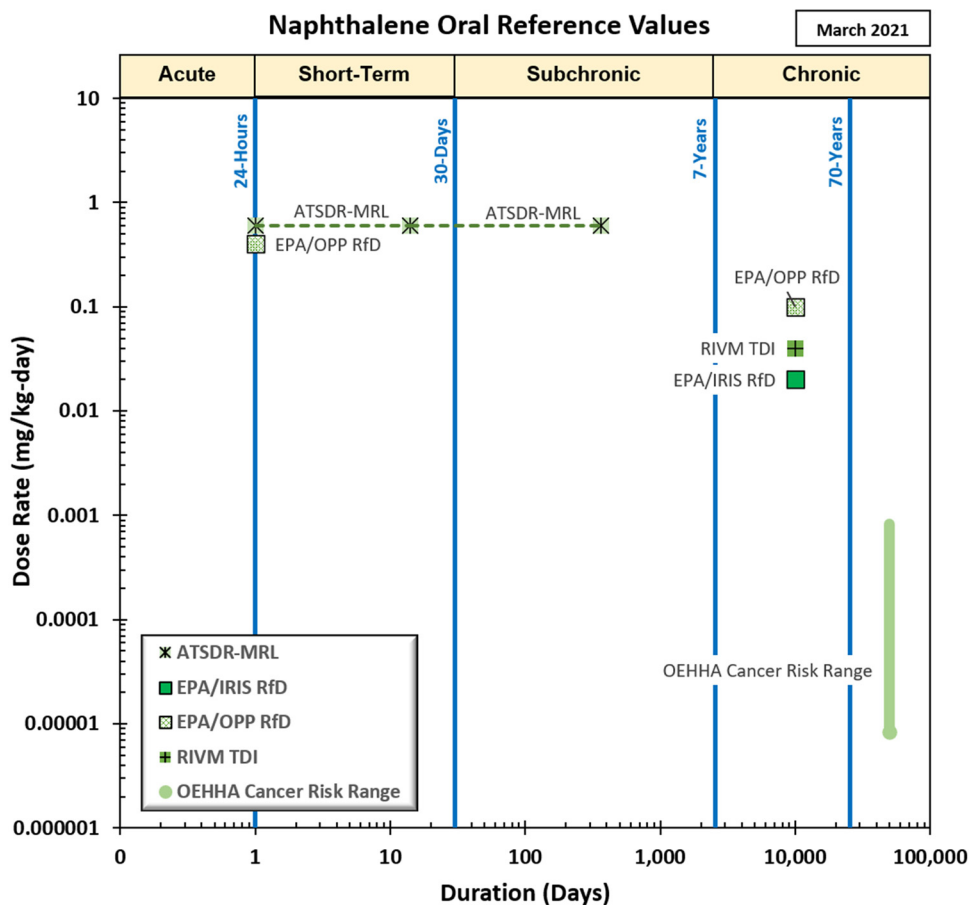
Figure 2. Available health effect reference values for oral exposure to naphthalene. See Supplemental Materials Appendix Table A2 for a tabular summary, including information on how each value was derived. All values in this figure are intended for application in the general public. Note: ATSDR, Agency for Toxic Substances and Disease Registry; IRIS, Integrated Risk Information System; MRL, minimal risk level; OPP, Office of Pesticide Programs; RfD, reference dose; RIVM, Rijksinstituut voor Volksgezondheid en Milieu (Netherlands Institute for Public Health and the Environment); TDI, tolerable daily intake.

Based Value are based on respiratory lesions in NTP's 2-y studies in mice and rats (NTP 1992b, 2000), whereas the IRIS RfC and California OEHHA's REL were both developed prior to the publication of NTP (2000) and are based on NTP (1992b). Health Canada's residential indoor RfC is based on NTP (2000). Additionally, several states (Nevada, Rhode Island, Oregon) derived chronic or 1-y inhalation values based on the OEHHA cancer unit risk factor, which is based on NTP's 2-y inhalation bioassays. Chronic oral RfDs from IRIS and the U.S. EPA Office of Pesticide Programs are based on the 90-d exposure study in rats by Battelle (1980b) (an unpublished report by a contract laboratory for NTP), and the ATSDR intermediate and acute oral MRLs are based on the gestational exposure study in rats by NTP (1991). The quality of these studies is discussed in subsequent sections.

Literature Search and Screening

Literature search and screening results are summarized in Figure 3. The database searches and additional search strategies identified 17,763 records, which were narrowed to 125 records that met PECO criteria (36 human, 78 animal, and 11 PBPK). Multiple records of the same data were available for some animal and human studies, so these 125 records corresponded to 115 distinct references (35 human, 69 animal, and 11 PBPK). (A list of multiple records and the study selected as the primary record are provided in the Excel file in the supplemental materials.) The screening process also identified 675 records that were tagged as potentially relevant supplemental information, including mechanistic studies, ADME/

toxicokinetics studies, animal studies with exposure routes other than oral or inhalation (e.g., injection studies), and human case reports.

Heat maps summarizing human and animal studies that met PECO criteria by route of exposure, population type (human) or exposure duration (animal), and health system are provided in Figure 4 and Figure 5, respectively. The available human studies consisted of epidemiology studies in several population types (occupational, general population, pregnant women/infants, children), some of which assessed inhalation exposures but the majority of which had unclear (nonspecific) routes of exposure. The available animal studies included inhalation, oral, dermal, and ocular exposure studies that covered a range of exposure durations. The designs and findings of all studies that met PECO criteria are summarized in an interactive dashboard in Tableau Public, with can be viewed at the following website: https://public.tableau.com/views/NaphthaleneEvidenceMap/ReadMe?:language=en&:display_count=y&:publish=yes&:origin=viz_share_link.

Selection of Human and Animal Studies for Further Evaluation

The human and animal studies that met PECO criteria were refined to a smaller subset that were considered most informative for deriving chronic reference value(s). This refinement of the evidence base is summarized in Figure 6 and described below.

- Because chronic and subchronic studies are preferred for developing chronic reference values, short-term or acute
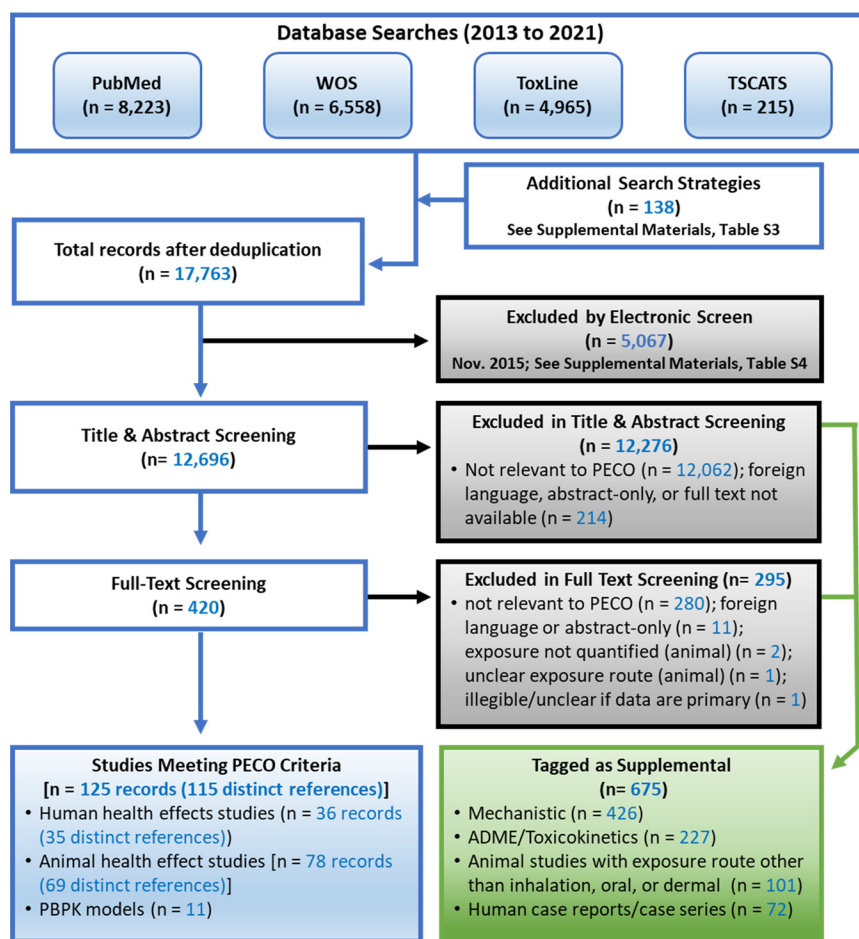
Database Searches (2013 to 2021)

| PubMed (n = 8,223) | WOS (n = 6,558) | ToxLine (n = 4,965) | TSCATS (n = 215) |

Additional Search Strategies (n = 138)
See Supplemental Materials, Table S3

Total records after deduplication (n = 17,763)

Excluded by Electronic Screen (n = 5,067)
Nov. 2015; See Supplemental Materials, Table S4

Title & Abstract Screening (n = 12,696)

Excluded in Title & Abstract Screening (n = 12,276)
• Not relevant to PECO (n = 12,062); foreign language, abstract-only, or full text not available (n = 214)

Full-Text Screening (n = 420)

Excluded in Full Text Screening (n = 295)
• not relevant to PECO (n = 280); foreign language or abstract-only (n = 11); exposure not quantified (animal) (n = 2); unclear exposure route (animal) (n = 1); illegible/unclear if data are primary (n = 1)

Studies Meeting PECO Criteria [n = 125 records (115 distinct references)]
• Human health effects studies (n = 36 records (35 distinct references))
• Animal health effect studies [n = 78 records (69 distinct references)]
• PBPK models (n = 11)

Tagged as Supplemental (n= 675)
• Mechanistic (n = 426)
• ADME/Toxicokinetics (n = 227)
• Animal studies with exposure route other than inhalation, oral, or dermal (n = 101)
• Human case reports/case series (n = 72)

Figure 3. Literature flow diagram for naphthalene. Multiple records were available for some studies, so the box titled "Studies Meeting PECO Criteria" lists both the total number of records and the number of unique studies. The box titled "Tagged as Supplemental" includes mechanistic studies, ADME/toxicokinetic studies, animal studies with exposure routes other than oral or inhalation (e.g., injection studies), and human case reports. Note: ADME, absorption, distribution, metabolism, and excretion; PBPK models, physiologically based pharmacokinetic models; PECO, populations, exposures, comparators, and outcomes; TSCATS, Toxic Substances Control Act Test Submissions; WOS, Web of Science.

Health system	Inhalation — Occupational	Inhalation — General population	Inhalation — Children	Nonspecific — Occupational	Nonspecific — General population	Nonspecific — Pregnant women/infants	Nonspecific — Children	Grand Total
Cardiometabolic					2		2	4
Developmental						3	1	4
Endocrine/Exocrine					2			2
Gastrointestinal	1							1
Hematological				1	1		1	3
Hepatic						2		2
Immunological	1		2				5	8
Nasal	1							1
Neurological	1							1
Pulmonary	1	1						2
Reproductive					6	2		8
Grand Total	4	1	2	1	11	7	9	35

references
1 ▬▬ 6

Figure 4. Survey of human studies that met PECO criteria, organized by route of exposure, population, and health systems evaluated. Numbers represent the number of references that investigated a health system, not the number of references that identified an association with exposure to naphthalene. If a reference evaluated multiple health systems, it is shown here multiple times. Column totals, row totals, and grand total indicate total numbers of distinct references. See the "Human Evidence" tab in the interactive dashboard in Tableau Public for a more detailed description of study design and results (https://public.tableau.com/views/NaphthaleneEvidenceMap/ReadMe?:language=en&:display_count=y&publish=yes&:origin=viz_share_link). Note: PECO, populations, exposures, comparators, and outcomes.

Health system	Inhalation Acute	Short term	Subchronic	Chronic	Oral Acute	Short term	Subchronic	Chronic	Gestational	Dermal Acute	Short term	Subchronic	Ocular Acute	Grand Total
Cardiometabolic						2	2							3
Cardiovascular		1		2			3							6
Dermal										2	2			4
Developmental									5					5
Endocrine/Exocrine			1	2			2				1			6
Gastrointestinal				2			2							4
Hematological				1		1	4				1	1		7
Hepatic			1	2	2	3	7		2		1			17
Immunological			1	2	2	1	3					1		9
Musculoskeletal				2			2							4
Nasal	3	1	1	2										6
Neurological			1	2		1	3							6
Ocular				2		14	23	2			1	1	1	43
Pulmonary	6			3	3	1	4							16
Renal/Urinary			1	2		3	6				1			12
Reproductive			1	2		1	3		5		1	1		13
Whole body	1	1	1	3	1	2	9	1	4		1	1		23
Grand Total	9	1	1	3	5	14	24	2	5	2	3	1	1	69

references
1 — 23

Figure 5. Survey of animal studies that met PECO criteria, organized by route of exposure, duration of exposure, and health systems evaluated. Numbers represent the number of references that investigated a particular health effect, not the number of references that identified an association with exposure to naphthalene. If a reference included multiple experiments (e.g., different species or exposure durations) or evaluated multiple health systems, it is shown here multiple times. Column totals, row totals, and grand total indicate total numbers of distinct references. Acute: exposures ≤24 h. Short-term: exposures >24 h but ≤30 d. Subchronic: exposures >30 d but ≤90 d. Chronic: exposures >90 d. Gestational: any exposure occurring during pregnancy. See "Animal Evidence" tab in the interactive dashboard in Tableau Public for a more detailed description of study design and results (https://public.tableau.com/views/NaphthaleneEvidenceMap/ReadMe?:language=en&:display_count=y&publish=yes&:origin=viz_share_link). Note: PECO, populations, exposures, comparators, and outcomes.

duration animal studies were excluded from further evaluation (*n* = 32). The exception to this is that the 1- and 5-d inhalation studies by Dodd et al. (2010) were included along with the 90-d study by this group (Dodd et al. 2012) because they provide information on the concentration- and time-dependent development of nasal and olfactory necrosis in rats exposed to naphthalene, which is anticipated to be useful for dose–response analysis. Likewise, the 14-d oral study by Shopp et al. (1984) was included along with the 90-d study from the same report to demonstrate dose- and time-dependent responses. Additionally, short-term studies that exposed animals during

gestation were included (NTP 1991, 1992a; Pharmakon Research 1985, 1986; Plasterer et al. 1985).

- A relatively large number of animal studies (*n* = 36) induced cataracts by exposing animals to a single high dose level of naphthalene (∼500 − 1,000 mg/kg-d) and were generally aimed at using naphthalene-induced cataracts as an animal model for age-related cataracts. Such studies are useful for hazard identification but are not likely to be useful for dose–response assessment because only a single high dose level was tested. Because there were several multidose studies that evaluated ocular effects, these cataract studies testing a
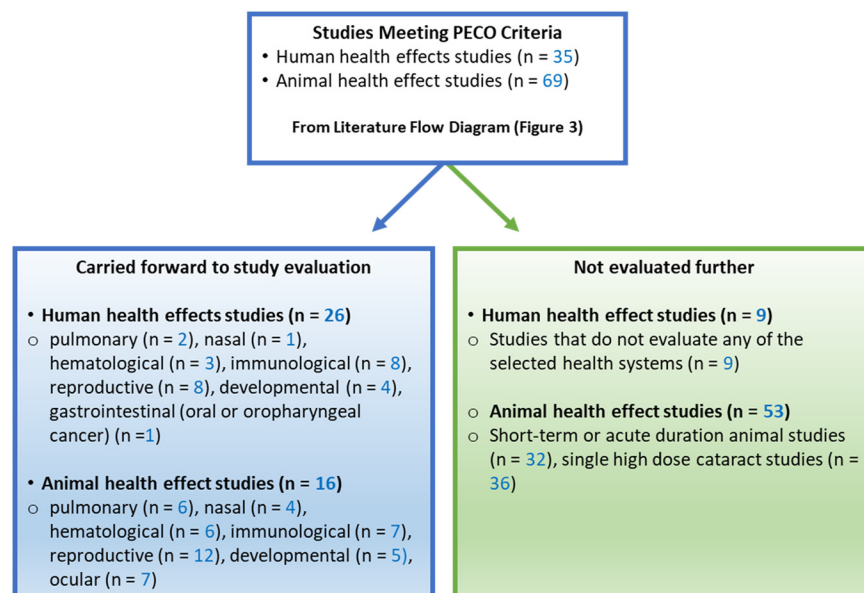
Studies Meeting PECO Criteria
- Human health effects studies (n = 35)
- Animal health effect studies (n = 69)

From Literature Flow Diagram (Figure 3)

Carried forward to study evaluation

- **Human health effects studies (n = 26)**
  o pulmonary (n = 2), nasal (n = 1), hematological (n = 3), immunological (n = 8), reproductive (n = 8), developmental (n = 4), gastrointestinal (oral or oropharyngeal cancer) (n =1)

- **Animal health effect studies (n = 16)**
  o pulmonary (n = 6), nasal (n = 4), hematological (n = 6), immunological (n = 7), reproductive (n = 12), developmental (n = 5), ocular (n = 7)

Not evaluated further

- **Human health effect studies (n = 9)**
  o Studies that do not evaluate any of the selected health systems (n = 9)

  o **Animal health effect studies (n = 53)**
  o Short-term or acute duration animal studies (n = 32), single high dose cataract studies (n = 36)

Figure 6. Selection of human and animal studies for further evaluation. The box titled "Carried forward to study evaluation" lists the total number of human and animal health effects references and the number of studies for each health system that underwent study evaluation; most animal studies evaluated multiple health systems. The box titled "Not evaluated further" lists the total number of human and animal health effects studies that were not carried forward for study evaluation and the reasons for exclusion. Some single high dose cataract studies also had short-term or acute exposure durations, so they are counted in both categories under the reason for exclusion.

