Using study evaluation to inform evidence integration: Application in a systematic review of hexavalent chromium male reproductive outcomes

Erin E. Yost[1], Xabier Arzuaga[2], Alan Sasso[2], Catherine Gibbons[2]

1. US Environmental Protection Agency, National Center for Environmental Assessment, Research Triangle Park, NC
2. US Environmental Protection Agency, National Center for Environmental Assessment, Washington, DC

www.epa.gov

Erin E. Yost ❙ yost.erin@epa.gov ❙ 919-541-3906

Background

Study evaluation is used in systematic reviews to identify the strengths and weaknesses of the evidence base in a consistent and transparent manner. These evaluations can be used to inform evidence integration by identifying factors that may affect the reliability and interpretability of the results. Here, we describe how this principle was applied in a systematic review of the male reproductive effects of hexavalent chromium [Cr(VI)].

Methods

Literature search and screening: This evaluation of male reproductive effects was conducted as part of a systematic review of the health effects of Cr(VI) exposure. Studies were identified by searching three online databases (PubMed, Web of Science, Toxline) through May 2018. Title/abstract screening followed by full-text screening was used to identify animal studies meeting the following PECO (**P**opulation, **E**xposure, **C**omparators, **O**utcomes) criteria:

- **P**: Nonhuman mammalian animals (whole organism) of any life stage
- **E**: Any exposure to Cr(VI) by oral or inhalation routes
- **C**: Concurrent vehicle control or untreated control group
- **O**: All cancer outcomes; noncancer outcomes in relevant target systems

The literature search identified 23 animal toxicology studies that examined effects on the male reproductive system. Studies included evaluation of:

- Male fertility
- Sperm parameters
- Reproductive hormones
- Reproductive organ weights
- Anogenital distance (AGD)
- Sexual behavior

Study evaluation: Each of these studies was evaluated by at least two independent reviewers for reporting quality, risk of bias, and sensitivity using the domain-based approach outlined in Figure 1. Based on the results of the evaluation, each study was rated overall as *high* confidence, *medium* confidence, *low* confidence, or *uninformative*. Evaluations were performed on an outcome-specific basis, as the utility of a study may vary across outcomes.

Figure 1. Study evaluation process

Judgment	Interpretation
Good	Appropriate study conduct relating to the domain & minor deficiencies not expected to influence results.
Adequate	A study that may have some limitations relating to the domain, but they are not likely to be severe or to have a notable impact on results.
Deficient	Identified biases or deficiencies interpreted as likely to have had a notable impact on the results or prevent reliable interpretation of study findings.
Critically Deficient	A serious flaw identified that is interpreted to be the primary driver of any observed effect or makes the study uninterpretable. Study is not used without exceptional justification.

Animal Study Evaluation Domains

- Reporting Quality
- Selection or Performance Bias
  - Allocation
  - Observational Bias / Blinding
- Confounding/Variable Control
- Selective Reporting and Attrition
- Exposure Methods Sensitivity
  - Chemical Administration and Characterization
  - Exposure Timing, Frequency, and Duration
- Outcome Measures and Results Display
  - Endpoint Sensitivity and Specificity
  - Results Presentation

Overall study rating

Rating	Interpretation
High	No notable deficiencies or concerns identified; potential for bias unlikely or minimal; sensitive methodology.
Medium	Possible deficiencies or concerns noted, but resulting bias or lack of sensitivity would be unlikely to be of a notable degree.
Low	Deficiencies or concerns were noted, and the potential for substantive bias or inadequate sensitivity could have a significant impact on the study results or their interpretation.
Uninformative	Serious flaw(s) makes study results unusable for hazard identification

Evidence synthesis: Evidence was synthesized across studies, using the following considerations to articulate the strengths and weaknesses of the dataset: consistency, biological gradient (dose-response), strength (effect magnitude) and precision, biological plausibility, and coherence. Careful examination was given to the potential impacts of risk of bias and sensitivity on the conclusions. Relevant mechanistic data identified in the literature search was considered as part of the weight of evidence for biological plausibility. Based on this synthesis, the evidence was assigned a conclusion of *robust*, *moderate*, *slight*, *indeterminate*, or *compelling evidence of no effect*.

Results

Table 1. Study evaluation results. These results represent the composite ratings for male reproductive outcomes within each evaluation domain; there were some instances where outcomes within the same study were rated differently due to outcome-specific concerns, in which case an average rating (representative of most outcomes) is shown here. In addition to the 15 studies shown in this table, 8 studies were considered *uninformative* due to serious flaws in the study design (e.g., use of wild-caught animals) or reporting (e.g., data could not be interpreted) and were excluded from consideration.

