

HSIA

halogenated
solvents
industry
alliance, inc.

October 6, 2015

Information Quality Guidelines Staff
Mail Code 2811R
Environmental Protection Agency
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460

Re: Request for Correction – TSCA Work Plan Chemical Risk Assessment
Trichloroethylene: Degreasing, Spot Cleaning and Arts & Craft Uses (June 2014)
(#740-R1-4002)

Dear Sir or Madam:

This request for the correction of information (“Request for Correction”) is submitted under the Information Quality Act (“IQA”)¹ and the implementing guidelines issued, respectively, by the Office of Management & Budget (“OMB”)² and the Environmental Protection Agency (“EPA”),³ on behalf of the Halogenated Solvents Industry Alliance, Inc. (“HSIA”). HSIA represents producers of trichloroethylene (“TCE”) and other chlorinated solvents. As discussed below, HSIA seeks the correction of information disseminated in an EPA document, TSCA Work Plan Chemical Risk Assessment for Trichloroethylene: Degreasing, Spot Cleaning and Arts & Craft Uses (June 2014) (#740-R1-4002), issued by the Office of Chemical Safety and Pollution Prevention (“Work Plan Assessment”).

This Request for Correction is separate and distinct from the request for correction filed on November 5, 2013, and denied by EPA on March 19, 2015 (#14001). That request was for correction of information disseminated in an EPA document, “Toxicological Review of Trichloroethylene (CAS No. 79-01-6) in Support of Summary Information on the Integrated Risk Information System (IRIS)” (“IRIS Assessment”). A request for reconsideration of that request was submitted on June 17, 2015 and is pending before EPA.

¹ Section 515(a) of the Treasury and General Government Appropriations Act for Fiscal Year 2001, P.L. 106-554; 44 U.S.C. § 3516 (notes).

² 67 Fed. Reg. 8452 (Feb. 22, 2002) (“OMB Guidelines”).

³ EPA, Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency, EPA/260R-02-008 (October 2002).

Our earlier request for correction addressed in detail the deficiencies of the IRIS Assessment. The IRIS Assessment contains a reference concentration (“RfC”) of 0.0004 ppm (0.4 ppb or 2 $\mu\text{g}/\text{m}^3$) and a reference dose (“RfD”) of 0.0005 mg/kg/day for TCE. These are values that are considered by EPA to be protective for all of the candidate critical effects. EPA’s derivation of the RfC/RfD for TCE is based, in part, on Johnson *et al.*, Threshold of Trichloroethylene Contamination in Maternal Drinking Waters Affecting Fetal Heart Development in the Rat, Environmental Health Perspectives 111: 289-92 (2003).

The Work Plan Assessment goes beyond the IRIS Assessment by expressly relying on hazard values derived directly from Johnson *et al.* (2003) to estimate acute risk:⁴ Specifically, the Work Plan Assessment states (p. 104):

“The acute inhalation risk assessment used developmental toxicity data to evaluate the acute risks for the TSCA TCE use scenarios. As indicated previously, EPA’s policy supports the use of developmental studies to evaluate the risks of acute exposures. This policy is based on the presumption that a single exposure of a chemical at a critical window of fetal development, as in the case of cardiac development, may produce adverse developmental effects (EPA, 1991).

“After evaluating the developmental toxicity literature of TCE, the TCE IRIS assessment concluded that the fetal heart malformations are the most sensitive developmental toxicity endpoint associated with TCE exposure (EPA, 2011e). Thus, EPA/OPPT based its acute risk assessment on the most health protective endpoint (i.e., fetal cardiac malformations; Johnson et al., 2003) representing the most sensitive human population (i.e., adult women of childbearing age and fetus >16 yrs).

“The acute risk assessment used the PBPK-derived hazard values (HEC_{50} , HEC_{95} , or HEC_{99}) from Johnson et al. (2003) developmental study for each degreaser and spot cleaner use scenario. Note that the variability among these hazard values is small and no greater than 3-fold (i.e., 2-fold for $\text{HEC}_{50}/\text{HEC}_{95}$ ratios; 3-fold for $\text{HEC}_{50}/\text{HEC}_{99}$ ratios; 1.4-fold for $\text{HEC}_{95}/\text{HEC}_{99}$ ratios).”

⁴ This is one of several reasons that EPA’s denial of the earlier request for correction is inapposite here. In that denial, EPA stated that the HSIA request “could be interpreted to assert that EPA’s Toxicological Review of TCE relies exclusively on a single study to support the derivation of reference values and this is factually incorrect. Rather, in developing the Toxicological Review of TCE, EPA reviewed more than one hundred toxicological studies to evaluate TCE hazards, including dozens of developmental toxicity studies. . . . The numerical values of the Reference Dose and Reference Concentration were also based on multiple studies.”

These extremely low values result in margin of exposure (“MOE”) values below 10 for almost all the occupational and residential exposure scenarios examined (p. 104):

“Acute inhalation risks were reported for most occupational and residential exposure scenarios based on concerns for developmental effects, irrespective of who is using the product (user vs. bystander), the type of exposure (typical vs. worst case scenario) and the room ventilation system (LEV vs no LEV). For instance, most of the degreaser and spot cleaner exposure scenarios and all of the residential use scenarios reported MOE values below the benchmark MOE of 10 irrespective of the percentile HEC value used to estimate the MOEs.”

I. Deficiencies in Johnson *et al.* (2003)

A single flawed study should not be the basis for the toxicological value that is expected to serve as the basis for regulation. Several other studies, including two GLP-compliant studies conducted under EPA guidelines to support pesticide registration (40 CFR § 870.3700) and Organization for Economic Coordination & Development (“OECD”) guidelines (414), one of which was sponsored by HSIA under a voluntary testing agreement with the Agency for Toxic Substances & Disease Registry (“ATSDR”), have been unable to reproduce the effect seen by Johnson *et al.* (2003), as described below.

Johnson *et al.* (2003) reported cardiac effects in rats from research carried out at the University of Arizona and originally published ten years earlier by the same authors.⁵ In the earlier-published study, there was no difference in the percentage of cardiac abnormalities in rats dosed during both pre-mating and pregnancy at drinking water exposures of 1100 ppm (9.2%) and 1.5 ppm (8.2%), even though there was a 733-fold difference in the concentrations. The authors reported that the effects seen at these exposures were statistically higher than the percent abnormalities in controls (3%). For animals dosed only during the pregnancy period, the abnormalities in rats dosed at 1100 ppm (10.4%) were statistically higher than at 1.5 ppm (5.5%), but those dosed at 1.5 ppm were not statistically different from the controls. Thus, no meaningful dose-response relationship was observed in either treatment group. Johnson *et al.* republished in 2003 data from the 1.5 and 1100 ppm dose groups published by Dawson *et al.* in 1993 and pooled control data from other studies, an inappropriate statistical practice, to conclude that rats exposed to levels of TCE greater than 250 ppb during pregnancy have increased incidences of cardiac malformations in their fetuses.