Figure 7. Heat map of human study evaluation results, listed by author, year, and HERO identification number. See interactive graphic in HAWC for ratings rationales: https://hawcprd.epa.gov/summary/visual/100500037/. Note: HAWC, U.S. EPA Health Assessment Workspace Collaborative; HERO, U.S. EPA Health & Environment Research Online.

single high dose of naphthalene were not moved forward for study evaluation.

- Based on the summary of studies meeting PECO criteria, the following health systems were selected for further evaluation: respiratory (nasal and pulmonary), hematologic, immune, reproductive, developmental, and ocular. Studies reporting evidence of cancer in any health system were also included. Although an association with severe neonatal jaundice was observed in an epidemiology study (Familusi and Dawodu 1985), this association is thought to be a secondary effect of naphthalene-induced hemolysis, so hepatic effects were not prioritized for further evaluation. Nine epidemiology studies did not evaluate any of the selected health systems; these studies did not undergo a full evaluation, but study designs and suitability for dose–response analysis are summarized in Table S8 of the supplemental materials. These epidemiology studies evaluated neurological, hepatic, endocrine/exocrine (thyroid hormones), and cardiometabolic effects and generally observed no association with naphthalene exposure, although positive associations were reported for obesity, hypertension, metabolic syndrome, type-2 diabetes, and (as mentioned above) neonatal jaundice. As described in Table S8, all of these studies were found to be not suitable or have limited suitability for dose–response analysis.

Study Evaluation Results

Study evaluation results for the subset of epidemiology and animal studies selected for further evaluation are summarized in Figures 7 and 8, respectively. The hyperlinks provided in the figure captions can be used to access interactive versions of these graphics in HAWC, where readers can click to view the detailed rationale for each study evaluation rating. The animal study evaluation heat map indicates instances where different outcomes in a study were rated differently due to outcome-specific concerns.

Human study evaluation summary. For epidemiology studies, the overall confidence was rated as *medium* in 13 studies and as *low* in 11 other studies. Two studies were found to be *critically deficient* and were excluded from further consideration: *a*) The study by Al-Daghri et al. (2014) on the relationship between PAH exposure and childhood asthma was excluded because no correlations or meaningful analyses were performed; and *b*) the occupational exposure study by Merletti et al. (1991) that evaluated oral cavity or oropharyngeal cancer in workers was excluded because there were no cases with naphthalene exposure higher than the general population, and therefore an odds ratio could not be calculated.

Among the epidemiology studies receiving *low* confidence ratings, the main concerns were related to exposure and confounding. Exposure misclassification (evaluated in the "exposure measurement" domain) was found to be a concern in most of the studies (9

Figure 8. Heat map of animal study evaluation results, listed by author, year, and HERO identification number. See interactive graphic in HAWC for ratings rationales: https://hawcprd.epa.gov/summary/visual/100500000/. Note: HAWC, U.S. EPA Health Assessment Workspace Collaborative; HERO, U.S. EPA Health & Environment Research Online.

of the 11) that received *low* confidence ratings. In particular, several epidemiology studies that used the urinary metabolite 1-hydroxynaphthalene (1N) as the only metric for exposure were considered to have a high probability of exposure misclassification because 1N is a metabolite of both naphthalene and the pesticide carbaryl; this concern was amplified in studies conducted in the general population where the source of the exposure was not known. The urinary metabolite 2-hydroxynaphthalene (2N) is specific to naphthalene exposure and therefore was considered an adequate metric for exposure. Another potential concern related to exposure was the measurement of naphthalene metabolites in single spot urine samples, which may not be a reliable surrogate for long-term exposure because naphthalene is rapidly metabolized and has a relatively short half-life in the body. Studies that relied on self-reported exposure history or occupational history without providing a quantitative measure of exposure were also considered a concern for exposure misclassification. Confounding issues (evaluated in the "confounding" domain) were found in 6 of the 11 *low* confidence epidemiology studies and included the lack of inclusion of pertinent information for potentially important confounders, lack of a strategy for identifying confounders, little to no information concerning identification of potential coexposures, and uncertainty about whether confounders were taken into account in the analysis.

Animal study evaluation summary. Almost all the available animal studies were found to be well-conducted, with minimal concerns for bias or sensitivity, and were rated as *high* or *medium* confidence overall for most outcomes. One exception was the 2-y inhalation bioassay in mice by NTP ([1992b](#)), which had a low survival rate among males in the control group in comparison with the naphthalene exposure groups (38% in controls, compared with 75% and 88% survival rates in naphthalene exposure groups). The authors attributed this to wound trauma and secondary lesions from fighting and stated that fighting was likely increased in control males in comparison with exposed males because exposed mice tended to huddle in cage corners during exposure to naphthalene. This aspect was identified as a potential source of bias due to confounding and attrition in the male animals, and therefore the chronic outcomes in this study were rated *low* confidence for males but *high* confidence for females. The study by Katsnelson et al. ([2014](#)) also received a *low* confidence rating primarily due to significant reporting limitations in this study, such as lack of information on the strain of rat used, source and purity of the test chemical, use of a vehicle control, and lack of details on the methods used for the hematologic evaluation. There were also some cases where certain outcomes were rated *low* confidence due to outcome-specific concerns, whereas other outcomes in the same study were rated *high* or *medium* confidence (see evaluations in HAWC for details). However, none of the animal studies were judged to have critical deficiencies that would exclude them from further consideration.

Many of the animal studies did not indicate whether blinding or other steps were taken to reduce observational bias during outcome evaluation ("observational bias/blinding" domain). This was not considered a concern for relatively simple, objective measurements (e.g., body and organ weights) or measurements made using automation or standard laboratory kits (e.g., hematologic evaluations). For the histopathological evaluations, lack of reported blinding was not generally considered a concern because blinding during the initial evaluation of tissues can make it more difficult to separate treatment-related changes from normal variation and may result in subtle lesions being overlooked ([Crissman et al. 2004](#)). However, masked evaluations are appropriate for identifying a no observed adverse effect level (NOAEL) or LOAEL in cases where a chemical is already known to produce a

defined toxic syndrome ([Crissman et al. 2004](#)). Naphthalene inhalation exposure is well known to produce respiratory lesions, so lack of reported blinding was flagged as a potential concern in studies that performed targeted histopathological evaluations of respiratory tissues. Concern for observational bias in histopathological evaluations can be further mitigated by performing peer-review of the diagnoses ([Crissman et al. 2004](#)), which was done by the NTP Pathology Working Group in the 2-y inhalation bioassays in mice and rats ([NTP 1992b](#), [2000](#)).

Summary of the Available PBPK Models for Naphthalene

The literature search identified 11 peer-reviewed publications that describe or make use of pharmacokinetic models for naphthalene. Four of these publications describe cell culture analogs (CCAs) of PBPK models ([Viravaidya et al. 2004](#); [Ghanem and Shuler 2000](#); [Shuler et al. 1996](#); [Sweeney et al. 1995](#)). CCA models are constructed as *in vitro* cell culture systems rather than *in silico* mathematical descriptions of whole organisms; thus, CCA models cannot be efficiently used for risk assessment dosimetry calculations. The six remaining publications all describe whole-organism PBPK models for naphthalene ([Campbell et al. 2014](#); [Morris 2013](#); [Kim et al. 2007](#); [Willems et al. 2001](#); [Quick and Shuler 1999](#); [Sweeney et al. 1996](#)). These models describe the fate of naphthalene once it enters an organism (e.g., a human, mouse, or rat) in terms of ADME and can be used to estimate internal doses (e.g., blood concentrations) experienced by an organism based on well-defined exposure scenarios. In some cases, such models also describe the production and fate of naphthalene metabolites within the body.

The first PBPK models for naphthalene in rats and mice ([Willems et al. 2001](#); [Quick and Shuler 1999](#); [Sweeney et al. 1996](#)) used parallel compartmental structures for liver, lung, and other tissues to describe the disposition of naphthalene and its metabolite naphthalene-1,2-oxide in those tissues. Although these early naphthalene PBPK models can be used to predict tissue dosimetry for lung, liver, blood and various other tissues, they do not include nasal compartments and cannot, therefore, predict dosimetry for nasal tissues.

To address dosimetry of the upper respiratory tract, Campbell et al. ([2014](#)) developed a hybrid computational fluid dynamics (CFD)-PBPK model for inhaled naphthalene in rats and humans. The model assumes that inhaled air flows through the nasal cavity, the nasopharynx and larynx, and then to the lungs. In the model, the nasal cavity consists of two parallel pathways: a dorsal pathway comprising sequential compartments for the respiratory epithelium and one (for humans) or two (for rats) olfactory compartments; and ventral pathway comprising two respiratory epithelium compartments. The authors used time-course data for concentrations of naphthalene in rat blood after single intravenous bolus doses ([Quick and Shuler 1999](#)), 6-h inhalation exposures ([NTP 2000](#)), and rat naphthalene upper respiratory tract extraction data at fixed inspiratory flow rates ([Morris and Buckpitt 2009](#)) to evaluate the accuracy of rat model predictions, but similar naphthalene kinetic data for humans were not available to evaluate human model predictions. Campbell et al. ([2014](#)) used their rat and human models to predict continuous exposure human equivalent concentrations; e.g., they estimated inhalation exposure concentrations that would produce in humans the same internal dose metrics as those predicted for rats exposed at the NOAEL reported by Dodd et al. ([2012](#)) for a 90-d exposure in rats.

Kim et al. ([2007](#)) developed a human PBPK model for naphthalene that can be used to simulate both dermal and inhalation exposure scenarios. Their model has five compartments: two skin compartments representing the exposed stratum corneum and the viable epidermis immediately below it; one central blood compartment (to which inhalation exposures are delivered directly);

one fat compartment; and one compartment representing all other tissues. Most physiological parameters, partition coefficients, and metabolism parameters for the model were extracted from the literature, but parameters related to dermal uptake and permeability and some partition coefficients were fit to blood time-course data from a laboratory study of dermal exposure to jet propellant 8 (JP-8) fuel, which is a jet fuel that contains naphthalene, in humans. Kim et al. (2007) demonstrated that their model was able to reasonably reproduce exhaled air concentration data from a field study of JP-8 dermal and inhalation exposures in U.S. Air Force personnel (Chao et al. 2006; Egeghy et al. 2003). The model's authors reported that a wide range of skin permeability parameters were necessary to fit individual human data, and this may indicate large interindividual variability in dermal uptake and/or systemic clearance of naphthalene.

Most recently, researchers in our group (Kapraun et al. 2020) published a model that extends the model of Campbell et al. (2014) by incorporating a skin route of exposure. Kapraun et al. (2020) evaluated their model by showing that it could reproduce time profiles of blood concentrations following controlled skin exposures in human subjects (Kim et al. 2006) and thus demonstrated the suitability of this model for human health risk assessment applications.

To the best of our knowledge, no existing naphthalene PBPK model is suitable for estimating human oral or inhalation exposures from metabolite concentrations in urine. Epidemiology studies that used urinary biomarkers as the only metric of exposure were therefore considered to have limited suitability for dose–response analysis, because a meaningful analysis is not possible using the currently available models.

Summary of Available Human and Animal Studies and Dose–Response Considerations for Each Health System

The following sections summarize the available evidence base for each of the health systems that were prioritized for further evaluation. This includes the available study designs, outcome(s) evaluated, outcome(s) observed, study evaluation results, and suitability of the study for dose–response analysis. Exposure measurements shown in Tables 2–7 for animal studies are nominal doses. The outcomes listed in the "Outcome(s) observed" column in Tables 2–7 reflect a high-level summary of the statistically and/or biologically significant effects reported by the authors of each study; when authors indicated there was no effect of treatment, the tables report no effects observed. Within each table, studies are organized into human (inhalation or nonspecific routes of exposure) and animal (inhalation, oral, or dermal exposure) and listed alphabetically by overall confidence level (*high* confidence studies listed first, followed by *medium* and *low* confidence). Body weight measurements and clinical observations from animal studies (categorized as "Whole Body" in the Tableau Public figures) are not explicitly discussed in these tables aside from maternal and offspring body weight changes in gestational exposure studies, but they are useful to consider during hazard identification as evidence of systemic toxicity to help interpret findings for other outcomes.

The only studies that evaluated naphthalene-induced carcinogenesis were animal studies that reported nasal and pulmonary lesions following inhalation exposure, so all evidence for cancer is summarized in the "Respiratory effects" section.

Respiratory effects (pulmonary and nasal). The respiratory tract has been demonstrated to be a sensitive target in rodents following inhalation exposure to naphthalene (ATSDR 2005). Table 1 summarizes the evidence base for respiratory effects of naphthalene exposure, consisting of studies that reported nasal or pulmonary outcomes. The evidence base for human respiratory effects consisted of a *medium* confidence study that evaluated

inflammatory effects and sensory irritation in workers exposed to naphthalene in the abrasives industry (Sucker et al. 2021) and a *medium confidence* study that evaluated changes in lung function associated with residential levels of volatile organic compounds (VOCs) (Cakmak et al. 2014). Sucker et al. (2021) used a cross-sectional cross-week design with pre- and post-shift outcome measurements and is suitable for dose–response analysis based on the study design and exposure measurements; however, the applicability for reference value derivation is somewhat limited because the outcome measurements are not generally considered to be apical health effects (self-reported nose and eye irritation; redness and mucus measured via otorhinolaryngological examination; and inflammatory biomarkers in serum, nasal lavage, or sputum). The study by Cakmak et al. (2014) used a cross-sectional study design that has limited ability to assess temporality between exposure and effect, and therefore this study would be more useful for hazard identification than for the derivation of a reference value.

Among the available animal studies, the most informative appear to be five inhalation studies in mice or rats evaluating nasal or pulmonary histopathological lesions. These include the NTP 2-y bioassays in mice and rats (NTP 1992b, 2000) that currently serve as the scientific basis for the NTP, IARC, and IRIS cancer determinations. There is also an earlier study that evaluated neoplastic lesions in female mice following a 6-month exposure (Adkins et al. 1986) and two more recent studies that reported nasal lesions in rats following shorter inhalation exposure durations [90-d: Dodd et al. (2012); 1- or 5-d: Dodd et al. (2010)] at lower concentration ranges. With the exception of the data in male mice from the 2-y inhalation bioassay by NTP (1992b), which was considered *low* confidence due to the high mortality rate in control animals, all of these studies were found to be *high* or *medium* confidence and were considered suitable for dose–response analysis. No respiratory effects were observed in three *high* confidence oral exposure studies (Shopp et al. 1984; Battelle 1980a, 1980b) aside from an increase in lung weight (Shopp et al. 1984). Although changes in organ weight may be indicative of adverse effects, interpretation of organ weight changes is limited in the absence of additional histopathological or functional data demonstrating respiratory effects in these animals, so this outcome is less likely to be useful for risk assessment in comparison with the histopathological outcomes observed in inhalation studies.

Hematologic effects. The evidence base for hematologic effects of naphthalene is summarized in Table 3. Although effects on leukocytes are included in this table as part of the hematology evaluation in several studies, these data should be considered more appropriate for the evaluation of immunological effects (discussed in the "Immune system effects" section).

In humans, the major hematologic outcome of concern for naphthalene exposure is hemolytic anemia. Hemolytic anemia has been frequently reported as a manifestation of naphthalene exposure in human case studies, particularly among children who have ingested mothballs and in infants whose clothing or bedding was stored in mothballs (ATSDR 2005). The three available epidemiology studies observed associations between naphthalene exposure and hemolytic anemia (Santucci and Shah 2000) or other hematologic outcomes (Kamal et al. 2014; Sudakin et al. 2013). However, this evidence base is limited because two out of three studies were rated as *low* confidence (Kamal et al. 2014; Santucci and Shah 2000), and all three studies were considered to have limited utility for dose–response analysis due to lack of quantitative exposure data (Santucci and Shah 2000), the use of urinary metabolites as the metric for exposure (Kamal et al. 2014; Sudakin et al. 2013), and/or the use of cross-sectional designs that have limited ability to assess temporality (Sudakin

Table 2. Summary of available studies evaluating respiratory effects.