Reference	Study description			Study evaluation										
	Species/ Strain	Exposure life stage and duration	Route of exposure	Reporting quality	Test animal allocation	Blinding of investigators	Confounding / variable control	Reporting or attrition bias	Characterization of exposure	Utility of exposure design	Sensitivity, specificity, and usability of results	Presentation of results	Overall confidence
NTP 1997	Mouse (BALBC)	Reproductive Assessment by Continuous Breeding	Diet	G	G	NR	G	G	G	G	G	G	High
NTP 1996a	Mouse (BALBC)	Adult males and females; 3, 6, or 9 weeks	Diet	G	G	G	G	A	G	G	G	G	High
NTP 1996b	Rat (Sprague-Dawley)	Adult males and females; 3, 6, or 9 weeks	Diet	G	G	G	G	G	G	G	G	A	High
NTP 2007	Rat (F344/N), Mouse (B6C3F1, BALB/c, C57BL-6)	Adult males and females; 3 months	Drinking water	G	G	A	G	A	G	G	G	G	High
Elbetieha et al. 1997	Mouse (Swiss)	Adult males and females; 12 weeks	Drinking water	G	NR	NR	D	A	D	G	G	A	Low
Bataineh et al. 1997	Rat (Sprague-Dawley)	Adult males; 12 weeks	Drinking water	D	A	NR	D	D	D	G	A	G	Low
Yousef et al. 2006	Rabbit (NZ white)	Adult males; 10 weeks	Oral gavage	G	A	NR	D	D	D	G	G	D	Low
Li et al. 2001	Rat (Wistar)	Adult males; 6 weeks	Oral feeding	D	NR	NR	D	A	D	A	A	A	Low
Rasool et al. 2014	Mouse (strain not reported)	Adult males; 30 or 60 days	Oral (unspecified)	D	A	NR	D	A	D	G	D	D	Low
Wang et al. 2015	Rat (Sprague-Dawley)	Adult males; 4 weeks	Drinking water	G	A	NR	G	G	A	G	D	A	Low
Kumar et al. 2017	Rat (Wistar)	F1 offspring; GD 9–14	Drinking water	A	NR	NR	G	D	D	A	A	D	Low
Al-Hamood et al. 1998	Mouse (BALBC)	F1 offspring; GD 12–PND 20	Drinking water	G	D	NR	D	D	D	G	G	D	Low
Glaser et al. 1986	Rat (Wistar)	Adult males; 18 months	Inhalation	A	A	NR	G	A	A	A	D	A	Low
Glaser et al. 1985	Rat (Wistar)	Adult males; 28 or 90 days	Inhalation	A	A	NR	G	A	A	G	D	A	Low
Kim et al. 2004	Rat (Sprague-Dawley)	Adult males; 90 days	Inhalation	G	A	NR	G	A	A	A	A	D	Low

Abbreviations: Gestation day (GD); Postnatal day (PND); Good (G); Adequate (A); Deficient (D); Not Reported (NR)

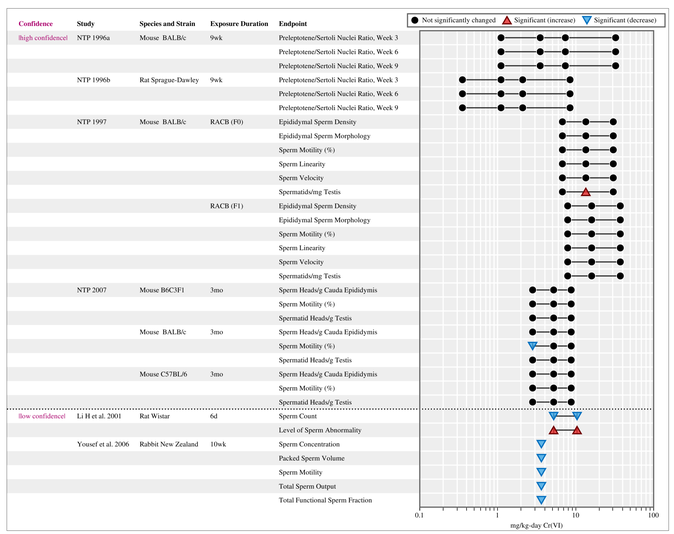

Figure 2. Summary of effects on sperm. Data is shown for all studies for which the ingested dose of Cr(VI) could be calculated. Decreased sperm count, mobility, and viability were also observed in the *low confidence* study by Kumar et al. 2017, but the ingested dose of Cr(VI) could not be calculated based on the reported information.

Summary of Effects:

- *High confidence* subchronic oral exposure studies in rats and mice (NTP 1996a, 1996b, 2007) and a continuous breeding study in mice (NTP 1997) generally indicated that the male reproductive system is not affected by Cr(VI) exposure.
- *Low confidence* oral exposure studies consistently observed effects on sperm quality and quantity, testicular histopathology, male reproductive organ weights, hormone levels, sexual behavior, and AGD.
- As an example, Figure 2 summarizes effects on sperm parameters across studies.
- Biological plausibility for male reproductive effects of Cr(VI) exposure was supported by mechanistic studies (in vivo and in vitro) demonstrating oxidative stress and apoptosis in male reproductive tissues, altered steroidogenic signaling, disruption of the blood-testis barrier, and alterations in meiosis.
- No effects were observed in three *low confidence* inhalation studies.