⁵ Dawson, B, *et al.*, Cardiac teratogenesis of halogenated hydrocarbon-contaminated drinking water, *J. Am. Coll. Cardiol.* 21: 1466-72 (1993).

Johnson *et al.* (2003) has been heavily criticized in the published literature.⁶ Indeed, its predecessor study was expressly rejected as the basis for MRLs by ATSDR in its last TCE Toxicological Profile Update.⁷ Moreover, the Johnson *et al.* (2003) findings were not reproduced in a study designed to detect cardiac malformations; this despite employing an improved method for assessing cardiac defects and the participation of Dr. Johnson herself.⁸ No increase in cardiac malformations was observed in the guideline study sponsored by HSIA,⁹ despite high inhalation doses and techniques capable of detecting most of the malformation types reported by Johnson *et al.* (2003). The dose-response relationship reported in Johnson *et al.* (2003) for doses spanning an extreme range of experimental dose levels is considered by many to be improbable, and has not been replicated by any other laboratory.¹⁰

We are hard pressed to find a better summary of Johnson *et al.* (2003) than the following statement by the California Office of Environmental Health Hazard Assessment (OEHHA) rejecting the study as deficient:

"Johnson et al. (2003) reported a dose-related increased incidence of abnormal hearts in offspring of Sprague Dawley rats treated during pregnancy with 0, 2.5 ppb, 250 ppb, 1.5 ppm, and 1,100 ppm TCE in drinking water (0, 0.00045, 0.048, 0.218, and 128.52 mg/kg-day, respectively). The NOAEL for the Johnson study was reported to be 2.5 ppb (0.00045 mg/kg-day) in this short exposure (22 days) study. The percentage of abnormal hearts in the control group was 2.2 percent, and in the treated groups was 0 percent (low dose), 4.5 percent (mid dose 1), 5.0 percent (mid dose 2), and 10.5 percent (high dose). The number of litters with fetuses with abnormal hearts was 16.4 percent, 0 percent, 44 percent, 38 percent, and 67 percent for the control,

⁶ Hardin, B, *et al.*, Trichloroethylene and cardiac malformations, *Environ. Health Perspect.* 112: A607-8 (2004); Watson, R., *et al.*, Trichloroethylene-contaminated drinking water and congenital heart defects: a critical analysis of the literature, *Repro. Toxicol.* 21: 117-47 (2006).

⁷ ATSDR concluded that "[t]he study is limited in that only two widely spaced exposure concentrations were used and that a significant dose-response was not observed for several exposure scenarios." *Toxicological Profile for Trichloroethylene Update* (September 1997), at 88.

⁸ Fisher, J, *et al.*, Trichloroethylene, trichloroacetic acid, and dichloroacetic acid: do they affect fetal rat heart development? *Int. J. Toxicol.* 20: 257-67 (2001).

⁹ Carney, E, *et al.*, Developmental toxicity studies in Crl:CD (SD) rats following inhalation exposure to trichloroethylene and perchloroethylene, *Birth Defects Research (Part B)* 77: 405-412 (2006).

¹⁰ "Johnson and Dawson, with their collaborators, are alone in reporting that TCE is a 'specific' cardiac teratogen." Hardin, B, *et al.*, Trichloroethylene and cardiac malformations, *Environ. Health Perspect.* 112: A607-8 (2004).

low, mid 1, mid 2, and high dose, respectively. The reported NOAEL is separated by 100-fold from the next higher dose level. The data for this study were not used to calculate a public-health protective concentration since a meaningful or interpretable dose-response relationship was not observed. *These results are also not consistent with earlier developmental and reproductive toxicological studies done outside this lab in mice, rats, and rabbits: The other studies did not find adverse effects on fertility or embryonic development, aside from those associated with maternal toxicity (Hardin et al., 2004).*¹¹

Remarkably, an EPA staff review that was recently placed in the docket for the Work Plan Assessment, and then cited in the denial of HSIA's original request for correction, reflects similar concerns. First, one staff member dissented over relying at all on the Arizona study:

“The rodent developmental toxicology studies conducted by Dawson et al. (1993), Johnson et al. (2003), and Johnson et al. (1998) that have reported cardiac defects resulting from TCE (and metabolite) drinking water exposures have study design and reporting limitations. Additionally, two good quality (GLP) inhalation and gavage rodent studies conducted in other laboratories, Carney et al. (2006) and Fisher et al. (2001), respectively, have not detected cardiac defects. These limitations and uncertainties were the basis of the single dissenting opinion of a team member regarding whether the database supports a conclusion that TCE exposures during development are likely to cause cardiac defects.”¹²

Second, even the EPA staff that agreed with use of the study had little confidence that it supported the dose-response assessment:

“[A] majority of the team members agreed that the Johnson et al. (2003) study was suitable for use in deriving a point of departure. However, confidence of team members in the dose response evaluation of the cardiac defect data from the Johnson et al. (2003) study was characterized as between “low” and “medium” (with 7 of 11

¹¹ California EPA Public Health Goal for Trichloroethylene in Drinking Water (July 2009), at 21 (emphasis added).

¹² TCE Developmental Cardiac Toxicity Assessment Update (available at <http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPPT-2012-0723-0045>).

team members rating confidence as “low” and four team members rating confidence as “low to medium”).”¹³

To provide further validation of its reliance on Johnson *et al.* (2003), an Appendix N (Weight-of-Evidence Analysis for Fetal Cardiac Malformation), excerpted from the staff review, was added to the final Work Plan Assessment. This analysis was not provided with the External Review Draft and thus was not available for evaluation by the peer review panel. In its written response to Peer Review Comment 44, however, EPA stated:

“A recent erratum (Johnson, 2014) and subsequent evaluation of the developmental toxicity data reaffirmed that the Johnson et al. studies are adequate to use in hazard identification and dose-response assessment (Appendix N). As explained in the TCE IRIS assessment, while the Johnson et al. studies have limitations, there is insufficient reason to dismiss their findings, especially when the findings are analyzed in combination with human, animal and mechanistic evidence. The comprehensive WOE evaluation of the developmental toxicity data, including fetal cardiac teratogenesis, is discussed in the TCE IRIS assessment and expanded in this assessment (Appendix N).”

Appendix N presents a weight-of-evidence analysis for the association between short-term exposure to TCE and fetal cardiac defects and classifies available information as being evidence of either a stronger or weaker weight of association. Unfortunately, there are several instances where the ‘evidence’ presented is contradictory or just incorrect. Several of these are presented below:

“Evidence for Stronger Weight of Association:

“Evidence

“The power of detection in the Johnson et al. (2003) study was enhanced by the use of historical controls that did not demonstrate a temporal shift in cardiac defects. A significant dose related trend in cardiac defects was observed even without large group sizes.