Author and year of publication	Study description	Route of exposure	Exposure measurement	Outcome(s) evaluated	Outcome(s) observed	Overall confidence level for outcome	Applicability for dose–response
Human studies (inhalation)							
Sucker et al. (2021)	Occupational cross-sectional cross-week study of healthy and nonsmoking male employees at an abrasives plant with either moderate ($n = 22$) or high ($n = 17$) exposure to naphthalene compared to 22 male employees from the same plants with no or only rare exposure to naphthalene ($n = 22$), in Germany or Austria.	Inhalation	Naphthalene levels measured during Thursday shift via personal air monitoring, and naphthalene metabolite levels (1N or 2N) in postshift urine	Preshift (Monday) and postshift (Thursday) measurements of self-reported nose and eye irritation, otorhinolaryngological examination, and evaluation of inflammatory markers in serum, NALF, and sputum.	Increased eye and nasal complaints and increased endoscopic score (based on clinical findings of reddening/swelling nasal mucosa and mucus production) in exposed groups compared to reference group. Decreased serum club cell secretory protein 16 (CC16) after shift Thursday compared to preshift Monday (no statistically significant differences between exposed groups and controls).	Medium. The small sample sizes and the cross-sectional study design may limit the study's ability to detect an effect. The "healthy worker effect" raises the potential of bias. However, these concerns are expected to have minimal impact on the interpretation of the study.	Suitable. Cross-week study design including air measurement from a full shift, with multiple exposure groups.
Cakmak et al. (2014)	General population cross-sectional health survey of ages 3–79 y old (total of 3,039 nonsmokers; smokers excluded from study) in Canada.	Inhalation	Residential indoor air levels of 84 VOCs, including naphthalene	Lung function as measured by FEV in the first second (FEV1), FVC, and the FEV_1/FVC ratio	Statistically significant negative association with FEV1, FVC and the FEV_1/FVC ratio	Medium. All domains were evaluated as at least adequate.	Limited suitability. Cross-sectional study design with limited ability to assess temporality.
Animal studies (inhalation)							
Dodd et al. (2012)	Male and female rats (Fischer 344); 90-d exposure	Inhalation (whole body)	0, 0.5, 5.2, 52,157 mg/m^3	Nasal histopathology	Increased nasal epithelium lesions in both sexes	High. This study was well-designed to evaluate these outcomes. The only notable concern is that the authors did not describe methods for reducing observational bias in the targeted histopathological evaluation.	Suitable. Multidose study with quantitative data.
Dodd et al. (2010)	Male and female rats (Sprague-Dawley, Fisher 344); 1- or 5-d exposure	Inhalation (whole body)	1-d exposure: 0, 0.5, 1.57, 5.2, 52, 157 mg/m^3 5-d exposure: 0, 0.5, 5.2, 52 mg/m^3	Nasal histopathology	1-d exposure: Increased nasal olfactory epithelium necrosis and respiratory epithelium necrosis in both sexes and strains 5-d exposure: Increased degeneration of nasal olfactory epithelium and nasopharyngeal goblet cell hyperplasia/hypertrophy in both sexes and strains	High. This study was well-designed to evaluate these outcomes. The only notable concern is that the authors did not describe methods for reducing observational bias in the targeted histopathological evaluation.	Suitable. Multidose study with quantitative data (quantitative data provided for males only).
NTP (2000)	Male and female rats (Fisher 344); 2-y (105-wk) exposure	Inhalation (whole body)	0, 52, 157, 314mg/m^3	Nasal and pulmonary histopathology	Increased incidence of respiratory epithelial adenomas (statistically significant in males), olfactory epithelial neuroblastomas (positive trend in both sexes, statistically significant in females), and nonneoplastic lesions (statistically significant in both sexes)	High. This study was well designed to evaluate these outcomes. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.
NTP (1992b)	Male and female mice (B6C3F1); 2-y (103-wk) exposure	Inhalation (whole body)	0, 52, 157 mg/m^3	Nasal and pulmonary histopathology	Increased incidence of alveolar/bronchiolar adenomas and combined alveolar/bronchiolar adenomas and carcinomas (statistically significant in females; marginal increase in males, but within the historical control rate), and increased incidence of nonneoplastic lesions (statistically significant in both sexes)	High (females). This study was well designed to evaluate these outcomes. Evidence was presented clearly and transparently. Low (males). The high mortality rate in control males has the potential to interfere with the interpretation of results.	Suitable. Multidose study with quantitative data.

Table 2. (*Continued.*)

Author and year of publication	Study description	Route of exposure	Exposure measurement	Outcome(s) evaluated	Outcome(s) observed	Overall confidence level for outcome	Applicability for dose–response
Adkins et al. (1986)	Female mice (A/J); 6-month exposure	Inhalation (whole body)	0, 52, 157 mg/m³	Pulmonary histopathology	Statistically significant increase in alveolar adenomas	Medium. Some concerns were raised about limited procedural details on the histopathology evaluation, and no description of methods to reduce observational bias.	Suitable. Multidose study with quantitative data.
Animal studies (oral)							
Battelle (1980b)	Male and female rats (Fischer 344); 90-d exposure	Oral gavage	0, 25, 50, 100, 200, 400 mg/kg-d	Pulmonary histopathology	No effects observed	High. This study was well designed to evaluate this outcome. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.
Battelle (1980a)	Male and female mice (B6C3F1); 90-d exposure	Oral gavage	0, 12.5, 25, 50, 100, 200 mg/kg-d	Pulmonary histopathology	No effects observed	High. This study was well designed to evaluate this outcome. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.
Shopp et al. (1984)	Male and female mice (CD-1); 14- or 90-d exposure	Oral gavage	14-d exposure: 0, 27, 53, 267 mg/kg-d; 90-d exposure: 0, 5.3, 53, 133 mg/kg-d	Lung weight	14-d exposure: Statistically significant increase in absolute and relative lung weight in females; no effects observed in males; 90-d exposure: No effects observed	High. This study was well-designed to evaluate this outcome. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.

Note: FVC, forced vital capacity; FEV1, forced expiratory volume; NALF, nasal lavage fluid; VOC, volatile organic compound.

et al. 2013; Santucci and Shah 2000). In addition to these studies that directly evaluated hematological effects, there were two epidemiology studies that evaluated neonatal jaundice (Sodeinde et al. 1995; Familusi and Dawodu 1985) (see Table S8, "Hepatic"), which may be a secondary effect of hemolytic anemia and therefore may provide additional information for hazard identification; we did not do a full evaluation of these studies, but we noted that they are not suitable for dose–response analysis.

Rats and mice have been described as having lower sensitivity to hemolytic agents in comparison with humans, which is speculated to be due to greater activity of methemoglobin reductase in rodents compared to humans (Abdo et al. 1992). Of the studies that observed hematologic effects, the most informative appeared to be the *medium* confidence studies in rats by Battelle (1980b) and in mice by NTP (1992b), both of which were found to be suitable for dose–response analysis. The study by Shopp et al. (1984) was rated as *medium confidence* but is not suitable for dose–response analysis because only qualitative data was reported, and the study by Katsnelson et al. (2014) was rated as *low* confidence and was found to have limited suitability for dose–response analysis because only a single dose level was tested. The majority of effects across animal studies were related to leukocytes and therefore are more applicable to the evaluation of immunological effects ("Immune system effects" section), although decreased hemoglobin and hematocrit were observed in the rat study by Battelle (1980b), and decreased hematocrit and mean corpuscular volume were observed in the mouse study by NTP (1992b).

Immune system effects. Table 4 summarizes the evidence base for immunological effects of naphthalene exposure. Seven epidemiology studies reported on the relationship between naphthalene exposure and the diagnosis of asthma or other allergic diseases (Lin et al. 2018; Al-Daghri 2008; Kim et al. 2005), biochemical markers of allergic sensitization such as serum IgE and T-cell cytokine profiles (Lin et al. 2018; Al-Daghri et al. 2013; Lehmann et al. 2001, 2002), and immune cell counts in peripheral blood (Rhodes et al. 2003), of which six of the seven studies reported a statistically significant relationship between naphthalene exposure and one or more of these measurements (Lin et al. 2018; Al-Daghri et al. 2013; Al-Daghri 2008; Kim et al. 2005; Rhodes et al. 2003; Lehmann et al. 2001, 2002); however, the majority of these studies were considered *low* confidence, and all were found to have limited suitability for dose–response analysis. The limiting factors for dose–response analysis varied across studies and included concerns about the nature of the exposure measurements (Lehmann et al. 2001, 2002), concerns about lack of adjustment for confounders (Al-Daghri et al. 2013), the use of urinary metabolites as the only metric for exposure (Lin et al. 2018; Kim et al. 2005), and the use of cross-sectional designs with limited ability to assess temporality (Lin et al. 2018; Al-Daghri et al. 2013; Al-Daghri 2008; Rhodes et al. 2003). As noted above, the epidemiology study by Kamal et al. (2014) that reported increased leukocyte counts in exposed vs. unexposed workers (Table 3) was also rated *low confidence* and was found to have limited suitability for dose–response analysis due to the use of urinary biomarkers as the metric for exposure.

The animal evidence base for immunological effects consisted of seven *high* or *medium* confidence studies that evaluated immune system histopathology and organ weights, although NTP (1992b) was considered *low* confidence for effects in males due to high mortality among control animals. Four of these studies observed no effects (NTP 1992b, 2000; Bushy Run 1986; Battelle 1980a), and the other three reported effects consisting of decreased thymus and spleen weights in mice (Shopp et al. 1984) and rats (Dodd et al. 2012) and a moderate depletion of thymic

Table 3. Summary of available studies evaluating hematologic effects.

Author and publication year	Study description	Route of exposure	Exposure measurement	Outcome(s) evaluated	Outcomes(s) observed	Overall confidence level for outcome	Applicability for dose-response
Human studies (nonspecific route of exposure)							
Sudakin et al. (2013)	Cross-sectional analysis of National Health and Nutrition Examination Survey (NHANES) 2003–2004 (a sample of noninstitutionalized U.S. civilians, selected by a complex, multistage probability design) data on 2,450 adults (18+ y old) without treated anemia and with complete data on naphthalene, hemoglobin, and hematocrit	Nonspecific route of exposure	Urinary biomarkers of naphthalene exposure (1N or 2N)	Hemoglobin and hematocrit levels	Statistically significant positive association of hemoglobin and hematocrit levels with 1N; no association with 2N	Medium. All domains were evaluated as at least adequate.	Limited suitability. Cross-sectional study design with limited ability to assess temporality. Also, insufficient availability of data or models to relate urinary metabolites to exposure levels.
Kamal et al. (2014)	Occupational case–control study of 46 male brick kiln workers compared with 34 nonoccupationally exposed workers in Pakistan	Nonspecific route of exposure	Naphthalene metabolite levels (1N or 2N) in postshift urine	Hemoglobin, mean corpuscle volume, and platelet, erythrocyte, and total leukocyte counts	Statistically significantly lower hemoglobin, erythrocytes, and mean corpuscular volume and increased leukocytes in exposed vs. unexposed workers	Low. Domains for exposure measures, confounding, analysis, and sensitivity were evaluated as deficient.	Limited suitability. Insufficient availability of data or models to relate urinary metabolites to exposure levels.
Santucci and Shah (2000)	General population cross-sectional health survey of 24 children (age 2 wk to 18 y) hospitalized with glucose 6-phosphate dehydrogenase deficiency and acute hemolysis in the United States (New York).	Nonspecific route of exposure	Retrospective chart review; no exposure level data	Hemolysis	Hemolytic anemia associated with self-reported exposure to naphthalene-containing moth repellents	Low. Domains for exposure measures, confounding, and analysis were evaluated as deficient.	Not suitable. No quantitative exposure level data.
Animal studies (inhalation)							
NTP (1992b)	Male and female mice (B6C3F1); 2-year (103-wk exposure), hematology evaluated at 14 d	Inhalation (whole body)	0, 52, 157 mg/m^3	Hematocrit, hemoglobin, mean cell (corpuscular) volume, and erythrocyte, reticulocyte, and total leukocyte counts	Statistically significant decrease in hematocrit and mean corpuscular volume and increase in leukocyte count in females	Medium. This study appears to be well-designed to evaluate these outcomes, but lacked details on the methods for the hematological evaluation. The high mortality rate in the control males is not a major concern for the hematological evaluation because it was performed on day 14, which was before there was a large difference in survival between experimental groups.	Suitable. Multidose study with quantitative data.
Animal studies (oral)							
Battelle (1980b)	Male and female rats (Fischer 344); 90-d exposure	Oral gavage	0, 25, 50, 100, 200, 400 mg/kg-d	Hemoglobin, hematocrit, mean corpuscular volume, erythrocyte count, and total and differential leukocyte count	Decreased hemoglobin and hematocrit and increased mature neutrophils in both sexes; decreased lymphocytes in males	Medium. This study appears to be well-designed to evaluate these outcomes, but lacked details on the methods for the hematological evaluation.	Suitable. Multidose study with quantitative data.
Battelle (1980a)	Male and female mice (B6C3F1); 90-d exposure	Oral gavage	0, 12.5, 25, 50, 100, 200 mg/kg-d	Hemoglobin, hematocrit, mean corpuscular volume, erythrocyte count, and total and differential leukocyte count	No effects observed	Medium. This study appears to be well-designed to evaluate these outcomes, but lacked details on the methods for the hematological evaluation.	Suitable. Multidose study with quantitative data.
Shopp et al. (1984)	Male and female mice (CD-1); 14- or 90-d exposure	Oral gavage	14-d exposure: 0, 27, 53, 267 mg/kg-d; 90-d exposure: 0, 5.3, 53, 133 mg/kg-d	Hemoglobin, erythrocyte, platelet, and differential leukocyte counts, plasma extrinsic activity (prothrombin time) and	14-d exposure: Statistically significant increase in eosinophils in males; decreased prothrombin time in females	Medium. This study was well-designed to evaluate these outcomes, but confidence was decreased because only qualitative results are provided.	Not suitable. Quantitative data not reported.

Table 3. (Continued.)

Author and publication year	Study description	Route of exposure	Exposure measurement	Outcome(s) evaluated	Outcome(s) observed	Overall confidence level for outcome	Applicability for dose–response
				intrinsic activity (activated partial thromboplastin time)	90-d exposure: Statistically significant increase in hemoglobin, activated partial thromboplastin time, and eosinophils in females		
Katsnelson et al. (2014)	Female rats (strain unspecified); 7-wk exposure	Oral gavage	0, 87.5 mg/kg-d	Hemoglobin, and erythrocyte, reticulocyte, lymphocyte, neutrophil, and eosinophil counts	Statistically significant decrease in segmented neutrophils	Low. Concerns were raised because no information was provided on the source and purity of the test chemical or the use of a vehicle control, animal strains were not reported, and no experimental details are provided on the hematological evaluation.	Limited suitability. Single dose study.
Animal studies (dermal)							
Bushy Run (1986)	Male and female rats (Sprague-Dawley CD); 90-d exposure	Dermal	0, 100, 300, 1,000 mg/kg-d	Hemoglobin, hematocrit, mean corpuscular volume, mean corpuscular hemoglobin, mean corpuscular hemoglobin concentration, erythrocyte count, total and differential leukocyte count, platelet count	No effects observed	High. This study was well designed to evaluate these outcomes. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.

lymphocytes in female rats (Battelle 1980b). These studies were found to be suitable for dose–response analysis because they tested multiple doses and reported quantitative data, although the biological significance of the organ weight and histopathological changes is unclear in the absence of functional measurements of immune toxicity. Only one animal study evaluated functional effects on the immune system (Shopp et al. 1984), with no dose-related effects reported. This limits the applicability of these outcomes for hazard identification and may make them a less desirable choice for reference value derivation. As noted above, the study by Battelle (1980b) also reported an increase in mature neutrophils in rats and was found to be suitable for dose–response analysis, whereas two studies that reported other effects on leukocytes (Katsnelson et al. 2014; Shopp et al. 1984) were found to have limited suitability for dose–response analysis (Table 3).

Reproductive effects. Table 5 summarizes the evidence base for reproductive effects of naphthalene exposure in males and females. Two *medium* confidence epidemiology studies evaluated the association between naphthalene exposure and outcomes in pregnant women, but both had limited suitability for dose–response analysis. The study by Singh et al. (2008) reported higher placental naphthalene levels among women with preterm vs. full-term birth but was considered to have limited suitability for dose–response analysis due to concerns that there was no adjustment for confounders in this analysis. The study by Yin et al. (2017) reported a significant negative association with cord anti-Müllerian hormones (AMH) in umbilical cord serum, but concerns were raised that this measurement may not be a reliable stand-in for fetal hormones; the fetal hypothalamic–pituitary–gonadal axis is suppressed by placental hormones during late gestation and is reactivated at birth (Kuiri-Hänninen et al. 2014); therefore, hormone levels including AMH (which is affected by gonadotropins) measured at birth may not reflect hormone levels during gestation.

In the six studies that evaluated effects in adult men, inverse associations were observed between urinary naphthalene metabolites and semen quality (Yang et al. 2017; Meeker et al. 2004a, 2004b), serum testosterone levels (Meeker et al. 2006), and serum estradiol levels (Meeker et al. 2008), with no association observed between urinary naphthalene metabolites and below-reference semen parameters in the study by Xia et al. (2009). None of these studies in men would be considered likely candidates for reference value derivation due to *low* confidence ratings and/or limited suitability for dose–response analysis. All six studies used urinary metabolites as the metric for exposure, and three of the studies by Meeker et al. were considered to have a significant risk of exposure misclassification due to the use of urinary 1N as the only exposure measurement (Meeker et al. 2006, 2004a, 2004b). Of note, the author of these studies later published a reanalysis of the data in these three studies in which subjects were stratified by the 1N/2N ratio (Meeker et al. 2007) (not shown in Table 5). The reanalysis found that carbaryl exposures were likely responsible for the associations between 1N and sperm motility (Meeker et al. 2004a) and serum testosterone (Meeker et al. 2006), whereas naphthalene is likely responsible for the association between 1N and sperm DNA damage (Meeker et al. 2004b). Most of these studies were conducted using cross-sectional designs, which have limited ability to assess temporality.