Summary of effects in high vs. low confidence studies

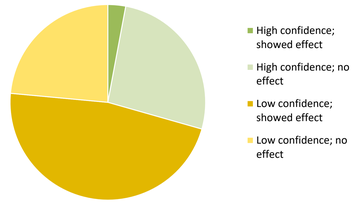

Figure 3. Incidence of outcomes indicative of male reproductive effects across *high* and *low confidence* Cr(VI) animal studies. One *high* confidence study observed decreased testis weight, but otherwise *high* confidence studies found no evidence of male reproductive effects. Comparative, male reproductive effects was frequently observed in *low* confidence studies.

Integration of evidence

It was concluded that animal toxicology studies along with supportive data from mechanistic studies provide *slight* evidence that Cr(VI) is a male reproductive toxicant. The rationale for this conclusion is documented in an evidence profile table (Table 2). Relatively severe male reproductive effects were observed across multiple *low* confidence studies and are supposed by mechanistic evidence. However, similar effects were not observed in *high* confidence studies, and concerns were raised about the potential impact of bias on the interpretation of the results in *low* confidence studies. Fertility (ability to produce offspring) was not affected in any studies but this did not affect overall conclusions, since rodents can remain fertile after large reductions in sperm count.

Table 2: Evidence profile table for Cr(VI) male reproductive effects

	Studies [confidence]	Factors that increase strength	Factors that decrease strength	Summary of findings
Fertility and Fecundity	NTP 1997 [high] Bataineh et al. 1997 [low] Elbetieha et al. 1997 [low] Al-Hamood et al. 1998 [low]		• Only study that observed an effect is considered *low* confidence • No effects observed in *high* confidence studies	No effects on male fertility (ability to produce offspring) were observed across studies in rats or mice, although one *low confidence* study in rats observed decreased fetal viability following paternal exposure (Elbetieha et al. 1997).
Sperm evaluation	NTP 1996a [high] NTP 1996b [high] NTP 1997 [high] NTP 2007 [high] Kumar et al. 2017 [low] Li et al. 2001 [low] Yousef et al. 2006 [low]	• Dose-response gradient • Biological plausibility (mechanistic evidence of oxidative stress, effects on blood-testis barrier, and altered meiosis)	• No effects observed in *high* confidence studies • Studies that observed effects were all considered *low* confidence	No effects on sperm parameters were observed in *high* confidence studies in rats or mice, whereas *low* confidence studies in rats and rabbits reported decreased sperm quality and quantity.
Histopathology	NTP 2007 [medium] Kumar et al. 2017 [low] Li et al. 2001 [low] Rasool et al. 2012 [low] Wang et al. 2015 [low] Glaser et al. 1985 [low] Kim et al. 2004 [low]	• Coherence with effects on sperm • Dose-response gradient • Biological plausibility (mechanistic evidence of oxidative stress and effects on blood-testis barrier)	• No effects observed in *high* confidence studies • Studies that observed effects were all considered *low* confidence	No histopathological effects were reported in the *high confidence* study in rats and a variety of mouse strains by NTP 2007, whereas three *low* confidence studies in rats and mice observed histopathological changes in the testis and seminiferous tubules.
Organ weight	NTP 1996a [high] NTP 1996b [high] NTP 1997 [high] NTP 2007 [high] Al-Hamood et al. 1998 [low] Bataineh et al. 1997 [low] Elbetieha et al. 1997[low] Kumar et al. 2017 [low] Yousef et al. 2006 [low] Wang et al. 2015 [low] Kim et al. 2004 [low] Glaser et al. 1986 [low]	• Coherence with decreased testosterone	• Unexplained inconsistency • Most studies that observed effects were considered *low* confidence	Decreased testis weight was observed in one out of three mouse strains in the *high confidence* study by NTP 2007, and decreased testis and accessory male reproductive organ weights were observed in four *low confidence* studies in rabbits (Yousef et al. 2006), rats (Bataineh et al. 1997, Kumar et al. 2017), and mice (Elbetieha et al. 1997). No effects were observed in the remaining 7 studies.
Hormones	Yousef et al. 2006 [low] Kumar et al. 2017 [low]	• Consistency • Biological plausibility (mechanistic evidence of decreased steroidogenesis)	• Few studies • Only *low* confidence studies available	Decreased testosterone was observed in rabbits exposed as adults, and decreased testosterone and gonadotropins were observed in F1 rats that had been exposed during gestation.
Sexual behavior	Bataineh et al. 1997 [low] Yousef et al. 2006 [low]	• Consistency	• Few studies • Only *low* confidence studies available	Decreased mounts, increased ejaculation latency and post-ejaculation interval, and decreased percentage of males ejaculating were observed in rats exposed as adults. Increased reaction time to mounting was observed in rabbits.
AGD	Kumar et al. 2017 [low]	• Coherence with decreased testosterone	• Single study • *Low* confidence	Decreased AGD was observed in F1 rats, which is consistent with the observation of decreased testosterone in these animals.

Disclaimer: The views expressed are those of the authors and do not necessarily represent the views/policies of the US EPA. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

U.S. Environmental Protection Agency
Office of Research and Development