* * * * *

“The Johnson et al. (2003) study reported data from several cohorts of animals, which were on study over a period of approximately 6 years. The data included

¹³ *Id.*

control cohorts, some of which were concurrent and some that were non-concurrent to the TCE treated groups (Johnson et al., 2005; Johnson, 2014).

* * * * *

“Johnson et al. (2003) reported that cardiac defect incidences were consistent across all control cohorts (55 litters over approximately 6 years). An EPA review of the available control data did not observe unusual heterogeneity in prevalence of malformations.

* * * * *

“It is unlikely that the cardiac defects observed by Johnson et al. (2003) were an artifact of the evaluation procedures used, since a study by Fisher et al. (2001), using the same fetal cardiac evaluation procedures, did not identify an association between TCE exposure and the incidence of cardiac defects.

“Evidence for Weaker Weight of Association:

“The Dawson et al. (1993) and Johnson et al. (2003) studies estimated doses based on the average water consumption. This method does not provide precise information to calculate TCE dose because variability in drinking water consumption among dams is not characterized.

* * * * *

“Some studies that did not identify treatment-related cardiac defects following developmental exposures to TCE (e.g., Carney et al., 2006; Fisher et al., 2001; Schwetz et al., 1975) were well-conducted and adequately-reported GLP and/or guideline studies with no substantive limitations identified.

Page 121 of final Work Plan Assessment:

“The potential hazard for congenital malformations is supported by a weight of evidence analysis of the *weakly suggestive epidemiological data* [emphasis added] in combination with the findings of the animal and mechanistic studies with TCE and its metabolites (EPA, 2011e). The robustness of the weight of evidence analysis gives greater confidence to the hazard conclusions for fetal cardiac defects (Appendix N).”

It is surprising that EPA would stretch to use for a dose-response value intended to serve as the basis for regulation a study in which seven of its own scientists expressed “low” confidence, and in which the other four could muster no more than “low to medium” confidence. The same report

notes: “In conclusion, there has not been a confirmation of the results of the Johnson et al. (2003) and Dawson et al. (1993) studies by another laboratory, but there has also not been a repeat of the exact same study design that would corroborate or refute their findings.”¹⁴ For over two years now, HSIA has had outstanding to EPA and ATSDR (with whom HSIA has a voluntary research agreement in place) an offer to sponsor just such a study so that the issue of reproducibility, at least, could be resolved. All we ask of EPA is that it participate in the study design so that all stakeholders can be assured that the protocol is scientifically appropriate.

A. No Other Rodent Study Replicates the Arizona Findings

TCE has been associated with cardiac malformations only in animal studies conducted at the University of Arizona laboratory. In this drinking water study (separately reported by Dawson *et al.* (1993) and Johnson *et al.* (2003)), TCE was reported to produce cardiac teratogenicity and no other adverse developmental effects.¹⁵ No other laboratory has been able to reproduce these results. In several well designed and conducted studies using standard techniques for identifying developmental hazards, rats, mice, and rabbits were exposed to TCE by inhalation at doses as high as 1800 ppm¹⁶ and rats have been exposed by oral gavage to 500 mg/kg/day of TCE.¹⁷ None of these studies reported exposure-related developmental toxicity, even in the presence of maternal toxicity. Furthermore, none of these studies reported evidence of specific cardiac teratogenicity, even when the micro-dissection technique of the one laboratory reporting cardiac anomalies was used and a member of that research group was part of the study team.¹⁸

¹⁴ *Id.*

¹⁵ The absence of data on maternal and fetal parameters other than cardiac malformations makes it impossible to assess the impact of exposure on critical factors like maternal body weight and fetal weight and development.

¹⁶ Schwetz, BA, *et al.*, The effect of maternally inhaled trichloroethylene, perchloroethylene, methyl chloroform, and methylene chloride on embryonal and fetal development in mice and rats, *Toxicol Appl Pharmacol* 32(1):84–96 (1975); Dorfmueller, MA, *et al.*, Evaluation of teratogenicity and behavioral toxicity with inhalation exposure of maternal rats to trichloroethylene, *Toxicology* 14(2):153–166 (1979); Beliles, RP, *et al.*, Teratogenic-mutagenic risk of workplace contaminants—trichloroethylene, perchloroethylene, and carbon disulfide, Final report, Contract 210-77-0047, NTIS PB 82-185075, National Institute for Occupational Safety and Health (NIOSH), Cincinnati, OH (1980); Healy, TE, *et al.*, Rat fetal development and maternal exposure to trichloroethylene 100 ppm, *Br J Anaesth* 54(3):337–341 (1982); Carney, E, *et al.*, Developmental toxicity studies in CrI:Cd (SD) rats following inhalation exposure to trichloroethylene and perchloroethylene, *Birth Defects Research (Part B)* 77: 405–412 (2006).

¹⁷ Fisher, J, *et al.*, Trichloroethylene, trichloroacetic acid, and dichloroacetic acid: do they affect fetal rat heart development?, *Int. J. Toxicol.* 20: 257-67 (2001).

¹⁸ *Id.*

Several possible explanations have been suggested for the positive findings of an association between TCE exposure and cardiac malformations in rats by Dawson *et al.* (1993) and Johnson *et al.* (2003) – higher TCE concentrations, the mode and timing of exposure, differences in detection techniques, or the use of non-standard statistical evaluations.¹⁹ The findings cannot be explained by the high concentrations included in the Arizona study (129 mg/kg/day) since Fisher *et al.* (2001) dosed dams at 500 mg/kg/day TCE by gavage. Exposure of the dams throughout pregnancy in the Arizona study, rather than limiting exposure to the most sensitive period of organogenesis (gestation days 6 through 15 or GD 6-15) as in Fisher *et al.* (2001), similarly is unlikely to explain the difference since the heart is formed during GD 6-15 and any exposure before or after this period would not increase cardiac anomalies. Other studies in which dams were exposed for all or most of the pregnancy also failed to observe an increase in cardiac effects.

Although only the Arizona study investigated cardiac anomalies in rats exposed to TCE through drinking water, the difference in route of exposure cannot explain the positive results reported. Oral gavage studies, such as that conducted by Fisher *et al.* (2001), will result in higher blood concentrations than those using drinking water exposures.²⁰ Consequently, a gavage study would be more likely to cause developmental effects than a drinking water study.

While the Arizona researchers have suggested that the dissection technique used in their studies may be more sensitive in detecting certain lesions, those lesions are not the predominant cardiac anomalies they reported. Perhaps more significantly, Fisher *et al.* (2001) failed to observe an increase in these effects despite collaborating with Dr. Johnson and using the same dissection method.