The animal evidence base consisted of 12 studies that evaluated male or female reproductive organ weights, histopathology of reproductive organs, number of dams pregnant or delivered at the time of sacrifice in gestational exposure studies, and/or maternal body weight gain, with most considered *high* confidence. However, confidence in the maternal weight gain measurement in the study by Plasterer et al. (1985) was reduced to *medium*

Table 4. Summary of available studies evaluating immune effects.

Author and publication year	Study description	Route of exposure	Exposure measurement	Outcome(s) evaluated	Outcome(s) observed	Overall confidence level for outcome	Applicability for dose–response
Human studies (inhalation)							
Rhodes et al. (2003)	Occupational cross-sectional study of 151 subjects from 3 U.S. Air Force bases, age 18–44 y in United States (southeastern states)	Inhalation	Naphthalene levels in breathing zone air during 4-h work assignments.	Immune cell counts in peripheral blood (total and differential leukocyte counts, total lymphocytes, T-cells, T-helper cells, T-suppressor cells, natural killer cells, B cells)	Statistically significantly higher counts of total leukocytes, neutrophils, and monocytes in high versus low/no exposure groups	Medium. All domains were evaluated as at least adequate.	Limited suitability. Cross-sectional study with limited ability to assess temporality.
Lehmann et al. (2002)	General population cohort study of 85 randomly selected newborns (43 boys and 42 girls) in Leipzig, Germany	Inhalation	Naphthalene levels in air samples collected during a 4-wk period in bedrooms of 3-y-olds and used as a proxy for maternal exposure	Cytokine secretion profiles of cord-blood T cells (IL-4, IL-2, interferon-γ, and tumor necrosis factor α)	Statistically significant positive association with elevated percentages of IL-4–producing type 2 T cells	Low. The exposure measures and confounding domains were evaluated as deficient.	Limited suitability. Postnatal exposure measurements may not be a viable proxy for prenatal/maternal exposure.
Lehmann et al. (2001)	General population cohort study of 200 3-y-old children in Leipzig, Germany	Inhalation	Naphthalene levels in air samples collected during a 4-wk period in bedrooms of 3-y-olds	Sensitization to allergens, measured as total serum IgE and specific IgE antibodies to food (milk, egg), indoor (cat, *D. pteronyssinus, D. farina, Penicillium notatum, Aspergillus fumigatus,* and *Candida albicans*), and outdoor (birch, timothy) allergens. Cytokine-producing peripheral T cells were measured in a subgroup of $n = 28$ children.	Higher prevalence of allergic sensitization in VOC-exposed groups (including naphthalene); statistically significantly relationship with decreased CD8+ peripheral T cells	Low. The confounding domain was evaluated as deficient and considered to have a severe impact on the interpretation of results.	Limited suitability. Exposure information was only available for 60% of subjects. High association between naphthalene and other VOCs is a potential confounder.
Human studies (nonspecific route of exposure)							
Lin et al. (2018)	General population cross-sectional health survey of 453 kindergarten children from the Childhood Environment and Allergic Diseases Study cohort in Taiwan	Nonspecific route of exposure	Urinary metabolite levels (2N)	Serum IgE levels and allergic diseases (atopic dermatitis, allergic rhinitis, and asthma)	Statistically significant association with asthma	Medium. All domains except participant selection were evaluated as at least adequate.	Limited suitability. Cross-sectional study design with limited ability to assess temporality. Also, insufficient availability of data or models to relate urinary metabolites to exposure levels.
Al-Daghri et al. (2013)	General population cross-sectional study of children ≤15 y old (98 asthma patients, 97 healthy controls) in Saudi Arabia.	Nonspecific route of exposure	Naphthalene levels in serum	Asthma-related biomarkers (e.g., resistin, cytokines, serum immunoglobins)	Statistically significant correlation of asthma biomarkers with serum naphthalene levels	Low. Domains on participant selection, exposure measures, confounding, and analysis were evaluated as deficient.	Limited suitability. Cross-sectional study with limited ability to assess temporality. Study does not adjust for possible confounders or interaction with other measured PAHs.
Al-Daghri (2008)	General population cross-sectional study of children ≤5 years of age (61 asthmatic, 14 nonasthmatic cases) in Saudi Arabia.	Nonspecific route of exposure	Naphthalene levels in serum	Diagnosis of asthma	Statistically significantly higher mean serum naphthalene levels in asthma cases versus nonasthmatic cases	Low. Domains on participant selection, confounding, and analysis were evaluated as deficient.	Limited suitability. Cross-sectional study with limited ability to assess temporality.

Table 4. (*Continued.*)

Author and publication year	Study description	Route of exposure	Exposure measurement	Outcome(s) evaluated	Outcomes(s) observed	Overall confidence level for outcome	Applicability for dose–response
Kim et al. (2005)	General population case–control study of children identified at Pediatric Allergy Clinic; 30 asthma cases (20 male, 10 female, mean age 4.9 y) and 30 controls with no history of wheezing episodes or allergic disease (13 male, 17 female, mean age 5.5 y) in Korea	Nonspecific route of exposure	Urinary metabolite levels (2N)	Diagnosis of asthma	No significant difference in urinary naphthalene metabolite levels in control vs. asthmatic groups	Low. Domains on participant selection, exposure measures, analysis, and sensitivity were evaluated as deficient.	Limited suitability. Insufficient availability of data or models to relate urinary metabolites to exposure levels.
Animal studies (inhalation)							
Dodd et al. (2012)	Male and female rats (Fischer 344); 90-d exposure	Inhalation (whole body)	0, 0.5, 5.2, 52, 157 mg/m³	Thymus and spleen weights	Statistically significant decrease in thymus weight in both sexes and spleen weight in males	High. This study was well-designed to evaluate these outcomes. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.
NTP (2000)	Male and female rats (Fischer 344); 2-y (105-wk) exposure	Inhalation (whole body)	0, 52, 157, 314 mg/m³	Histopathology of lymph nodes (mandibular, mesenteric, bronchial, and mediastinal), thymus, and spleen	No effects observed	High. This study was well designed to evaluate these outcomes. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.
NTP (1992b)	Male and female mice (B6C3F1); 2-y (103-wk) exposure	Inhalation (whole body)	0, 52, 157 mg/m³	Histopathology of lymph nodes (mandibular, mesenteric, bronchial, and mediastinal), thymus, and spleen	No effects observed	High (females). This study was well designed to evaluate these outcomes. Evidence was presented clearly and transparently. Low (males). The high mortality rate in control males has the potential to interfere with the interpretation of results.	Suitable. Multidose study with quantitative data.
Animal studies (oral)							
Battelle (1980b)	Male and female rats (Fischer 344); 90-d exposure	Oral gavage	0, 25, 50, 100, 200, 400 mg/kg-d	Histopathology of thymus and spleen	Moderate depletion of thymic lymphocytes in females	High. This study was well designed to evaluate these outcomes. Evidence was presented clearly and transparently.	Suitable. Multidose studies with quantitative data.
Battelle (1980a)	Male and female mice (B6C3F1); 90-d exposure	Oral gavage	0, 12.5, 25, 50, 100, 200 mg/kg-d	Histopathology of thymus and spleen	No effects observed	High. This study was well designed to evaluate these outcomes. Evidence was presented clearly and transparently.	Suitable. Multidose studies with quantitative data.
Shopp et al. (1984)	Male and female mice (CD-1); 14- or 90-d exposure	Oral gavage	14-d exposure: 0, 27, 53, 267 mg/kg-d 90-d exposure: 0, 5.3, 53, 133 mg/kg-d	Immunotoxicity assays including antibody forming cell counts, lymphocyte responsiveness, analysis of delayed-type hypersensitivity and popliteal lymph node response, and bone marrow function; thymus and spleen weights	14-d exposure: Statistically significant decrease in absolute thymus weight in males and absolute and relative spleen weight in females 90-d exposure: Statistically significant decrease in absolute and relative spleen weight in females	High (organ weights). This study was well-designed to evaluate these outcomes. Evidence was presented clearly and transparently. Medium (immunotoxicity assays). This study was well-designed to evaluate these outcomes, but confidence was decreased because only qualitative results are provided.	Suitable (organ weights). Multidose study with quantitative data. Not suitable (immunotoxicity assays). Data not shown for 14-d study; qualitative data is presented for 90-d study.
Animal studies (dermal)							
Bushy Run (1986)	Male and female rats (Sprague-Dawley CD); 90-d exposure	Dermal	0, 100, 300, 1,000 mg/kg-d	Histopathology of thymic region, spleen, and lymph nodes	No effects observed	High. This study was well designed to evaluate these outcomes. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.

Note: IL, interleukin.

Table 5. Summary of studies evaluating reproductive effects.

Author and publication year	Study description	Route of exposure	Exposure measurement	Outcome(s) evaluated[a]	Outcome(s) observed	Overall confidence level for outcome	Applicability for dose–response
Human studies (nonspecific route of exposure)							
Singh et al. (2008)	General population case–control study of 60 mothers [31 with full-term deliveries (mean age 25.55 y) and 29 with preterm deliveries (mean age 26.41 y)] in India (Lucknow City).	Nonspecific route of exposure	Placental tissue levels of naphthalene	Preterm birth	Higher placental naphthalene levels among women with preterm vs. full-term birth, but the difference was not statistically significant.	Medium. All domains were evaluated as adequate.	Limited suitability. Study does not adjust for possible confounders. (Information on confounders was collected, but not used directly in the analysis.)
Xia et al. (2009)	General population cross-sectional health survey of 542 men diagnosed with unexplained male factor infertility between 2004 and 2007 (176 controls) in China.	Nonspecific route of exposure	Creatinine-adjusted urinary metabolite levels (1N, 2N)	Sperm parameters (semen volume, sperm concentration, sperm number per ejaculum, and sperm motility)	No association between urinary naphthalene metabolites and below-reference semen parameters	Medium. All domains except sensitivity were evaluated as at least adequate.	Limited suitability. Cross-sectional study design with limited ability to assess temporality. Also, insufficient availability of data or models to relate urinary metabolites to exposure levels.
Yang et al. (2017)	General population cohort study of 933 men from subfertile couples; recruited from an infertility clinic in Wuhan, China in China. Men had no known occupational exposure to PAHs.	Nonspecific route of exposure	Urinary metabolite levels (1N, 2N)	Semen quality (sperm count, sperm concentration, sperm morphology, sperm motility, semen volume)	Significant inverse association between at least one metabolite level and sperm concentration, count, normal morphology, and linearity	Medium. All domains were evaluated as at least adequate.	Limited suitability. Insufficient availability of data or models to relate urinary metabolites to exposure levels.
Yin et al. (2017)	General population cohort study of 109 mother–child pairs (singleton births without congenital problems); recruited between July 2011 and May 2012 at the Shengsi People's Hospital on the Shengsi Islands (East China Sea) in China.	Nonspecific route of exposure	Umbilical cord serum levels of naphthalene	Reproductive hormone levels in umbilical cord blood serum (testosterone, estradiol, gonadotropins, AMH)	Significant negative association with cord AMH	Medium. All domains except sensitivity were evaluated as at least adequate.	Limited suitability. Hormone measurements at childbirth may not accurately reflect the situation in the infant endocrine system during pregnancy because of negative feedback by placental hormones.
Meeker et al. (2008)	General population cross-sectional health survey of 322 adult males in United States (Massachusetts).	Nonspecific route of exposure	Urinary metabolite levels (1N; 2N also measured but is not shown in the analysis)	Serum reproductive hormone levels: estradiol and prolactin	Nonsignificant inverse association between 1N and estradiol levels	Low. All domains except outcome and exposure measurement were evaluated as adequate. "However, because 1N is a major urinary metabolite of both naphthalene and the insecticide carbaryl, exposure misclassification stemming from differences in exposure source is possible and interpretation of the results is limited." (Meeker et al. 2007)	Limited suitability. Cross-sectional study design with limited ability to assess temporality. Significant concern for exposure misclassification due to use of urinary 1N. Also, insufficient availability of data or models to relate urinary metabolites to exposure levels.
Meeker et al. (2006)[a]	General population cross-sectional health survey of 322 adult males (mean age 36.1 y) in United States (Massachusetts).	Nonspecific route of exposure	Urinary metabolite levels (1N)	Serum reproductive hormone levels: gonadotropins, inhibin B, testosterone, and sex hormone-binding globulin	Inverse association with testosterone	Low. All domains except outcome and exposure measurement were evaluated as adequate. "However, because 1N is a major urinary metabolite of both naphthalene and the insecticide carbaryl, exposure misclassification stemming from differences in exposure source is possible and interpretation of the results is limited." (Meeker et al. 2007)	Limited suitability. Cross-sectional study design with limited ability to assess temporality. Significant concern for exposure misclassification due to use of urinary 1N. Also, insufficient availability of data or models to relate urinary metabolites to exposure levels.
Meeker et al. (2004a)[a]	General population cross-sectional health survey of 272 adult men in United States (Massachusetts).	Nonspecific route of exposure	Urinary metabolite levels (1N)	Sperm parameters	Statistically significant association with below-reference sperm concentration and percent motile sperm; nonsignificant association with abnormal sperm morphology	Low. All domains except exposure measurement were evaluated as at least adequate. "However, because 1N is a major urinary metabolite of both naphthalene and the insecticide carbaryl, exposure misclassification stemming from differences in exposure source is possible and interpretation of the results is limited." (Meeker et al. 2007)	Limited suitability. Cross-sectional study design with limited ability to assess temporality. Significant concern for exposure misclassification due to use of urinary 1N. Also, insufficient availability of data or models to relate urinary metabolites to exposure levels.

Table 5. (*Continued.*)

Author and publication year	Study description	Route of exposure	Exposure measurement	Outcome(s) evaluated[a]	Outcome(s) observed	Overall confidence level for outcome	Applicability for dose–response
Meeker et al. (2004b)[a]	General population cross-sectional health survey of 260 adult men (mean age 36.1 y) in United States (Massachusetts).	Nonspecific route of exposure	Urinary metabolite levels (1N)	DNA damage in sperm, measured using comet assay	Association with DNA damage in sperm (statistically significant for % DNA located in comet tail; not significant for comet extent and tail distributed moment)	Low. All domains except exposure measurement were evaluated as at least adequate. "However, because 1N is a major urinary metabolite of both naphthalene and the insecticide carbaryl, exposure misclassification stemming from differences in exposure source is possible and interpretation of the results is limited." (Meeker et al. 2007)	Limited suitability. Cross-sectional study design with limited ability to assess temporality. Significant concern for exposure misclassification due to use of urinary 1N. Also, insufficient availability of data or models to relate urinary metabolites to exposure levels.
Animal studies (inhalation)							
Dodd et al. (2012)	Male and female rats (Fischer 344); 90-d exposure	Inhalation (whole body)	0, 0.5, 5.2, 52, 157 mg/m³	Testis, ovary, and uterus weight	Statistically significant decrease in absolute testis weight	High. This study was well-designed to evaluate these outcomes. Evidence was presented clearly and transparently.	Suitable for testis data only. Multidose study; quantitative data reported for testis but not ovary or uterus.
NTP (2000)	Male and female rats (Fischer 344); 2-y (105-wk) exposure	Inhalation (whole body)	0, 52, 157, 314 mg/m³	Histopathology of testis, epididymis, seminal vesicles, preputial gland, prostate, ovary, uterus, clitoral gland, and mammary glands	No effects observed	High. This study was well-designed to evaluate these outcomes. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.
NTP (1992b)	Male and female mice (B6C3F1); 2-y (103-wk) exposure	Inhalation (whole body)	0, 52, 157 mg/m³	Histopathology of testis, epididymis, seminal vesicles, prostate, ovaries, uterus, and mammary glands	No effects observed	High (females). This study was well designed to evaluate these outcomes. Evidence was presented clearly and transparently. Low (males). The high mortality rate in control males has the potential to interfere with the interpretation of results.	Suitable. Multidose study with quantitative data.
Animal studies (oral)							
Battelle (1980b)	Male and female rats (Fischer 344); 90-d exposure	Oral gavage	0, 25, 50, 100, 200, 400 mg/kg-d	Histopathology of testis, prostate, ovaries, uterus, and mammary glands	No effects observed	High. This study was well-designed to evaluate these outcomes. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.
Battelle (1980a)	Male and female mice (B6C3F1); 90-d exposure	Oral gavage	0, 12.5, 25, 50, 100, 200 mg/kg-d	Histopathology of testes, prostate, ovaries, uterus, and mammary glands	No effects observed	High. This study was well-designed to evaluate these outcomes. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.
NTP (1992a)[b]	Female rabbits (New Zealand); exposure from GD 6–19, sacrifice on GD 30	Oral gavage	0, 20, 80, 120 mg/kg-d	Maternal body weight gain, gravid uterine weight, number pregnant at sacrifice	Decreased maternal body weight gain during treatment period (statistically significant trend)	Low (maternal body weight, uterine weight). Rabbit maternal body weight gain is known to have high variability, and therefore this is not an ideal species to use for evaluating this end point. High (number pregnant at sacrifice). This study was well-designed to evaluate this outcome. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.
NTP (1991)[b]	Female rats (Sprague-Dawley); exposure from GD 6–15, sacrifice on GD 20	Oral gavage	0, 50, 150, 450 mg/kg-d	Maternal body weight gain, gravid uterine weight, number pregnant at sacrifice	Statistically significant decrease in maternal body weight gain during gestation (corrected for gravid uterine weight)	High. This study was well-designed to evaluate these outcomes. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.
Pharmakon Research (1985)[b,c]	Female rabbits (New Zealand); exposure from GD 6–18, sacrifice on GD 29	Oral gavage	0, 50, 250, 630, 1,000 mg/kg-d	Maternal body weight gain, uterine weight, number pregnant at sacrifice	Statistically significant decrease in maternal body weight during gestation (occurring at doses with high maternal mortality)	Low (maternal body weight, uterine weight). Rabbit maternal body weight gain is known to have high variability, and therefore this is not an ideal species to use for evaluating this end point. High (pregnant at sacrifice). This study was well-designed to evaluate this outcome. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.

Table 5. (*Continued.*)

Author and publication year	Study description	Route of exposure	Exposure measurement	Outcome(s) evaluated[a]	Outcomes(s) observed	Overall confidence level for outcome	Applicability for dose–response
Shopp et al. (1984)	Male and female mice (CD-1); 14- or 90-d exposure	Oral gavage	14-d exposure: 0, 27, 53, 267 mg/kg-d 90-d exposure: 0, 5.3, 53, 133 mg/kg-d	Testis weight	No effects observed	High. This study was well-designed to evaluate this outcome. Evidence was presented clearly and transparently.	Suitable. Multidose studies with quantitative data.
Plasterer et al. (1985)	Female mice (CD-1); exposure from GD 7–14, dams allowed to deliver	Oral gavage	0, 300 mg/kg-d	Maternal body weight gain, reproductive index (survivors delivered/pregnant survivors)	Statistically significant decrease in maternal body weight gain during gestation	Medium (maternal body weight). Body weight gain was not able to be corrected for gravid uterine weight (due to the dams giving live birth), which affects the usability of the data because maternal toxicity cannot be distinguished from fetal toxicity. High (reproductive index). This study was well-designed to evaluate this outcome. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.
Pharmakon Research (1986)[b]	Female rabbits (New Zealand); exposure from GD 6–18, sacrifice on GD 29	Oral gavage	0, 40, 200, 400 mg/kg-d	Maternal body weight gain, premature delivery, number pregnant at sacrifice	No effects observed	Low (maternal body weight). Rabbit maternal body weight gain is known to have high variability, and therefore this is not an ideal species to use for evaluating this end point. High (premature delivery, pregnant at sacrifice). This study was well-designed to evaluate these outcomes. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.
Animal studies (dermal)							
Bushy Run (1986)	Male and female rats (Sprague-Dawley CD); 90-d exposure	Dermal	0, 100, 300, 1,000 mg/kg-d	Gonad weights; histopathology of testes, epididymis, seminal vesicles, prostate, vagina, ovaries, uterus, and mammary glands.	Statistically significant decrease in absolute and relative testes weight (no effects on testis weight observed in subgroup allowed to recover for 4 wk after exposure)	High. This study was well designed to evaluate these outcomes. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.

Note: AMH, anti-Müllerian hormone; GD, gestational day; HAWC, U.S. EPA Health Assessment Workspace Collaborative; PAH, polycyclic aromatic hydrocarbons.
[a]Meeker et al. (2007) presents a reanalysis of data from Meeker et al. 2004a, 2004b, and 2006; not shown in this table.
[b]Corpora lutea counts from gestational exposure studies would not have been affected by naphthalene exposure because all exposures began after animals were already pregnant. Corpora lutea counts from these studies may be useful as a metric of baseline fertility and are extracted in HAWC but are not discussed in this table.
[c]Pharmakon Research (1985) was a dose-range finding study with a 50% maternal mortality rate at 630 mg/kg-d and 100% mortality at 1,000 mg/kg-d.

Table 6. Summary of studies evaluating developmental effects.

Author and publication year	Study description	Route of exposure	Exposure measurement	Outcome(s) evaluated	Outcomes(s) observed	Overall confidence level for outcome	Applicability for dose–response
Human studies (nonspecific route of exposure)							
Agarwal et al. (2020)	General population cross-sectional health survey of 110 pregnant women (18–40 y old excluding smokers, and those who had previous history of serious chronic disease or had pregnancy complications) with attending antenatal care at Gynecology Department, S. N. Medical College, Agra, India, during the period from March 2017 to September 2018	Nonspecific route of exposure	Placental tissue levels of naphthalene	Birth weight	No statistically significant association of birth weight with naphthalene exposure	Medium. All domains were evaluated as at least adequate.	Limited suitability. Cross-sectional study design with limited ability to assess temporality. Also, insufficient availability of data or models to relate urinary metabolites to exposure levels.
Gong et al. (2018)	General population cross-sectional health survey of all registered births in Texas by Texas residents from 1996–2008; case–control design was used to investigate the associations between maternal residential proximity to the emission sources of these chemicals and LBW.	Nonspecific route of exposure	Estimated based on residential proximity to industrial facilities reporting to U.S. EPA's Toxic Release Inventory	LBW	Statistically significantly higher odds of LBW in maternal residences exposed to naphthalene	Medium. All domains are evaluated as at least adequate.	Not suitable. No quantitative exposure level data. Cross-sectional study design with limited ability to assess temporality.
Nie et al. (2018)	General population cross-sectional health survey of 287 pregnant women in Taiyuan, China	Nonspecific route of exposure	Creatinine-adjusted urinary metabolite levels (2-OH-naphthalene)	Birth weight, birth length, birth head circumference, and two growth proportionality indices.	Statistically significant association with lower birth weight and higher cephalization index	Medium. All domains are evaluated as at least adequate.	Limited suitability. Cross-sectional study design with limited ability to assess temporality. Also, insufficient availability of data or models to relate urinary metabolites to exposure levels.
Wang et al. (2014)	General population cross-sectional health survey of preschool-age children (4–5 y) in Nanjing, China	Nonspecific route of exposure	Indoor and outdoor dust levels of naphthalene	Neurobehavioral scores based on the Child Behavior Checklist and Gesell Development Inventory	Statistically significant association with higher "internalizing problem" score on the Child Behavior Checklist, which is a measure of anxiety, depression, withdrawal, and/or somatic complaints	Medium. All domains except sensitivity were evaluated as adequate.	Limited suitability. Exposure levels are reported only as overall mean, median, and range. Cross-sectional study design with limited ability to assess temporality.
Animal studies (oral)							
NTP (1992a)	Female rabbits (New Zealand); exposure from GD 6–19, sacrifice on GD 30	Oral gavage	0, 20, 80, 120 mg/kg-d	Fetal viability (live and dead fetuses, implantations, resorptions), fetal body weight, and gross, visceral, and skeletal alterations	Increased incidence of malformed female fetuses (statistically significant trend); the malformation increased in the fetuses was fused sternebrae.	High. This study was well-designed to evaluate these outcomes. Evidence was presented clearly and transparently.	Suitable. Multidose studies with quantitative data.
NTP (1991)	Female rats (Sprague-Dawley); exposure from GD 6–15, sacrifice on GD 20	Oral gavage	0, 50, 150, 450 mg/kg-d	Fetal viability (live and dead fetuses, implantations, resorptions), fetal body weight, and gross, visceral, and skeletal alterations	Increased incidence of malformed fetuses and adversely affected implants per litter (statistically significant trend), and decreased fetal body weight (statistically significant trend)	High. This study was well-designed to evaluate these outcomes. Evidence was presented clearly and transparently.	Suitable. Multidose studies with quantitative data.
Pharmakon Research (1986)	Female rabbits (New Zealand); exposure from GD 6–18, sacrifice on GD 29	Oral gavage	0, 40, 200, 400 mg/kg-d	Fetal viability (live and dead fetuses, implantations, resorptions), fetal body weight, and gross, visceral, and skeletal alterations	Statistically significant decrease in fetuses with extra 13th right-sided rib	High. This study was well-designed to evaluate these outcomes. Evidence was presented clearly and transparently	Suitable. Multidose studies with quantitative data.
Pharmakon Research (1985)	Female rabbits (New Zealand); exposure from GD 6–18, sacrifice on GD 29	Oral gavage	0, 50, 250, 630, 1,000 mg/kg-d	Fetal viability (live and dead fetuses, implantations, resorptions), fetal body weight, and gross, visceral, and skeletal alterations	No effects in the 50 or 250 mg/kg-d groups. (Fetuses in 630 and 1,000 mg/kg-d groups were not evaluated due to mortality or abortion in all does in these dose groups.)	High. This study was well-designed to evaluate these outcomes. Evidence was presented clearly and transparently	Suitable. Multidose studies with quantitative data.
Plasterer et al. (1985)	Female mice (CD-1); exposure from GD 7–14, dams allowed to deliver	Oral gavage	0, 300 mg/kg-d	Pup viability, postnatal body weight	Statistically significant decrease in live pups per litter on PND 1 and 3, with no change in dead pups per litter; suggests that this effect was due to early resorptions	High. This study was well-designed to evaluate these outcomes. Evidence was presented clearly and transparently	Suitable. Multidose studies with quantitative data.

Note: GD, gestational day; LBW, low birth weight; PND, postnatal day.

Table 7. Summary of studies evaluating ocular effects.

Author and publication year	Study description	Route of exposure	Exposure measurement	Outcome(s) evaluated	Outcomes(s) observed	Overall confidence level for outcome	Applicability for dose–response
Animal studies (inhalation)							
NTP (2000)	Male and female rats (Fischer 344); 2-y (105-wk) exposure	Inhalation (whole body)	0, 52, 157, 314 mg/m^3	Histopathology of eyes	No effects observed	High. This study was well designed to evaluate this outcome. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.
NTP (1992b)	Male and female mice (B6C3F1); 2-y (103-wk) exposure	Inhalation (whole body)	0, 52, 157 mg/m^3	Examination of eyes by slit-lamp biomicroscopy and indirect ophthalmoscopic procedures	No effects observed	High (females). This study was well designed to evaluate this outcome. Evidence was presented clearly and transparently. Low (males). The high mortality rate in control males has the potential to interfere with the interpretation of results.	Suitable. Multidose study with quantitative data.
Animal studies (oral)							
Battelle (1980b)	Male and female rats (Fischer 344); 90-d exposure	Oral gavage	0, 25, 50, 100, 200, 400 mg/kg-d	Histopathological evaluation of eyes	No effects observed	High. This study was well designed to evaluate this outcome. Evidence was presented clearly and transparently.	Suitable. Multidose studies with quantitative data.
Battelle (1980a)	Male and female mice (B6C3F1); 90-d exposure	Oral gavage	0, 12.5, 25, 50, 100, 200 mg/kg-d	Histopathological evaluation of eyes	No effects observed	High. This study was well designed to evaluate this outcome. Evidence was presented clearly and transparently.	Suitable. Multidose studies with quantitative data.
Holmén et al. (1999)	Female rats (Brown Norway); 10-wk exposure	Oral gavage	0, 100, 500, 1,000, 1,500 mg/kg-d (dosed twice per week)	Evaluation of cataractous changes using lens photography	Cataractous changes correlated with naphthalene dose	Medium. Some concerns were raised over the small sample size of the control and some reporting limitations (method of dosing not reported, chemical purity not reported), but overall appears to be a well-conducted study.	Suitable. Multidose studies with quantitative data.
Shopp et al. (1984)	Male and female mice (CD-1); 14- or 90-d exposure	Oral gavage	14-d exposure: 0, 27, 53, 267 mg/kg-d; 90-d exposure: 0, 5.3, 53, 133 mg/kg-d	Examination of eyes (methods not reported)	No effects observed	Medium. This study was well-designed to evaluate these outcomes, but confidence was decreased because only qualitative results are provided.	Not suitable. Quantitative results not reported.
Animal studies (dermal)							
Bushy Run (1986)	Male and female rats (Sprague-Dawley CD); 90-d exposure	Dermal	0, 100, 300, 1,000 mg/kg-d	Examination with indirect ophthalmoscope prior to dosing and prior to final sacrifice	No effects observed	High. This study was well designed to evaluate this outcome. Evidence was presented clearly and transparently.	Suitable. Multidose study with quantitative data.

because it was not corrected for gravid uterine weight (due to the animals giving live birth), which affects the interpretation of the data because maternal toxicity cannot be distinguished from fetal toxicity (U.S. EPA 1991); and three studies that evaluated maternal weight gain in rabbits were rated *low* confidence for this end point (NTP 1992a; Pharmakon Research 1985, 1986) because body weight changes in rabbits are reported to have high variability, so this measurement may be a less reliable indicator of maternal toxicity in comparison with that of other species (U.S. EPA 1991). The histopathology data in males from NTP (1992b) was also rated *low* confidence due to the high mortality rate in control animals. Although most studies reported no effects, there were some effects consisting of decreased testis weight (Dodd et al. 2012; Bushy Run 1986) and decreased maternal body weight gain (NTP 1991, 1992a; Pharmakon Research 1985; Plasterer et al. 1985). However, the authors of the dermal exposure study by Bushy Run (1986) doubted the biological significance of the decreased testis weight in their study because the magnitude of change was small, effects were not observed in a subgroup of animals that was allowed to recover for 4 wk post exposure, and there was little evidence of histopathological effects to support an effect on the testis. All studies reported quantitative data suitable for dose–response analysis. Effects on testis weight and maternal body weight gain are considered indicative of reproductive effects (U.S. EPA 1996), but the strength of the animal evidence base is limited by the absence of studies evaluating functional end points such as fertility or sperm quality that could provide more specific insight into how the reproductive system is affected by naphthalene exposure. This lack of functional end point measurements limits the applicability of these studies for hazard identification and reference value derivation.

Developmental effects. Table 6 summarizes the evidence base for developmental effects. The four available epidemiology studies were all considered *medium* confidence, and three reported statistically significant associations between naphthalene exposure and several developmental outcomes, including low birth weight (Gong et al. 2018; Nie et al. 2018), high cephalization index (Nie et al. 2018), and increased measures of "internalizing problems" on the Child Behavior Checklist (Wang et al. 2014), whereas the study by Agarwal et al. (2020) reported no statistically significant association of birth weight with naphthalene exposure. All studies were found to have limited suitability for dose–response analysis. The study by Gong et al. (2018) estimated exposure to naphthalene emissions using a geocomputational approach and therefore did not include quantitative exposure level data. The study by Wang et al. (2014) reported exposure level (naphthalene levels in house dust) only as an overall mean, median, and range for all individuals included in the study, which provides little insight into the exposure–response relationship. Nie et al. (2018) and Agarwal et al. (2020) used urinary biomarkers as the only metric for exposure. All four studies used cross-sectional designs that have limited ability to assess temporality. In addition to these four studies that evaluated traditional developmental end points, there were epidemiology studies that evaluated other effects in children that risk assessors may wish to consider for hazard identification of developmental effects: childhood obesity (Bushnik et al. 2020; Scinicariello and Buser 2014) (see Table S8, "Cardiometabolic") and neonatal jaundice (Sodeinde et al. 1995; Familusi and Dawodu 1985) (see Table S8, "Hepatic"). We did not do a full evaluation of the childhood obesity or jaundice studies but noted that they are not suitable for dose–response analysis.

The animal evidence base for developmental effects consisted of five *high* confidence gestational exposure studies in rats, mice,

and rabbits, all of which were found to be suitable for dose–response analysis. Four of these studies observed effects consisting of increased incidence of fetal malformations (NTP 1991, 1992a), increased adversely affected implants (nonlive or malformed) (NTP 1991), decreased incidence of extra 13th right-side rib (Pharmakon Research 1986), and decreased live fetuses per litter (Plasterer et al. 1985). Fetal malformations, adversely affected implants, and decreased live fetuses per litter are the most informative of these outcomes for risk assessment, whereas the biological significance of decreased extra ribs (a skeletal variation) is less clear (Chernoff and Rogers 2004). The fifth study by Pharmakon Research (1985) is a dose–range finding study in which fetuses in the highest dose groups were not evaluated due to mortality and abortion in all dams (100% mortality of does in the 1,000 mg/kg-d group and 50% mortality and 50% abortion in dams in the 630 mg/kg-d group), and no developmental effects were reported in the lowest two dose groups.