B. Inappropriate Statistical Practices Preclude Reliance on Johnson *et al.* (2003)

One of the principal criticisms of Johnson *et al.* (2003) is that it employed inappropriate statistical practices:

“Johnson *et al.* (2003) provided no rationale for designing their study with a concurrent control five times larger than the treatment groups, which leads us to ask whether the control group reported here is, in fact, a composite of controls from multiple, perhaps five, different studies.. The immediate impact of this large control group is that the very

¹⁹ Watson, RE, *et al.*, Trichloroethylene-contaminated drinking water and congenital heart defects: a critical analysis of the literature, *Reprod Toxicol* 21(2):117–147 (2006).

²⁰ *Id.*

cardiac ‘abnormalities’ at the 1.5 ppm dose that did not differ significantly from controls in 1993 become statistically significant in 2003.’²¹

It appears that one significant reason for the positive results reported at the University of Arizona is that the statistics were performed differently than in traditional developmental studies. Original statistics were performed on a per-fetus basis, rather than on a per-litter basis, despite the fact that per-litter analysis is the accepted method for developmental effects related to chemical exposure during pregnancy, as recommended by EPA.²² Statistics should be conducted on a per-litter basis because, during gestation, the dam is the unit of treatment and exposure of the pups is dependent on her.²³ Performing statistics in a per-fetus manner artificially inflates the significance of the findings. Had the correct statistical unit been used in these studies, a positive correlation between TCE and cardiac anomalies probably would not have been reported in the original drinking water studies.²⁴

Johnson *et al.* (2003) re-published data from Dawson *et al.* (1993) using pooled controls in the statistical evaluation.²⁵ Pooling of controls is not an appropriate statistical practice, however,

²¹ Hardin, B, et al., Trichloroethylene and cardiac malformations, *Environ. Health Perspect.* 112: A607-8 (2004).

²² EPA Risk Assessment Forum, Guidelines for Developmental Toxicity Risk Assessment (EPA/600/FR-91/001) (1991) (available at http://www2.epa.gov/sites/production/files/2014-11/documents/dev_tox.pdf).

²³ Haseman, JK and Hogan, MD, Selection of the experimental unit in teratology studies, *Teratology* 12:165–72 (1975).

²⁴ Watson, RE, *et al.*, Trichloroethylene-contaminated drinking water and congenital heart defects: a critical analysis of the literature, *Reprod Toxicol* 21(2):117–147 (2006).

²⁵ In Dawson *et al.* (1993), water containing 1.5 or 1100 ppm TCE was provided before conception, during pregnancy, or both before and during pregnancy. With drinking water exposure before conception, no impact on mating success or intrauterine survival was observed, and pre-gestational exposure alone had no influence on heart defects. The number of fetuses reported with “abnormal hearts” was significantly increased for dams exposed to 1.5 ppm TCE before conception and during pregnancy, but not for those exposed at that level only during pregnancy. Abnormal hearts were significantly increased in fetuses in both groups (before/during conception, during conception only) of dams exposed to 1100 ppm TCE. In Johnson *et al.* (2003), rats were given drinking water containing TCE at 2.5 ppb (equal to 0.0025 ppm), 250 ppb (0.25 ppm), 1.5 ppm, or 1100 ppm. Cardiac defects were reported to be significantly increased at the 0.25 ppm and 1100 ppm exposure concentrations, but not at intermediate level of 1.5 ppm. Further evaluation by Hardin *et al.* (2004) confirmed that results from the two higher doses were the same as those reported in 1993 and that the control group represented a combination of an unspecified number of historic controls rather than one control group run concurrently with the two lower exposed groups, as asserted in Johnson, PD, *et al.*, Trichloroethylene: Johnson *et al.*’s Response [Letter], *Environ Health Perspect* 112(11):A608–A609 (2004). In the absence of a clear dose-response relationship, it is difficult to conclude that the observed effects were treatment-related.

and is likely to have exaggerated the alleged statistical significance.²⁶ Fisher *et al.* (2001), moreover, express concern that “[t]he high background of fetal heart malformations on a per litter basis provides a challenge for using these data in regulatory decisions relating to risk characterization of TCE, TCA, and DCA.”

Fisher *et al.* (2001) also note that the lack of clear dose-related effects in the study by Dawson *et al.* (1993) and their own study provide “data of questionable utility for risk assessment applications.” In an attempt to provide support for a dose-response, Johnson *et al.* (2003) present a dose-response curve, based on a probit analysis, at concentrations up to 4878 ppm. The concentration of 4878 ppm is well above water solubility for TCE, however, and the authors fail to explain how they could generate a curve using concentrations for which no data exist.

Finally, in a recent erratum,²⁷ Johnson and co-authors note that the dates listed for conduct of the “2.5-ppb and 250-ppb trichloroethylene (TCE) groups *and their concurrent controls*” were incorrect (emphasis added) (see Table 1 below, from 2005 erratum). The authors now note the correct study start dates were in 1994, not 1995, although exact start dates could “no longer be confirmed.” The 2014 author erratum now explicitly states that “all of the animal exposure experiments were run *with concurrent controls*” (emphasis added).

Strangely, the Work Plan Assessment states (p. 98) that this erratum “reaffirmed that the Johnson et al. studies are adequate to use in hazard identification and dose-response assessment.” Far from reaffirming the adequacy of Johnson *et al.* (2003), we believe that the erratum provides additional and irrefutable evidence of the incurable deficiencies of that study.

Examination of the data in Table 1, even as corrected in the 2014 erratum, indicates that the claim of concurrent controls is incorrect. The times described for evaluation of control data presented as “concurrent” to the 1,100 and 1.5 ppm TCE treatments (assuming the individual data lines in the Table represent control data on the left for corresponding TCE treatment groups on the right) still are dramatically different from each other. For example, the start times listed for the 1,100 ppm TCE treatment are 29 Jun 1989-12 Mar 1990, while the “concurrent” controls include evaluation times up to 10 Oct 1992, over 2-1/2 years *later*. Similarly, all of the dates for the apparent controls for the 1.5 ppm TCE treatment are listed as starting two to almost three years

²⁶ Hardin, B, *et al.*, Trichloroethylene and cardiac malformations, *Environ. Health Perspect.* 112: A607-8 (2004). Moreover, EPA’s benchmark dose guidance requires concurrent controls in key studies to be used in the calculation of an RfD or RfC. EPA, Benchmark Dose Technical Guidance (EPA/100/R-12/001) (2012) (available at http://www2.epa.gov/sites/production/files/2015-01/documents/benchmark_dose_guidance.pdf).

²⁷ Johnson *et al.*, *Environ. Health Perspect.* 122: A94 (2014), further correcting an earlier erratum (Johnson *et al.*, *Environ. Health Perspect.* 113: A18 (2005)).

later. In fact, with the possible exception of an unidentified number of controls listed as starting between 14 Jun 1989 and 10 Oct 1992, the exposure dates of all other controls listed in Table 1 are listed as occurring at least two to four years later than either of the 1,100 ppm and 1.5 ppm TCE treatments (note that the 2014 erratum corrects the exposure dates for “concurrent controls” for the 2.5 ppb and 250 ppb treatments to “unconfirmed” times in 1994, not 1995 as shown in Table 1). The description of having conducted “concurrent controls” as noted in the 2014 author erratum is inconsistent with the accepted technical definition of “concurrent control” as control that occurs *while an activity is in progress* (i.e., controls parallel in time to treatments).

The 2005 erratum states that control values of cardiac malformations were “statistically consistent across and throughout all treatment groups” (data not provided). However, an examination of the data in Table 1 indicates a potential of substantial variability in the average number of fetuses per mother within the various control groups as well as relative to TCE treatments, a factor that could be an untested confounder to cardiac malformation outcomes. The average number of fetuses per mother, calculated from the data in Table 1, is 9, 11.9, 10.3, 12, and 12.2 for the respective control groups, and 11.7, 13.9, 12.0 and 12.2 for the four respective TCE treatments. In addition to this intergroup variability and apparent lack of concurrent controls, it also appears that other confounders were present within the pooled control group population used in Johnson *et al.* (2003) as the basis for their conclusion that TCE induced cardiac malformations. In the initial developmental toxicity reporting exposures to 1.5 and 1,100 ppm TCE, both controls and treatment groups were described as exposed to “normal tap water” (Dawson *et al.*, 1993), while in the subsequent Johnson *et al.* (2003) study, adding the 2.5 ppb and 250 ppb groups, the control animals were described as exposed to “distilled water.” Another potential confounder across these studies is that the Dawson *et al.* (1993) study also included TCE treatments with pre-gestation treatments of approximately two months, implying that the age of the animals at the time of fetal evaluations, and the length of concurrent control pretreatments, likewise was varied across experiments. The data provided in the errata do not allow any assessment of these potential confounders.

Importantly, the data presented by Johnson *et al.* (2003), and subsequently clarified in the two errata, do not allow calculation of the incidence of cardiac malformations per litter that is time-matched to concurrent controls (the standard practice for evaluation of developmental toxicity studies). Accepting the author claims in the 2014 erratum that exposure times cannot be confirmed for substantial amounts of either control or treatment data, it also can be presumed that it is now impossible to reconstruct a calculation of per litter incidence of cardiac malformations that is appropriately matched to concurrent controls. Thus, the data reported in Johnson *et al.* (2003) and as amended in two subsequent errata do not allow for data analysis generally accepted as essential to interpreting adverse outcomes of developmental toxicity study findings. The lack of data availability and clarity sufficient to construct key analyses associated with a key study should

disqualify the use of that study in important agency decisions such as a Work Plan Assessment intended to support rulemaking.

From Johnson (2005, erratum):

Table 1. Control versus TCE treatment groups and dates of exposure.

Control		Dose	TCE	
Fetuses/mothers ^a	Dates		Fetuses/mothers	Dates
135/15	14 Jun 1989–10 Oct 1992	1,100 ppm	105/9	29 Jun 1989–12 Mar 1990
155/13	11 Dec 1992–20 Oct 1993 ^a	1.5 ppm	181/13	29 Dec 1989–26 Dec 1990
62/6	15 Apr 1994–23 May 1994 ^a			
120/10	6 Jul 1994–7 Jul 1995	2.5 ppb	144/12	6 Jun 1995–13 Jun 1995
134/11	18 Jul 1995–6 Oct 1995	250 ppb	110/9	5 Jul 1995–21 Jul 1995

^aThe total number of control rat fetuses/mothers was 606/55. ^bOther studies that coincided with these control groups were carried out during December 1989–June 1995 (e.g., metabolites that were reported in other articles (Johnson et al. 1998a, 1998b)).

II. EPA Ignored the Peer Review of the Draft Work Plan Assessment

Pursuant to the IQA, OMB in 2005 issued a Final Information Quality Bulletin for Peer Review.²⁸ This bulletin recognizes that different types of peer review are appropriate for different types of information. For assessments considered “highly influential,” the agency is required to prepare a written response to the peer review report, indicating whether the agency agrees with the reviewers and how the agency will address the points made by reviewers.

These requirements have not been followed in the case of the Work Plan Assessment. EPA has ignored the commentary of its peer review panel, which concluded that the Assessment needed much more work before EPA could use the conclusions to justify regulations.²⁹ Moreover, the assessment was only a screening-level assessment, not sufficient to support regulation.

²⁸ 70 Fed. Reg. 2664 (Jan.14, 2005).

²⁹ Peer Review Meeting for EPA’s Draft TSCA Work Plan Chemical Risk Assessment for Trichloroethylene: Degreaser and Arts/Crafts Uses (CASRN: 79-01-6) 1,1,2-Trichloroethene (July 9 – August 21, 2013) (available at <http://www.scgcorp.com/tcl2013/prcomments.asp>). Attachments containing more detailed critiques of Johnson *et al.* (2003) are also available via this link.

One of the most important comments not adequately addressed was raised by the chair of the peer review panel, Dr. Penny Fenner-Crisp, a well-respected risk assessor who spent more than 22 years at EPA, including serving as Director of the OPPT Health and Environmental Review Division. With such a background, her comments on the Work Plan Assessment should have carried significant weight. In her response to Question 1-1 [“Please comment on whether the characterization provides a clear and logical summary of EPA’s analysis. Please provide specific suggestions for improving the document.”] she wrote:

“The draft document fails to articulate satisfactorily that the analysis described within should be characterized as a screening level assessment (emphasis added).

* * * * *

“I believe that the Agency acted prematurely in issuing this (screening level) assessment for public comment and in convening a formal scientific expert peer review, given the conclusions reached in it. If all of the conclusions had indicated “no problem, then that assessment should have been peer reviewed externally, to determine if there were outside expert agreement. Presumably, if so, then no further risk assessments would be needed. However, most (and, perhaps, all, if the Agency reverses its decision not to include an assessment of dermal exposure) of the exposure scenarios assessed in the present draft resulted in the conclusion of “indicates potential risks of concern.” This begs for refinement of the assessments, on both the exposure and hazard side of the equation. This is essential for any defensible regulatory actions to be undertaken.