Ocular effects. Cataracts have been frequently observed in human case reports involving acute or occupational exposure to naphthalene, although exposure levels associated with these effects have not been identified (ATSDR 2005). No epidemiology studies were identified in our literature search that evaluated cataract formation or other ocular effects. Six multidose animal studies evaluated ocular effects in rats or mice, most of which reported quantitative data suitable for dose–response analysis and all of which were considered *high* or *medium* confidence, with the exception of data from male mice in NTP (1992b) (Table 7). However, only one study (Holmén et al. 1999) observed a dose-related increase in the development of cataracts (in dose groups 500 mg/kg-d and above), whereas the remaining studies in Table 7 observed no effect. As mentioned in the "Literature Search and Screening" section, naphthalene at high oral exposure levels (>500 mg/kg-d) is well known to induce cataracts in rodent models and is frequently used in animal models of cataractogenesis; many such studies were identified in our literature search and may be used to support hazard identification but were not moved forward for study evaluation because they used single high doses of naphthalene and are unlikely to be useful for dose–response analysis. The lack of effect in the studies by NTP (1992b, 2000), Battelle (1980a, 1980b), Shopp et al. (1984), and Bushy Run (1986) may therefore be related to the lower dose levels used in most of those studies or due to differences in strain or route of exposure.

Summary across health systems. Figure 9 presents a heat map summarizing the overall confidence level and suitability for dose–response analysis for each health system for the studies described in Tables 2–7. Overall, one *medium* confidence epidemiology study evaluating respiratory effects was found to be suitable for dose-response analysis, whereas there were *high* or *medium* confidence animal studies available for each health system that were found to be suitable for dose–response analysis. This heat map summarizes the number of references that investigated a particular health effect, not the number of studies that identified a positive association with exposure to naphthalene; LOAELs for each health system are summarized in the next section.

Comparison of LOAELs across Health Systems

Figures 10 and 11 present the dose ranges and lowest LOAELs within each of the health systems across *high* and *medium* confidence animal studies for inhalation and oral exposures, respectively. Results are summarized according to duration of exposure (acute, short-term, subchronic, or chronic). LOAELs are not summarized for dermal exposure because we only evaluated one study [90-d study by Bushy Run (1986)], and the only effect in

that study was decreased testis weight at 1,000 mg/kg-d (no hematologic, immune, or ocular effects observed).

Respiratory lesions were observed in the chronic (2-y) inhalation studies in rats and mice at the lowest dose tested in these studies (52 mg/m³). Effects at this level consisted of nasal respiratory epithelial adenomas in male rats in the study by NTP (2000) and nonneoplastic lesions in both studies (NTP 1992b, 2000). Lower LOAELs for respiratory lesions were observed in the subchronic (90-d) exposure study by Dodd et al. (2012) (5.2 mg/m³) and in the acute (1-d) and short-term (5-d) studies by Dodd et al. (2010) (5.2 mg/m³), which tested lower dose ranges compared with the 2-y NTP studies. LOAELs for immune, reproductive, and hematologic outcomes in inhalation studies occurred at a similar concentration range compared with respiratory effects. The LOAEL for immune effects was 5.1 mg/m³ [decreased thymus weight in female rats (Dodd et al. 2012)], and the LOAEL for reproductive and hematologic effects was 52 mg/m³ [decreased testis weight in rats (Dodd et al. 2012) and increased leukocyte count in mice (NTP 1992b)]. The LOAEL for hematologic effects in the 2-y study by NTP (1992b) is classified here as "short-term," because this evaluation was performed on day 14 of the study.

In oral exposure studies, LOAELs were within an order of magnitude of one another across outcome categories and durations of exposure, with ocular effects being the least sensitive outcome. LOAELs among subchronic oral exposure studies were 133 mg/kg-d for immune effects [decreased spleen weight in female mice, Shopp et al. (1984)], 400 mg/kg-d for hematologic effects [decreased hemoglobin and hematocrit in male and female rats, increased mature neutrophils in male and female rats, and decreased lymphocytes in male rats (Battelle 1980b)], and 500 mg/kg-d for ocular effects [cataractous changes (Holmén et al. 1999)]. LOAELs among short-term oral exposure studies (including gestational exposure studies) were 150 mg/kg-d for reproductive effects [decreased maternal body weight gain (NTP 1991)], 267 mg/kg-d for immune effects [decreased thymus, spleen, and lung weight (Shopp et al. 1984)], and 300 mg/kg-d for developmental effects [decreased live pups per litter (Plasterer et al. 1985)].

Discussion

This systematic evidence map summarizes human and animal studies that evaluated health effects associated with naphthalene exposure, with a focus on studies that may be potentially used to develop updated chronic reference value(s) for human health risk assessment. Within each of the health systems that were prioritized for further evaluation [respiratory (pulmonary and nasal), hematologic, immune, reproductive, developmental, ocular], there were one or more *high* or *medium* confidence multidose animal studies that observed exposure-related effects and appeared to be suitable for dose–response analysis. The available epidemiology studies had more concerns for bias and sensitivity (rated *medium* confidence, *low* confidence, or *uninformative*), and almost all were found to have limited suitability or no suitability for dose–response analysis. Among the epidemiology studies, the most frequent limitation for dose–response analysis was the use of urinary biomarkers (1N, 2N) as the only metric for exposure, which cannot be extrapolated to exposure concentrations because urine is not included as a compartment in any of the currently available PBPK models. Other common limitations for dose–response analysis included lack of exposure level data, lack of adjustment for confounding variables in the analysis, and inability to assess temporality due to the use of cross-sectional study designs.

Although none of the existing PBPK models are suitable for estimating human oral or inhalation exposure levels from metabolite concentrations in urine, the models of Campbell et al. (2014), Kim et al. (2007), and Kapraun et al. (2020) can be used to inform route-to-route and interspecies extrapolation. These models had not yet been published when most of the available reference values for naphthalene were developed. The chronic reference values from federal and state agencies (e.g., U.S. EPA, ATSDR, OEHHA) summarized in Figures 1 and 2 were all derived by applying interspecies uncertainty factors to extrapolate from animal NOAELS or LOAELs. Applying PBPK models in future assessments of naphthalene would allow for more mathematically rigorous quantification of the uncertainty inherent in inter- and intraspecies extrapolations and thus provide a more scientifically defensible estimate of the risk to human health.

We did not undertake a formal evaluation of the quality of all of the PBPK models available for naphthalene. However, when a more recent model builds on previously published models it is usually assumed to improve on the previous models through use of newer data, computational methods, and/or the most recent developments in the understanding of pharmacological principles. In the event that there are competing models, the technical criteria outlined by U.S. EPA (2018c) can be applied to identify

Health system	Human					Animal					Grand Total
	Medium			Low		High	Medium		Low		
	Suitable	Limited suitability	Not suitable	Limited suitability	Not suitable	Suitable	Suitable	Not suitable	Suitable	Limited suitability	
Nasal	1					4					5
Pulmonary	1	1				5	1				8
Hematological		1		1	1	1	3	1		1	9
Immunological		2		5		7					14
Reproductive		4		4		8	1		3		20
Developmental		3	1			5					9
Ocular						6	1	1			7

references
1 � 8

Figure 9. Visualization of the overall confidence ratings (high, medium, low) and suitability for dose–response analysis (suitable, limited suitability, not suitable) within each health system for the studies presented in Tables 2–7. See Tables 2–7 for details. Numbers represent the number of distinct references that investigated a particular health system, not the number that identified a positive association with exposure to naphthalene. The grand total column indicates the total number of human and animal references for each health system. An interactive version of this figure is available in the "Summary of Included Studies" tab in the Tableau Public dashboard (https://public.tableau.com/views/NaphthaleneEvidenceMap/ReadMe?:language=en&:display_count=y&publish=yes&:origin=viz_share_link).
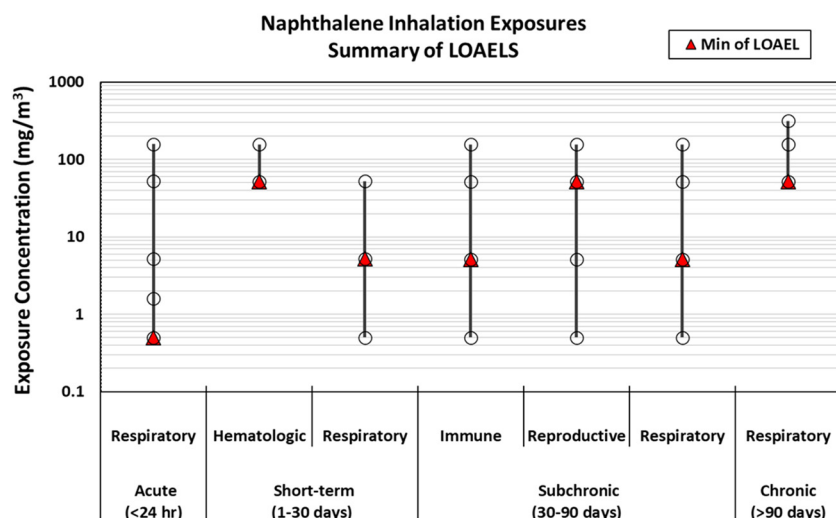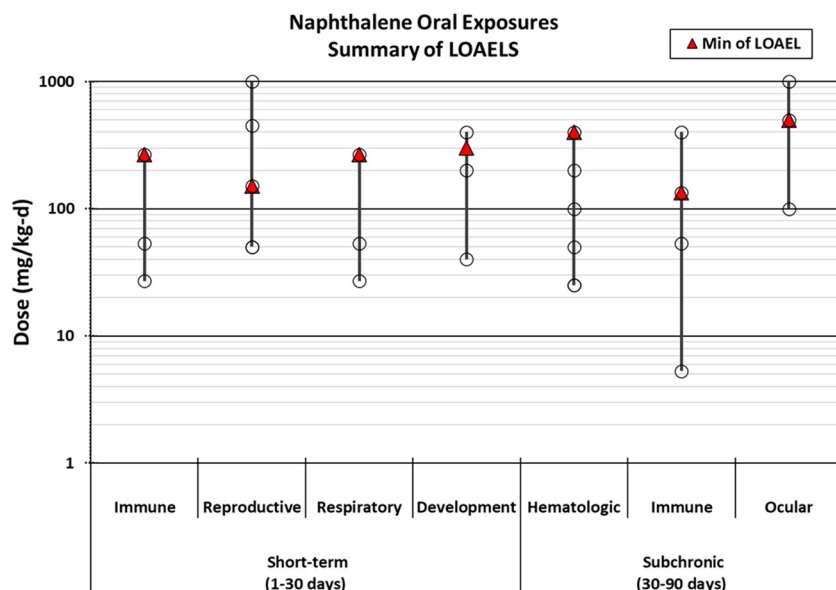
Figure 10. Summary of the dose ranges and lowest LOAELs within each health system across high and medium confidence animal inhalation exposure studies. Open circles represent the doses tested across all studies that reported a LOAEL for each health system. Note: LOAEL, lowest observed adverse effect level.

the most suitable or preferred model and determine that it meets minimal quality criteria. Because the number of models that exist for a given chemical is generally small and a large array of features can differ among models, this determination is usually accomplished by having a review team read each of the corresponding publications and discuss the features and overall quality of each model. In some cases, model selection may be clear and simple, such as when one model includes a route of exposure that is needed for the assessment and an alternate model does not. Other aspects may require more nuanced evaluation, such as when different data sets are used to identify some of the parameters for each model and there is not a clear distinction between the quality of the data sets. In any case the rationale for selecting among alternate models would be provided, where it is not readily established due to model "lineage" as described above. In the case of naphthalene, the model of Campbell et al. (2014) is the penultimate model in its lineage and explicitly describes the distribution of dosimetry of specific regions in the upper respiratory

tract, a feature that distinguishes it from all previous models. Kapraun et al. (2020) extended the model of Campbell et al. (2014) by incorporating a skin route of exposure and demonstrated that their model could be used to reproduce human pharmacokinetic data; they also performed quality assurance procedures (U.S. EPA 2018c) for their model. This most recently published naphthalene PBPK model (Kapraun et al. 2020) would therefore be the clear choice for human health risk assessment applications.

The existing chronic reference values for naphthalene from U.S. federal and state agencies and Health Canada that are intended to assess risk in the general public were all derived based on studies evaluated in this evidence map, specifically the 2-y rodent inhalation bioassays by NTP (1992a, 2000), the 90-d study in rats by Battelle (1980b) and the gestational exposure study in rats by NTP (1991). We generally found these studies to be *high confidence*, although there were limitations; most notably, the high attrition rate of control males in the 2-y study in mice by NTP (1992b)

Figure 11. Summary of the dose ranges and lowest LOAELs within each health system across high and medium confidence animal oral exposure studies. Open circles represent the doses tested across all studies that reported a LOAEL for each health system. Note: LOAEL, lowest observed adverse effect level.

(resulting in a *low confidence* rating for males in this study but a *high confidence* rating for females). Inhalation reference values from IRIS and OEHHA were developed prior to the publication of NTP's 2-y study in rats (NTP 2000) and are based on the 2-y study in mice (NTP 1992b). An updated inhalation reference value for naphthalene would benefit from considering both of the NTP 2-y studies as well as the more recent subchronic study by Dodd et al. (2012) that tested a lower dose range and observed lower LOAELs compared to the 2-y studies. Acute or short-term studies such as the 1- and 5-d inhalation studies by Dodd et al. 2010 are less informative for developing a chronic reference dose, but they provide useful information for the weight of evidence on the dose- and time-related progression of respiratory effects. No new oral exposure studies for naphthalene have been published since the release of the existing assessments.

Respiratory effects were the most frequent outcomes evaluated in the inhalation studies, although a comparison of LOAELs indicates that reproductive, hematologic, and immune effects occurred at a similar concentration range compared to respiratory effects in inhalation studies. The available animal studies are informative of the progression of nasal and pulmonary lesions following naphthalene exposure, including the development of tumors in both rats and mice in 2-y studies, and the epidemiology studies examining respiratory effects provided measurements of respiratory inflammation and changes in lung function associated with naphthalene exposure (Table 1). Although case reports in humans have documented increased laryngeal cancer following occupational exposure to naphthalene (Wolf 1976, 1978), there were no epidemiology studies evaluating the association between naphthalene exposure and cancer that would allow for quantitative evaluation of the exposure–response relationship.

To develop reliable estimates of low-dose cancer risk from animal studies of naphthalene, two key areas of scientific interest are *a*) the MOA for tumorigenesis and *b*) interspecies differences in naphthalene bioactivation (Bogen et al. 2008; North et al. 2008). The mechanistic studies and ADME/toxicokinetic studies identified in this evidence mapping exercise can play a critical role in the evaluation of these issues. Following exposure, naphthalene is rapidly metabolized via a CYP-mediated pathway into a reactive epoxide intermediate (1,2-naphthalene oxide) that then undergoes further metabolism into multiple quinone intermediates (1,2-naphthoquinone; 1,4-naphthoquinone) (Carratt et al. 2016; Li et al. 2011; Waidyanatha et al. 2002). Naphthalene-induced lesions are thought to be related to the binding of these reactive naphthalene metabolites to cellular proteins, as well as oxidative damage resulting from the reactive epoxide intermediate and from the participation of naphthalene quinone metabolites in redox cycles. Both cytotoxic and genotoxic MOAs have been hypothesized for naphthalene metabolites (ATSDR 2005), and it will be necessary to consider the plausibility of these respective MOAs when selecting a modeling approach for low-dose extrapolation of carcinogenic effects (U.S. EPA 2005). Regarding interspecies differences in metabolism, it has been established via *in vitro* and *in vivo* in mouse models that the initial epoxidation of naphthalene is catalyzed by Cyp2f2 and Cyp2a5 in the lung and olfactory epithelia, respectively, with histopathological lesions correlating with the rate of *in situ* metabolism (Hu et al. 2014; Li et al. 2011; Cruzan et al. 2009; Buckpitt et al. 2002; Ritter et al. 1991; Nagata et al. 1990). There is controversy surrounding whether CYP2F1 (the human ortholog of Cyp2f2) is capable of metabolizing naphthalene into the initial epoxide metabolite that leads to the two toxic quinone moieties. An *in vitro* study by Baldwin et al. (2005) found that CYP2F1 was unable to metabolize naphthalene into its epoxide form, whereas a recent *in vivo* study by Li et al. (2017) found that mice with humanized CYP2A13/2F1 displayed greater toxicity than Cyp2abfgs-

null mice and indicated that human CYP2F1 is an active mediator of naphthalene-induced focal toxicity in the lung. Our group is currently undertaking a MOA analysis for naphthalene using data tagged in this evidence map as mechanistic supplemental material, which will provide a systematic evaluation of these interspecies differences in metabolism and the evidence to support cytotoxic and genotoxic MOAs for each naphthalene metabolite.

This evidence map identified several key data gaps that would benefit from additional research. The human evidence base would be strengthened through the addition of studies providing quantitative evidence of the exposure–response relationship for naphthalene-induced carcinogenesis, hemolytic anemia, and cataracts, particularly because there are human case reports indicating that naphthalene is associated with these outcomes. [Human case reports were identified in this evidence map as potentially relevant supplemental material and not evaluated further; however, they have been summarized elsewhere, e.g., by ATSDR (2005).] The available epidemiology studies provided evidence of other types of effects associated with naphthalene exposure (e.g., effects on the respiratory system, hemoglobin, asthma, sperm quality, birth weight) but were largely limited by weaknesses in experimental design and reporting. Other issues that may warrant further investigation are age- or sex-related differences in naphthalene susceptibility. For instance, a 4-h inhalation study in mice by Carratt et al. (2019) (which was not selected for further evaluation in this evidence map due to the acute exposure duration but is summarized in the Tableau Public database) reported that juvenile (3-wk-old) mice were more susceptible to naphthalene-induced airway epithelial cytotoxicity compared with neonatal (7-d-old) or adult mice, with female juveniles being the most susceptible. The authors hypothesized this could be due to age- or sex-specific deficiencies in the ability to synthesize glutathione (a detoxification mechanism for naphthalene). None of the subchronic or chronic study designs selected for further evaluation included neonatal or juvenile animals, which potentially limits the sensitivity of these studies; however, because glutathione conjugation may not be a dominant pathway of naphthalene detoxification in primates including humans (NTP 2000; Rozman et al. 1982; Summer et al. 1979), it is unclear whether the age- and sex-specific differences in animal models are relevant to humans. Hazard identification and dose–response assessment for naphthalene could be strengthened by conducting additional studies addressing these limitations.

The considerations for reference value derivation extend beyond what is presented in this evidence map and will vary across entities, depending on stakeholder needs. We anticipate that the aggregate data set identified in this evidence map—including human and animal health effect studies, PBPK models, mechanistic studies, and ADME/toxicokinetic studies—can provide a useful starting point toward developing an updated reference value that reflects the state of the science on the health effects of naphthalene exposure. Although we focused our evaluations on studies that could be used to develop chronic reference value(s), acute and short-term studies can be used for deriving reference value(s) for shorter duration exposure scenarios. Risk assessors interested in these studies can refer to our interactive dashboard in Tableau Public, where study designs and results are summarized.

Acknowledgments

References

Abdo KM, Eustis SL, McDonald M, Jokinen MP, Adkins B, Haseman JK. 1992. Naphthalene: a respiratory tract toxicant and carcinogen for mice. Inhal Toxicol 4(4):393–409, https://doi.org/10.3109/08958379209145317.

Abdo KM, Grumbein S, Chou BJ, Herbert R. 2001. Toxicity and carcinogenicity study in F344 rats following 2 years of whole-body exposure to naphthalene vapors. Inhal Toxicol 13(10):931–950, PMID: 11696867, https://doi.org/10.1080/089583701752378179.

Adkins B Jr, Van Stee EW, Simmons JE, Eustis SL. 1986. Oncogenic response of strain a/J mice to inhaled chemicals. J Toxicol Environ Health 17(2–3):311–322, PMID: 3083111, https://doi.org/10.1080/15287398609530825.

Agarwal P, Anand M, Chakraborty P, Singh L, Masih J, Taneja A. 2020. Placental levels of polycyclic aromatic hydrocarbons (PAHs) and their association with birth weight of infants. Drug Chem Toxicol 1–10, PMID: 32597233, https://doi.org/10.1080/01480545.2020.1783285.

Ajao OG, Adenuga MO, Lapido JK. 1988. Colorectal carcinoma in patients under the age of 30 years: a review of 11 cases. J R Coll Surg Edinb 33(5):277–279, PMID: 3230552.

Al-Daghri NM. 2008. Serum polycyclic aromatic hydrocarbons among children with and without asthma: correlation to environmental and dietary factors. Int J Occup Med Environ Health 21:211–217, PMID: 18980880, https://doi.org/10.2478/v10001-008-0021-0.

Al-Daghri NM, Alokail MS, Abd-Alrahman SH, Draz HM. 2014. Polycyclic aromatic hydrocarbon distribution in serum of saudi children using HPLC-FLD: marker elevations in children with asthma. Environ Sci Pollut Res Int 21(20):12085–12090, PMID: 24923226, https://doi.org/10.1007/s11356-014-3108-0.

Al-Daghri NM, Alokail MS, Abd-Alrahman SH, Draz HM, Yakout SM, Clerici M. 2013. Polycyclic aromatic hydrocarbon exposure and pediatric asthma in children: a case–control study. Environ Health 12:1, PMID: 23286340, https://doi.org/10.1186/1476-069X-12-1.

ATSDR (Agency for Toxic Substances and Disease Registry). 2005. Toxicological profile for naphthalene, 1-methylnaphthalene, and 2-methylnaphthalene. (NTIS/02937293). https://ntrl.ntis.gov/NTRL/dashboard/searchResults.xhtml?searchQuery=PB2006100004 [accessed 26 June 2021].

Baldwin RM, Shultz MA, Buckpitt AR. 2005. Bioactivation of the pulmonary toxicants naphthalene and 1-nitronaphthalene by rat CYP2F4. J Pharmacol Exp Ther 312(2):857–865, PMID: 15509722, https://doi.org/10.1124/jpet.104.075440.

Battelle (Battelle Columbus Laboratories). 1980a. Unpublished Subchronic Toxicity Study: Naphthalene (C52904), B6C3F1 Mice. (subcontract no. 76-34-106002). Research Triangle Park, NC: National Toxicology Program.

Battelle. 1980b. Unpublished Subchronic Toxicity Study: Naphthalene (C52904), Fischer 344 Rats. (subcontract no. 76-34-106002). Research Triangle Park, NC: National Toxicology Program.

Bogen KT, Benson JM, Yost GS, Morris JB, Dahl AR, Clewell HJ, et al. 2008. Naphthalene metabolism in relation to target tissue anatomy, physiology, cytotoxicity and tumorigenic mechanism of action. Regul Toxicol Pharmacol 51(suppl 2):S27–S36, PMID: 18191315, https://doi.org/10.1016/j.yrtph.2007.10.018.

Buckpitt A, Boland B, Isbell M, Morin D, Shultz M, Baldwin R, et al. 2002. Naphthalene-induced respiratory tract toxicity: metabolic mechanisms of toxicity. Drug Metab Rev 34(4):791–820, PMID: 12487150, https://doi.org/10.1081/dmr-120015694.

Bushnik T, Wong SL, Holloway AC, Thomson EM. 2020. Association of urinary polycyclic aromatic hydrocarbons and obesity in children aged 3–18: Canadian Health Measures Survey 2009–2015. J Dev Orig Health Dis 11(6):623–631, PMID: 31806062, https://doi.org/10.1017/S2040174419000825.

Bushy Run (Bushy Run Research Center). 1986. Ninety-day (sub-chronic) dermal toxicity study with naphthalene in albino rats [TSCA Submission]. (EPA/OTS Doc #86-870000565). Beacon, NY: Texaco. https://ntrl.ntis.gov/NTRL/dashboard/searchResults/titleDetail/OTS0513643.xhtml.

Cakmak S, Dales RE, Liu L, Kauri LM, Lemieux CL, Hebbern C, et al. 2014. Residential exposure to volatile organic compounds and lung function: results from a population-based cross-sectional survey. Environ Pollut 194:145–151, PMID: 25108490, https://doi.org/10.1016/j.envpol.2014.07.020.

Campbell JL, Andersen ME, Clewell HJ. 2014. A hybrid CFD-PBPK model for naphthalene in rat and human with IVIVE for nasal tissue metabolism and cross-species dosimetry. Inhal Toxicol 26(6):333–344, PMID: 24666369, https://doi.org/10.3109/08958378.2014.896059.

Carratt SA, Kovalchuk N, Ding X, Van Winkle LS. 2019. Metabolism and lung toxicity of inhaled naphthalene: effects of postnatal age and sex. Toxicol Sci 170(2):536–548, PMID: 31020322, https://doi.org/10.1093/toxsci/kfz100.

Carratt SA, Morin D, Buckpitt AR, Edwards PC, Van Winkle LS. 2016. Naphthalene cytotoxicity in microsomal epoxide hydrolase deficient mice. Toxicol Lett 246:35–41, PMID: 26840748, https://doi.org/10.1016/j.toxlet.2016.01.019.

Chao YCE, Kupper LL, Serdar B, Egeghy PP, Rappaport SM, Nylander-French LA. 2006. Dermal exposure to jet fuel JP-8 significantly contributes to the production of urinary naphthols in fuel-cell maintenance workers. Environ Health Perspect 114(2):182–185, PMID: 16451852, https://doi.org/10.1289/ehp.8288.

Chernoff N, Rogers JM. 2004. Supernumerary ribs in developmental toxicity bioassays and in human populations: incidence and biological significance. J Toxicol Environ Health B Crit Rev 7(6):437–449, PMID: 15586878, https://doi.org/10.1080/10937400490512447.

Crissman JW, Goodman DG, Hildebrandt PK, Maronpot RR, Prater DA, Riley JH, et al. 2004. Best practices guideline: toxicologic histopathology. Toxicol Pathol 32(1):126–131, PMID: 14713558, https://doi.org/10.1080/01926230490268756.

Cruzan G, Bus J, Banton M, Gingell R, Carlson G. 2009. Mouse specific lung tumors from CYP2F2-mediated cytotoxic metabolism: an endpoint/toxic response where data from multiple chemicals converge to support a mode of action. Regul Toxicol Pharmacol 55(2):205–218, PMID: 19589367, https://doi.org/10.1016/j.yrtph.2009.07.002.

Dishaw L, Yost E, Arzuaga X, Luke A, Kraft A, Walker T, et al. 2020. A novel study evaluation strategy in the systematic review of animal toxicology studies for human health assessments of environmental chemicals. Environ Int 141:105736, PMID: 32434117, https://doi.org/10.1016/j.envint.2020.105736.

Dodd DE, Gross EA, Miller RA, Wong BA. 2010. Nasal olfactory epithelial lesions in F344 and SD rats following 1- and 5-day inhalation exposure to naphthalene vapor. Int J Toxicol 29(2):175–184, PMID: 20086191, https://doi.org/10.1177/1091581809357955.

Dodd DE, Wong BA, Gross EA, Miller RA. 2012. Nasal epithelial lesions in F344 rats following a 90-day inhalation exposure to naphthalene. Inhal Toxicol 24(1):70–79, PMID: 22182220, https://doi.org/10.3109/08958378.2011.636086.

Egeghy PP, Hauf-Cabalo L, Gibson R, Rappaport SM. 2003. Benzene and naphthalene in air and breath as indicators of exposure to jet fuel. Occup Environ Med 60(12):969–976, PMID: 14634191, https://doi.org/10.1136/oem.60.12.969.

Familusi JB, Dawodu AH. 1985. A survey of neonatal jaundice in association with household drugs and chemicals in Nigeria. Ann Trop Paediatr 5(4):219–222, PMID: 2418773, https://doi.org/10.1080/02724936.1985.11748397.

Ghanem A, Shuler ML. 2000. Combining cell culture analogue reactor designs and PBPK models to probe mechanisms of naphthalene toxicity. Biotechnol Prog 16(3):334–345, PMID: 10835232, https://doi.org/10.1021/bp9901522.

Gong X, Lin Y, Bell ML, Zhan FB. 2018. Associations between maternal residential proximity to air emissions from industrial facilities and low birth weight in Texas, USA. Environ Int 120:181–198, PMID: 30096612, https://doi.org/10.1016/j.envint.2018.07.045.

Holmén JB, Ekesten B, Lundgren B. 1999. Naphthalene-induced cataract model in rats: a comparative study between slit and retroillumination images, biochemical changes and naphthalene dose and duration. Curr Eye Res 19(5):418–425, PMID: 10520218, https://doi.org/10.1076/ceyr.19.5.418.5296.

Hu J, Sheng L, Li L, Zhou X, Xie F, D'Agostino J, et al. 2014. Essential role of the cytochrome P450 enzyme CYP2A5 in olfactory mucosal toxicity of naphthalene. Drug Metab Dispos 42(1):23–27, PMID: 24104196, https://doi.org/10.1124/dmd.113.054429.

IARC (International Agency for Research on Cancer). 2002. IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans: Some Traditional Herbal Medicines, Some Mycotoxins, Naphthalene, and Styrene. Lyon, France: IARC. http://monographs.iarc.fr/ENG/Monographs/vol82/mono82.pdf [accessed 26 June 2021].

Jia C, Batterman S. 2010. A critical review of naphthalene sources and exposures relevant to indoor and outdoor air [review]. Int J Environ Res Public Health 7(7):2903–2939, PMID: 20717549, https://doi.org/10.3390/ijerph7072903.

Kamal A, Malik RN, Martellini T, Cincinelli A. 2014. PAH exposure biomarkers are associated with clinico-chemical changes in the brick kiln workers in Pakistan. Sci Total Environ 490:521–527, PMID: 24878696, https://doi.org/10.1016/j.scitotenv.2014.05.033.

Kapraun DF, Schlosser PM, Nylander-French LA, Kim D, Yost EE, Druwe IL. 2020. A physiologically based pharmacokinetic model for naphthalene with inhalation and skin routes of exposure. J Toxicol Sci July, PMID: 32687177, https://doi.org/10.1093/toxsci/kfaa117.

Katsnelson BA, Minigaliyeva IA, Degtyareva TD, Privalova LI, Beresneva TA. 2014. Does a concomitant exposure to lead influence unfavorably the naphthalene subchronic toxicity and toxicokinetics? Environ Toxicol Chem 33(1):152–157, PMID: 24114755, https://doi.org/10.1002/etc.2405.

Keshava C, Davis JA, Stanek J, Thayer KA, Galizia A, Keshava N, et al. 2020. Application of systematic evidence mapping to assess the impact of new research when updating health reference values: a case example using acrolein. Environ Int 143:105956, PMID: 32702594, https://doi.org/10.1016/j.envint.2020.105956.

Kim D, Andersen ME, Chao YC, Egeghy PP, Rappaport SM, Nylander-French LA. 2007. PBTK modeling demonstrates contribution of dermal and inhalation exposure components to end-exhaled breath concentrations of naphthalene. Environ Health Perspect 115(6):894–901, PMID: 17589597, https://doi.org/10.1289/ehp.9778.

Kim D, Andersen ME, Nylander-French LA. 2006. Dermal absorption and penetration of jet fuel components in humans. Toxicol Lett 165(1):11–21, PMID: 16497449, https://doi.org/10.1016/j.toxlet.2006.01.009.

Kim JH, Kim JK, Son BK, Oh JE, Lim DH, Lee KH, et al. 2005. Effects of air pollutants on childhood asthma. Yonsei Med J 46(2):239–244, PMID: 15861497, https://doi.org/10.3349/ymj.2005.46.2.239.

Kuiri-Hänninen T, Sankilampi U, Dunkel L. 2014. Activation of the hypothalamic-pituitary-gonadal axis in infancy: minipuberty. Horm Res Paediatr 82(2):73–80, PMID: 25012863, https://doi.org/10.1159/000362414.

Lehmann I, Rehwagen M, Diez U, Seiffart A, Rolle-Kampczyk U, Richter M, et al. 2001. Enhanced in vivo IgE production and T cell polarization toward the type 2 phenotype in association with indoor exposure to VOC: results of the LARS study. Int J Hyg Environ Health 204(4):211–221, PMID: 11833293, https://doi.org/10.1078/1438-4639-00100.

Lehmann I, Thoelke A, Rehwagen M, Rolle-Kampczyk U, Schlink U, Schulz R, et al. 2002. The influence of maternal exposure to volatile organic compounds on the cytokine secretion profile of neonatal T cells. Environ Toxicol 17(3):203–210, PMID: 12112628, https://doi.org/10.1002/tox.10055.

Li L, Carratt S, Hartog M, Kovalchik N, Jia K, Wang Y, et al. 2017. Human CYP2A13 and CYP2F1 mediate naphthalene toxicity in the lung and nasal mucosa of CYP2A13/2F1-humanized mice. Environ Health Perspect 125(6):067004, PMID: 28599267, https://doi.org/10.1289/EHP844.

Lin TJ, Guo YL, Hsu JC, Wang IJ. 2018. 2-Naphthol levels and allergic disorders in children. Int J Environ Res Public Health 15(7):1449, PMID: 29987264, https://doi.org/10.3390/ijerph15071449.

Li L, Wei Y, Van Winkle L, Zhang QY, Zhou X, Hu J, et al. 2011. Generation and characterization of a Cyp2f2-null mouse and studies on the role of CYP2F2 in naphthalene-induced toxicity in the lung and nasal olfactory mucosa. J Pharmacol Exp Ther 339(1):62–71, PMID: 21730012, https://doi.org/10.1124/jpet.111.184671.

Long PH, Herbert RA, Peckham JC, Grumbein SL, Shackelford CC, Abdo K. 2003. Morphology of nasal lesions in F344/N rats following chronic inhalation exposure to naphthalene vapors. Toxicol Pathol 31(6):655–664, PMID: 14585734, https://doi.org/10.1080/01926230390242016.

Meeker JD, Barr DB, Serdar B, Rappaport SM, Hauser R. 2007. Utility of urinary 1-naphthol and 2-naphthol levels to assess environmental carbaryl and naphthalene exposure in an epidemiology study. J Expo Sci Environ Epidemiol 17(4):314–320, PMID: 16721410, https://doi.org/10.1038/sj.jes.7500502.

Meeker JD, Ravi SR, Barr DB, Hauser R. 2008. Circulating estradiol in men is inversely related to urinary metabolites of nonpersistent insecticides. Reprod Toxicol 25(2):184–191, PMID: 18249523, https://doi.org/10.1016/j.reprotox.2007.12.005.

Meeker JD, Ryan L, Barr DB, Hauser R. 2006. Exposure to nonpersistent insecticides and male reproductive hormones. Epidemiology 17(1):61–68, PMID: 16357596, https://doi.org/10.1097/01.ede.0000190602.14691.70.