* * * * *

“Prior to the July 17 meeting, I was inclined to suggest that some revision to the current document would be productive. Revision would be minimal and limited to embellishment with the necessary contextual information I recommended above and in my comments below on Question 1-2. After listening carefully to the comments and contributions from the other members of the Panel, I have concluded that there would [be] little benefit in revising this draft screening assessment. Rather, I would suggest that the effort be put into a higher tier, more refined assessment which would include empirical data gathered during the course of real-world uses, e.g., as OPP regularly asks be done for occupational exposures and sometimes for residential exposures, consumer use survey data, evaluation of exposure using additional modeling tools and a revisiting and reanalysis of the choices of toxicity and epidemiologic studies used to describe the health benchmark at the MEC99 level and the rationale for selecting the singular MOE of 30 to apply to the selected studies, each of which have varying degrees of credibility. This current draft screening level

assessment could then be attached as an appendix to the new second-generation assessment, and described, in summary form, in the early chapter(s) of the new assessment. I would have saved the resources expended for the current external peer review and spent them on the next-generation assessment.”

The Work Plan Assessment completely ignored these comments from the peer review panel chair and failed either to acknowledge or to dispute her assertion that the Assessment is actually a screening-level assessment. In terms of “disposition” of this comment, the final Work Plan Assessment barely mentions the word “screening.” In the 212-page document, “screening” is only mentioned nine times, three times in reference titles and six times in referring to the EPA model E-FAST2. The E-FAST2 model was also the basis of a comment from Dr. Fenner-Crisp, who felt that its use, in providing exposure estimates, would result in a screening-level risk assessment. Her concern appears to be well-founded, as the EPA website for the model (<http://www.epa.gov/oppt/exposure/pubs/efast.htm>) states that it “[p]rovides screening-level estimates of the concentrations of chemicals released to air, surface water, landfills, and from consumer products. . . . Modeled estimates of concentrations and doses are designed to reasonably overestimate exposures, for use in screening level assessment.”

The response to Comment 27 states that the “exposure assessment is not a theoretical bounding estimate or a worst case assessment” as the modelers “did not use the screening level parameters that are set as the defaults in the EFAST2 modeling software.” Instead, it appears that a mixture of high end (*i.e.*, consumer use patterns) and median (*i.e.*, air exchange rates) were used. In seeking to marginalize the comments from the chair of its peer review panel, however, EPA failed to refute her assertion that the Work Plan Assessment is actually a screening assessment. As discussed above, under OMB guidance a screening-level assessment is not be a suitable basis for the development of regulations.

Most compelling was the fact that EPA did not even acknowledge, much less respond to, the detailed critique of EPA’s reliance on Johnson *et al.* (2003) by peer reviewer Calvin Wilhite, which echoes most of the points outlined above:

“It is not clear why OPPT relied on the results of the Johnson et al. (2003) study to the exclusion of all other inhalation and oral developmental toxicity studies in rodents and rabbits. If in fact the OPPT is reliant upon only the inhalation data, why is it the Carney et al. (2001), the Schwetz et al. (1975), the Hardin et al. (1981), the Beliles et al. (1980) or the Dorfmueller et al. (1979) study was not used? Why is there no discussion of all of the available developmental toxicity inhalation bioassays in the present analysis?

* * * * *

“As submitted, the exposure parameters appear arbitrary (e.g., 0.5 and 1 hr/day) and may have been selected for sake of convenience. The data upon which conclusions put forward by OPPT on risk for developmental toxicity associated with arts and crafts use of TCE are not reliable. Nearly all developmental toxicity studies with TCE in rodents find no sign of teratogenicity (e.g., Beliles et al., 1980) or find only slight developmental delay (Dormueller et al., 1979). Chiu et al. (2013) cite the NRC (2006) report as verification of their risk assessment for TCE developmental toxicity, but actually the NRC (2006) concluded:

“Additional studies evaluating the lowest-observed-adverse-effect-level and mode of action for TCE-induced developmental effects are needed to determine the most appropriate species for human modeling.”

“In its present assessment, the OPPT ignored the serious deficiencies already identified in conduct of the Johnson et al. (2003) rat drinking water study upon which the BMD01 was based (Kimmel et al., 2009; Watson et al., 2006) [Attachments 1 and 2]. In their weight-of-evidence assessment, Watson et al. (2006) concluded:

“...application of Hill’s causality guidelines to the collective body of data revealed no indication of a causal link between gestational TCE exposure at environmentally relevant concentrations and congenital heart defects.”

“Those conclusions were consistent with Hardin et al. (2005). Perhaps most disturbing of all in US EPA’s reliance upon Johnson et al. (2003) as the key study (which for the basis for their lowest non-cancer TCE hazard index and margin of exposure) is the observation by Hardin and associates (2004):

“Conventional developmental and reproductive toxicology assays in mice, rats and rabbits consistently fail to find adverse effects of TCE on fertility or embryonic development aside from embryo- or fetotoxicity associated with maternal toxicity. Johnson and Dawson, with their collaborators, are alone in reporting that TCE is a “specific” cardiac teratogen.”

“One of the fundamental tenants in science is the reliability and reproducibility of results of scientific investigations. In this regard, one of the most damning of the TCE developmental toxicity studies in rats is that by Fisher et al. (2005) who stated:

“The objective of this study was to orally treat pregnant CDR(CD) Sprague-Dawley rats with large bolus doses of either TCE (500 mg/kg), TCA (300 mg/kg) or DCA (300 mg/kg) once per day on days 6 through 15 of gestation to determine the effectiveness of these materials to induce cardiac defects in the fetus. All-trans-retinoic acid (RA) dissolved in soybean oil was used as a positive control.”

“The heart malformation incidence for fetuses in the TCE-, TCA- and DCA-treated dams did not differ from control values on a per fetus or per litter basis. The RA treatment group was significantly higher with 33% of the fetuses displaying heart defects.”

“Unfortunately, Johnson et al. (2005) failed to report the source or age of their animals, their husbandry or provide comprehensive historical control data for spontaneous cardiovascular malformations in their colony. The Johnson study with 55 control litters compared to 4 affected litters of 9 treated was apparently conducted over a prolonged period of time (perhaps years); it is possible this was due to the time required to dissect and inspect fresh rodent fetuses by a small academic research group. However, rodent background rates for malformations, anomalies and variants show temporal fluctuations (WHO, 1984) and it is not clear whether the changes reported by Johnson et al. (2005) were due to those fluctuations or to other factors. Surveys of spontaneous rates of terata in rats and other laboratory animals are common particularly in pharmaceutical and contract laboratory safety assessment (e.g., Fritz et al., 1978; Grauwiler, 1969; Palmer, 1972; Perraud, 1976). The World Health Organization (1984) advised:

“Control values should be collected and permanently recorded. They provide qualitative assurance of the nature of spontaneous malformations that occur in control populations. Such records also monitor the ability of the investigator to detect various subtle structural changes that occur in a variety of organ systems.”