Meeker JD, Ryan L, Barr DB, Herrick RF, Bennett DH, Bravo R, et al. 2004a. The relationship of urinary metabolites of carbaryl/naphthalene and chlorpyrifos with human semen quality. Environ Health Perspect 112(17):1665–1670, PMID: 15579410, https://doi.org/10.1289/ehp.7234.

Meeker JD, Singh NP, Ryan L, Duty SM, Barr DB, Herrick RF, et al. 2004b. Urinary levels of insecticide metabolites and DNA damage in human sperm. Hum Reprod 19(11):2573–2580, PMID: 15333606, https://doi.org/10.1093/humrep/deh444.

Merletti F, Boffetta P, Ferro G, Pisani P, Terracini B. 1991. Occupation and cancer of the oral cavity or oropharynx in Turin, Italy. Scand J Work Environ Health 17(4):248–254, PMID: 1925436, https://doi.org/10.5271/sjweh.1706.

Morris JB. 2013. Nasal dosimetry of inspired naphthalene vapor in the male and female B6C3F1 mouse. Toxicology 309:66–72, PMID: 23619605, https://doi.org/10.1016/j.tox.2013.04.009.

Morris JB, Buckpitt AR. 2009. Upper respiratory tract uptake of naphthalene. Toxicol Sci 111(2):383–391, PMID: 19648534, https://doi.org/10.1093/toxsci/kfp138.

Nagata K, Martin BM, Gillette JR, Sasame HA. 1990. Isozymes of cytochrome P-450 that metabolize naphthalene in liver and lung of untreated mice. Drug Metab Dispos 18(5):557–564, PMID: 1981702.

NASEM (National Academies of Sciences, Engineering, and Medicine). 2018. Progress Toward Transforming the Integrated Risk Information System (IRIS) Program. A 2018 Evaluation (2018). Washington, DC: National Academies Press. http://dx.doi.org/10.17226/25086.

Nie J, Li J, Cheng L, Li Y, Deng Y, Yan Z, et al. 2018. Maternal urinary 2-hydroxynaphthalene and birth outcomes in Taiyuan, China. Environ Health 17(1):91, PMID: 30572877, https://doi.org/10.1186/s12940-018-0436-4.

North DW, Abdo KM, Benson JM, Dahl AR, Morris JB, Renne R, et al. 2008. A review of whole animal bioassays of the carcinogenic potential of naphthalene. Regul Toxicol Pharmacol 51(suppl 2):S6–S14, PMID: 18364246, https://doi.org/10.1016/j.yrtph.2007.09.022.

NRC (National Research Council). 1983. Risk Assessment in the Federal Government: Managing the Process. Washington, DC: National Academy Press. https://doi.org/10.17226/366.

NRC (National Research Council). 2009. Science and Decisions: Advancing Risk Assessment. Washington, DC: National Academies Press. http://dx.doi.org/10.17226/12209.

NTP (National Toxicology Program). 1991. Final Study Report and Appendix: Developmental Toxicity of Naphthalene (CAS No. 91-20-3) in Sprague-Dawley (CD) Rats on Gestational Days 6 through 15. (TER91006). Research Triangle Park, NC: U.S. Department of Commerce National Technical Reports Library. https://ntrl.ntis.gov/NTRL/dashboard/searchResults.xhtml?searchQuery=PB92135623 [accessed 26 June 2021].

NTP. 1992a. Final Report on the Developmental Toxicity of Naphthalene (CAS No. 91-20-3) in New Zealand White Rabbits. Research Triangle Park, NC: National Toxicology Program. https://ntrl.ntis.gov/NTRL/dashboard/searchResults.xhtml?searchQuery=PB92219831 [accessed 26 June 2021].

NTP. 1992b. Toxicology and Carcinogenesis Studies of Naphthalene (CAS No. 91-20-3) in B6C3F1 Mice (Inhalation Studies). (TR 410). Research Triangle Park, NC: U.S. Department of Health and Human Services, Public Health Service. http://ntp.niehs.nih.gov/ntp/htdocs/LT_rpts/tr410.pdf [accessed 26 June 2021].

NTP. 2000. NTP Technical Report on the Toxicology and Carcinogenesis Studies of Naphthalene (CAS No. 91-20-3) in F344/N Rats (Inhalation Studies) (pp.). Research Triangle Park, NC: National Toxicology Program, 1–173. https://ntp.niehs.nih.gov/ntp/htdocs/lt_rpts/tr500.pdf?utm_source=direct&utm_medium=prod&utm_campaign=ntpgolinks&utm_term=tr500 [accessed 26 June 2021].

NTP. 2016. Naphthalene. Report on Carcinogens. 14th ed. Research Triangle Park, NC: National Toxicology Program, https://ntp.niehs.nih.gov/ntp/roc/content/profiles/naphthalene.pdf [accessed 26 June 2021].

OEHHA (California Environmental Protection Agency Office of Environmental Health Hazard Assessment). 2000. Chronic RELs and Toxicity Summaries Using the Previous Version of the Hot Spots Risk Assessment Guidelines, Appendix D3. Sacramento, CA: California Environmental Protection Agency. http://oehha.ca.gov/media/downloads/crnr/appendixd3final.pdf [accessed 26 June 2021].

Pharmakon Research International. 1985. Dose-range-finding–Developmental toxicity study in rabbits using test article 5601-56-1 [TSCA Submission]. (EPA/OTS Doc #86-870000562). Beacon, NY: Texaco. https://ntrl.ntis.gov/NTRL/dashboard/searchResults/titleDetail/OTS0513640.xhtml.

Pharmakon Research International. 1986. Developmental toxicity study in rabbits using test article 5601-56-1 [TSCA Submission]. (EPA/OTS Doc #86-870000563). Beacon, NY: Texaco. https://ntrl.ntis.gov/NTRL/dashboard/searchResults/titleDetail/OTS0513641.xhtml.

Plasterer MR, Bradshaw WS, Booth GM, Carter MW, Schuler RL, Hardin BD. 1985. Developmental toxicity of nine selected compounds following prenatal exposure in the mouse: naphthalene, p-nitrophenol, sodium selenite, dimethyl phthalate, ethylenethiourea, and four glycol ether derivatives. J Toxicol Environ Health 15(1):25–38, PMID: 3981663, https://doi.org/10.1080/15287398509530633.

Quick DJ, Shuler ML. 1999. Use of in vitro data for construction of a physiologically based pharmacokinetic model for naphthalene in rats and mice to probe species differences. Biotechnol Prog 15(3):540–555, PMID: 10356275, https://doi.org/10.1021/bp990057t.

Radke EG, Braun JM, Meeker JD, Cooper GS. 2018. Phthalate exposure and male reproductive outcomes: a systematic review of the human epidemiological evidence. Environ Int 121(pt 1):764–793, PMID: 30336412, https://doi.org/10.1016/j.envint.2018.07.029.

Radke EG, Glenn B, Galizia A, Persad A, Nachman R, Bateson T, et al. 2019. Development of outcome-specific criteria for study evaluation in systematic reviews of epidemiology studies. Environ Int 130:104884, PMID: 31299560, https://doi.org/10.1016/j.envint.2019.05.078.

Rhodes AG, Lemasters GK, Lockey JE, Smith JW, Yiin JH, Egeghy P, et al. 2003. The effects of jet fuel on immune cells of fuel system maintenance workers. J Occup Environ Med 45(1):79–86, PMID: 12553182, https://doi.org/10.1097/00043764-200301000-00016.

Ritter JK, Owens IS, Negishi M, Nagata K, Sheen YY, Gillette JR, et al. 1991. Mouse pulmonary cytochrome P-450 naphthalene hydroxylase: cDNA cloning, sequence, and expression in Saccharomyces cerevisiae. Biochemistry 30(48):11430–11437, PMID: 1742282, https://doi.org/10.1021/bi00112a009.

Rozman K, Summer KH, Rozman T, Greim H. 1982. Elimination of thioethers following administration of naphthalene and diethylmaleate to the rhesus monkey. Drug Chem Toxicol 5(3):265–275, PMID: 7151720, https://doi.org/10.3109/01480548209041057.

Santucci K, Shah B. 2000. Association of naphthalene with acute hemolytic anemia. Acad Emerg Med 7(1):42–47, PMID: 10894241, https://doi.org/10.1111/j.1553-2712.2000.tb01889.x.

Scinicariello F, Buser MC. 2014. Urinary polycyclic aromatic hydrocarbons and childhood obesity: NHANES (2001–2006). Environ Health Perspect 122(3):299–303, PMID: 24380973, https://doi.org/10.1289/ehp.1307234.

Shopp GM, White KL Jr, Holsapple MP, Barnes DW, Duke SS, Anderson AC, et al. 1984. Naphthalene toxicity in CD-1 mice: general toxicology and immunotoxicology. Toxicol Sci 4(3 pt 1):406–419, PMID: 6745531.

Shuler ML, Ghanem A, Quick D, Wong MC, Miller P. 1996. A self-regulating cell culture analog device to mimic animal and human toxicological responses. Biotechnol Bioeng 52(1):45–60, PMID: 18629851, https://doi.org/10.1002/(SICI)1097-0290(19961005)52:1<45::AID-BIT5>3.0.CO;2-Z.

Singh VK, Singh J, Anand M, Kumar P, Patel DK, Krishna Reddy MM, et al. 2008. Comparison of polycyclic aromatic hydrocarbon levels in placental

tissues of Indian women with full- and preterm deliveries. Int J Hyg Environ Health 211(5–6):639–647, PMID: 18308633, https://doi.org/10.1016/j.ijheh.2007.11.004.

Sodeinde O, Chan MC, Maxwell SM, Familusi JB, Hendrickse RG. 1995. Neonatal jaundice, aflatoxins and naphthols: report of a study in Ibadan, Nigeria. Ann Trop Paediatr 15(2):107–113, PMID: 7677410, https://doi.org/10.1080/02724936.1995.11747757.

Sucker K, Zschiesche W, Aziz M, Drews T, Hummel T, Raulf M, et al. 2021. Naphthalene: irritative and inflammatory effects on the airways. Int Arch Occup Environ Health Preprint posted online January 19, 2021, PMID: 33462664, https://doi.org/10.1007/s00420-020-01636-0.

Sudakin DL, Smit E, Cardenas A, Harding A. 2013. Naphthalene biomarkers and relationship with hemoglobin and hematocrit in White, Black, and Hispanic adults: results from the 2003-2004 National Health and Nutrition Examination Survey. J Med Toxicol 9(2):133–138, PMID: 23007805, https://doi.org/10.1007/s13181-012-0262-y.

Summer KH, Rozman K, Coulston F, Greim H. 1979. Urinary excretion of mercapturic acids in chimpanzees and rats. Toxicol Appl Pharmacol 50(2):207–212, PMID: 505452, https://doi.org/10.1016/0041-008X(79)90145-5.

Sweeney LM, Shuler ML, Babish JG, Ghanem A. 1995. A cell culture analog of rodent physiology: application to naphthalene toxicology. Toxicol In Vitro 9(3):307–316, PMID: 20650092, https://doi.org/10.1016/0887-2333(95)00007-U.

Sweeney LM, Shuler ML, Quick DJ, Babish JG. 1996. A preliminary physiologically based pharmacokinetic model for naphthalene and naphthalene oxide in mice and rats. Ann Biomed Eng 24(2):305–320, PMID: 8678360, https://doi.org/10.1007/BF02667357.

U.S. EPA (U.S. Environmental Protection Agency). 1991. Guidelines for Developmental Toxicity Risk Assessment (EPA/600/FR-91/001). Washington, DC: U.S. Environmental Protection Agency Risk Assessment Forum, 1-71. http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=23162 [accessed 26 June 2021].

U.S. EPA. 1996. Guidelines for Reproductive Toxicity Risk Assessment (EPA/630/R-96/009). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum, 1-143. https://www.epa.gov/sites/production/files/2014-11/documents/guidelines_repro_toxicity.pdf [accessed 26 June 2021].

U.S. EPA. 1998. Toxicological Review of Naphthalene (CAS No. 91-20-3) in Support of Summary Information on the Integrated Risk Information System (IRIS). (NTIS/03010276_a). Washington, DC: Integrated Risk Information System. https://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=56434 [accessed 26 June 2021].

U.S. EPA. 2005. Guidelines for Carcinogen Risk Assessment. (EPA/630/P-03/001F). Washington, DC: U.S. Environmental Protection Agency Risk Assessment Forum. https://www.epa.gov/sites/production/files/2013-09/documents/cancer_guidelines_final_3-25-05.pdf [accessed 26 June 2021].

U.S. EPA. 2009. Graphical Arrays of Chemical-Specific Health Effect Reference Values for Inhalation Exposures (2009 Final Report). (EPA/600/R-09/061). Washington, DC: U.S. Environmental Protection Agency. https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=211003 [accessed 26 June 2021].

U.S. EPA. 2016. Chemical Data Reporting (CDR) public database (May 2020 update). https://www.epa.gov/chemical-data-reporting [accessed 26 June 2021].

U.S. EPA. 2018a. EPA: Toxicity Values Version 5 (Aug 2018). https://comptox.epa.gov/dashboard/chemical_lists/TOXVAL_V5 [accessed 17 March 2021].

U.S. EPA. 2018b. IRIS Assessment Plan for Naphthalene (Scoping and Problem Formulation Materials) (EPA/635/R-18/007). Washington, DC: U.S. Environmental Protection Agency. https://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=340791 [accessed 26 June 2021].

U.S. EPA. 2018c. An Umbrella Quality Assurance Project Plan (QAPP) for PBPK Models [EPA Report]. ORD QAPP ID No: B-0030740-QP-1-1. Research Triangle Park, NC: U.S. Environmental Protection Agency. https://hero.epa.gov/hero/index.cfm/reference/details/reference_id/4326432 [accessed 26 June 2021].

U.S. EPA. 2020. Office of Research and Development (ORD) Staff Handbook for Developing IRIS Assessments (public comment draft). (EPA/600/R-20/137). Washington, DC: Center for Public Health and Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency. https://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=350086 [accessed 26 June 2021].

U.S. EPA. 2021. Health and Environmental Research Online (HERO) Database for Naphthalene. https://hero.epa.gov/hero/index.cfm/project/page/project_id/3064 [accessed 15 March 2021].

Viravaidya K, Sin A, Shuler ML. 2004. Development of a microscale cell culture analog to probe naphthalene toxicity. Biotechnol Prog 20(1):316–323, PMID: 14763858, https://doi.org/10.1021/bp0341996.

Waidyanatha S, Troester MA, Lindstrom AB, Rappaport SM. 2002. Measurement of hemoglobin and albumin adducts of naphthalene-1,2-oxide, 1,2-naphthoquinone and 1,4-naphthoquinone after administration of naphthalene to F344 rats. Chem Biol Interact 141(3):189–210, PMID: 12385719, https://doi.org/10.1016/S0009-2797(02)00048-0.

Wang BL, Pang S, Zhang XL, Li X, Sun YG, Lu X, et al. 2014. Levels and neurodevelopmental effects of polycyclic aromatic hydrocarbons in settled house dust of urban dwellings on preschool-aged children in Nanjing, China. Atmos Pollut Res 5(2):292–302, https://doi.org/10.5094/APR.2014.035.

Willems B, Melnick R, Kohn M, Portier C. 2001. A physiologically based pharmacokinetic model for inhalation and intravenous administration of naphthalene in rats and mice. Toxicol Appl Pharmacol 176(2):81–91, PMID: 11601884, https://doi.org/10.1006/taap.2001.9269.

Wolf O. 1976. Krebserkrankungen bei chemiearbeitern einer ehemaligen naphthalinreinigung [Cancer diseases in chemical workers in a former naphthalene cleaning plant]. Dtsch Gesundheitsw 31:996–999.

Wolf O. 1978. Cancer of the larynx in naphthalene cleaners. Z Gesamte Hyg Grenzgeb 24:737–739, PMID: 735200.

Xia Y, Han Y, Zhu P, Wang S, Gu A, Wang L, et al. 2009. Relation between urinary metabolites of polycyclic aromatic hydrocarbons and human semen quality. Environ Sci Technol 43(12):4567–4573, PMID: 19603678, https://doi.org/10.1021/es9000642.

Yang P, Wang YX, Chen YJ, Sun L, Li J, Liu C, et al. 2017. Urinary polycyclic aromatic hydrocarbon metabolites and human semen quality in China. Environ Sci Technol 51(2):958–967, PMID: 27966341, https://doi.org/10.1021/acs.est.6b04810.

Yin S, Tang M, Chen F, Li T, Liu W. 2017. Environmental exposure to polycyclic aromatic hydrocarbons (PAHs): the correlation with and impact on reproductive hormones in umbilical cord serum. Environ Pollut 220(part B):1429–1437, PMID: 27838061, https://doi.org/10.1016/j.envpol.2016.10.090.

Yost EE, Euling SY, Weaver JA, Beverly BEJ, Keshava N, Mudipalli A, et al. 2019. Hazards of diisobutyl phthalate (DIBP) exposure: a systematic review of animal toxicology studies. Environ Int 125:579–594, PMID: 30591249, https://doi.org/10.1016/j.envint.2018.09.038.