“Rates of spontaneous congenital defects in rodents can vary with temperature and housing conditions. For example, depending on the laboratory levocardia and cardiac hypertrophy occur in rats at background rates between 0.8-1.25% (Perraud, 1976). Laboratory conditions can also influence study outcome; for instance, maternal hyperthermia (as a result of ambient elevated temperature or infection) can induce congenital defects (including cardiovascular malformations) in rodents and it acts

synergistically with other agents (Aoyama et al., 2002; Edwards, 1986; Zinskin and Morrissey, 2011). Thus while the anatomical observations made by Johnson et al. (2003) may be accurate, in the absence of data on maternal well-being (including body weight gain), study details (including investigator blind evaluations), laboratory conditions, positive controls and historical rates of cardiac terata in the colony it is not possible to discern the reason(s) for the unconventional protocol, the odd dose-response and marked differences between the Johnson et al. (2003) results and those of other groups.

“As noted by previous investigators, the rat fetus is ‘clearly at risk both to parent TCE and its TCA metabolite’ given sufficiently high prenatal TCE exposures that can induce neurobehavioral deficits (Fisher et al., 1999; Taylor et al., 1985), but to focus on cardiac terata limited to studies in one laboratory that have not been reproduced in other (higher dose) studies and apply the BMD01 with additional default toxicodynamic uncertainty factors appears misleading.”³⁰

EPA seems to support dismissing Dr. Wilhite’s points with the following:

“EPA/OPPT used developmental endpoints for the acute risk assessment based on U.S. EPA’s policy that a single exposure of a chemical within a critical window of fetal development may produce adverse developmental effects (EPA, 1991). Particularly, this assessment used the PBPK-derived HECs reported for developmental animal studies reporting fetal cardiac defects. TCE-induced fetal cardiac malformations are biologically plausible based on the weight of evidence analysis presented in the TCE IRIS assessment, which considered human and animal findings as well as mechanistic data.

* * * * *

“[R]isk estimates are focused on the most susceptible life stage, which are pregnant women and their developing fetus. This focus is supported by the hazard findings in the TCE IRIS assessment, which conclude that developmental toxicity is the most sensitive health effect associated to TCE exposure.”

Dr. Wilhite’s extensive comments raise serious questions regarding the quality of the University of Arizona studies and, in particular, the study by Johnson *et al.* (2003) selected by EPA

³⁰ *Id.* at 56-73.

as the driver for its acute risk values. The foregoing EPA statement does not even begin to address these quality concerns.

EPA's denial of the earlier HSIA request for correction of the IRIS Assessment relied extensively on the peer review of the draft Toxicological Review by its Science Advisory Board, which EPA characterized as supporting its approach. Here, EPA has paid no heed to any of the extensive peer review it received on the draft Work Plan Assessment, which did not support its approach at all.

Moreover, other relevant reviews do not support EPA's reliance on Johnson *et al.* (2003). A 2006 report by the National Academy of Sciences, issued in response to an interagency request for independent guidance on scientific issues to support an objective and scientifically balanced health risk assessment for TCE, concluded:

"The committee noted that the rodent studies showing trichloroethylene-induced cardiac teratogenesis at low doses were performed by investigators from a single institution. Also noted were the unusually flat dose-response curves in the low-dose studies from these investigators. For example, the incidences of heart malformations at trichloroethylene concentrations of 1.5 and 1,100 ppm (almost three orders of magnitude greater) were 8.2% to 9.2% (prepregnancy and during pregnancy) to 10.4% (during pregnancy only) (Dawson *et al.* 1993). The same pattern occurred with dichloroethylene. Thus, the animal data are inconsistent, and the apparent species differences have not been addressed."³¹

A more recent National Academy of Sciences report was even more dismissive:

"Early rodent studies using inhalation exposure (Schwetz *et al.* 1975; Dorfmueller *et al.* 1979) indicated little or no developmental toxicity as a result of exposure, whereas later studies by Dawson *et al.* (1990, 1993) and Johnson *et al.* (1998a,b, 2003) reported an increase in cardiovascular malformations at concentrations as low as 0.25 ppm. However, the latter studies used direct delivery of TCE to the gravid uterus or in drinking water and a novel examination process for examining the heart and great vessels. Fisher *et al.* (2001), using the same examination process as the Dawson and Johnson groups and in collaboration with them, reported no increase in cardiac or vascular defects. Warren *et al.* (2006) examined fetuses from the Fisher *et al.* (2001) study and found no ocular defects after TCE exposure. More recently, Carney *et al.*

³¹ National Research Council, Addressing the Human Health Risks of Trichloroethylene: Key Scientific Issues (2006), at 210.

(2006), using a standard test protocol (inhalation exposure to TCE at 0, 50, 150, or 600 ppm for 6 h/day on gestation days 6-20), reported no effect of TCE on development in rats at up to 600 ppm, a concentration that produced minimal maternal toxicity. . . .

The recent studies by Carney et al. (2006) address some of the recommendations of the 2006 National Research Council report that additional studies are needed to evaluate a LOAEL. The Carney study clearly shows no effects on heart or other organ development in the rat at exposure concentrations up to a minimal maternally toxic concentration. Several studies have been published to address mode of action but have not made clear which species is most appropriate for human modeling. Otherwise, the more recent data reviewed here do not change the conclusions of the 2006 National Research Council report on the prenatal toxicity of TCE. *An in-depth review of the animal and human data on cardiovascular defects by Watson et al. (2006) concluded that there is no indication of a causal link between TCE and cardiovascular defects at environmentally relevant concentrations. On the basis of that review and the Carney et al. (2006) study results, the conclusion is appropriate.*

“In summary, the database on the prenatal developmental effects of TCE is robust and indicates a lack of pregnancy outcomes up to concentrations that are minimally toxic in adults. The in vitro and whole-embryo studies are intriguing, but effects reported in them are probably due to the degree of exposure. On the basis of the Carney et al. (2006) study, the LOAEL of inhalation exposure during prenatal development in rats is greater than 600 ppm, and the NOAEL is also 600 ppm. The LOAEL for maternal or adult toxicity is 600 ppm, and the NOAEL is 150 ppm.”³²

Again, the Work Plan Assessment fails to address or even acknowledge these conflicting reviews of the same study.

III. Other Information Quality Act Issues

Congress passed the IQA in 2000 to ensure the quality, objectivity, utility, and integrity of the scientific, technical, and statistical information that federal agencies adopt and disseminate to the public. HSIA submits that EPA’s exclusive reliance on a single inappropriate study as the

³² National Research Council, Contaminated Water Supplies at Camp Lejeune: Assessing Potential Health Effects (2009), at 102-03 (emphasis added).

basis for the Work Plan Assessment constitutes erroneous information, the dissemination of which contravenes the IQA.

EPA's IQA Guidelines "contain EPA's policy and procedural guidance for ensuring and maximizing the quality of information [it] disseminate[s]" as well as specifically describing "new mechanisms to enable affected persons to seek and obtain corrections from EPA regarding disseminated information that they believe does not comply with EPA or OMB guidelines."³³ Accordingly, the Guidelines expressly set out a pathway for seeking correction of information disseminated by EPA that falls short of the "basic standard of quality, including objectivity, utility, and integrity," contained in the EPA Guidelines and those issued by OMB.³⁴

Both the "objectivity" and "utility" criteria are implicated by EPA's reliance on Johnson *et al.* (2003) as the basis for its Work Plan Assessment. As does OMB, EPA considers the "objectivity" inquiry for IQA purposes to be "whether the disseminated information is being presented in an accurate, clear, complete, and unbiased manner, and as a matter of substance, is accurate, reliable, and unbiased." The "utility" criterion refers to "the usefulness of the information to the intended users."³⁵

For giving content to the concept of ensuring the "objectivity" of "influential scientific risk assessment information," EPA, in developing the Guidelines, adapted the quality principles in the Safe Drinking Water Act Amendments ("SDWA") of 1996 as follows:

"(A) The substance of the information is accurate, reliable and unbiased.
This involves the use of:

- (i) the best available science and supporting studies conducted in accordance with sound and objective scientific practices, including, when available, peer reviewed science and supporting studies; and
- (ii) data collected by accepted methods or best available methods (if the reliability of the method and the nature of the decision justifies the use of the data).

³³ EPA Guidelines, at 3.

³⁴ *Id.*

³⁵ *Id.* at 15; OMB Guidelines § V.2, V.3, 67 Fed. Reg. at 8459.

(B) The presentation of information on human health, safety, or environmental risks, consistent with the purpose of the information, is comprehensive, informative, and understandable.”³⁶

Like OMB, EPA recognizes that the “influential scientific, financial, or statistical information” it disseminates “should meet a higher standard of quality.”³⁷ Under the EPA Guidelines, information is considered influential if “the Agency can reasonably determine that dissemination of the information will have or does have a clear and substantial impact (*i.e.*, potential change or effect) on important public policies or private sector decisions.”³⁸

For influential scientific information, EPA requires a “higher degree of transparency about data and methods” to “facilitate the reproducibility of such information by qualified third parties.” The Guidelines further state: “*For disseminated influential original and supporting data, EPA intends to ensure reproducibility according to commonly accepted scientific, financial, or statistical standards*” and “It is important that analytic results for influential information have a higher degree of transparency regarding (1) the source of the data used. . . and (4) the statistical procedures employed.”³⁹ “Reproducibility” means that the information is capable of being substantially reproduced, *i.e.*, “that independent analysis of the original or supporting data using identical methods would generate similar analytic results.”⁴⁰

Johnson *et al.* (2003) clearly does not meet the applicable IQA objectivity, integrity, or reproducibility criteria. EPA’s exclusive reliance on a single flawed and unreproducible study as the basis for the critical toxicological value in the Work Plan Assessment contravenes the IQA.

IV. Epidemiological Evidence Relating to Cardiac Anomalies

³⁶ EPA Guidelines, at 22.

³⁷ EPA Guidelines, at 19.

³⁸ *Id.*

³⁹ *Id.* at 20-21 (emphasis added).

⁴⁰ OMB Guidelines, 67 Fed. Reg. at 8460.

HSIA recently sponsored a critical review of the epidemiologic literature regarding the association between congenital heart defects (“CHD”) and exposure to TCE.⁴¹ It concluded that overall, the reviewed studies provide no substantive or consistent epidemiologic evidence of a causal relationship between TCE exposure and CHD. The literature assessing this association is relatively sparse, consisting of only about a dozen studies covering eight different populations. A positive association with exposure was reported for four of these, but these positive studies contained substantial design or analytic limitations that could easily have explained the elevated results.

The strongest associations were reported by Yauck *et al.* (2004) and Forand *et al.* (2012), each finding a significant 5- to 6-fold increase among certain subgroups. However, the former finding was the result of *post-hoc* model shopping for interaction, with no main effect reported for TCE exposure (OR: 1.0). The latter finding was based on only three CHD cases using models adjusted for nine strata from six covariates, suggesting sparse-data concerns and model over-fitting. The 3-fold increased risk reported by Goldberg *et al.* (1990) was unadjusted for confounding despite the fact that several risk factors were 2- to 3-fold more common among exposed cases than among either controls or cases without exposure. Finally, the marginally increased CHD risk (OR: 1.2 – 1.3) reported by Bove *et al.* (1995) was not statistically significant and was not an *a priori* study focus.

Bove *et al.* (1995) suggest that the marginal nature of their reported increase might have been due to the relatively low-level TCE exposures found in NJ water. Yet this logic is not consistent with the 2- to 6-fold increased risks reported by other positive studies in which exposure estimates were in the same general range (Table 1). Furthermore, some negative studies had higher TCE exposures (Table 1) but reported fewer CM/CHD cases than expected. Such results suggest that risk is unrelated to level of exposure, and that positive findings might be better explained by analytical flaws or study bias.

Several of the studies were highly exploratory, making dozens to a hundred or more comparisons in search of positive findings. Chief among these was Bove *et al.* (1995) and Bove (1996), which explored 117 planned comparisons in addition to post-hoc analyses. In such an instance, a non-significant 20 – 30% increase is inconsequential. Forand *et al.* (2012) reported dozens of tests of association, and likely explored other unreported associations as well. Indeed,

⁴¹ Bukowski, J., Critical Review of the Epidemiologic Literature Regarding the Association between Congenital Heart Defects and Exposure to Trichloroethylene, *Crit Rev Toxicol*, 2014; Early Online: 1–9. (All references in text in this section are cited in this review.)

ATSDR acknowledged more than 200 tests of exposure/disease associations in its assessment of the Endicott population.

Similarly, the finding by Yauck *et al.* (2004) of significant effect modification is not unusual given that main effects were ignored, thereby reducing the model building to a shotgun approach looking for all possible interactions with exposure. The highly unlikely finding of a 6-fold increase in risk among older women with exposure, despite no overall effect of exposure and a protective effect among younger exposed women (OR: 0.90), is highly suggestive of a chance or spurious result (Yauck *et al.* 2004).

Uncontrolled/residual confounding was also an issue in this literature. This concern was paramount in Goldberg *et al.* (1990), given the strong propensity for exposed subjects to be both more Hispanic and lower SES. However, residual confounding was also a concern for both Forand *et al.* (2012) and Bove *et al.* (1995), given that these studies could adjust only for variables available on birth/death certificates. This precluded adjustment for more established risk factors such as maternal diabetes or alcohol consumption. Furthermore, all studies were confounded by coexposures to various chemicals such as benzene, chromium, vinyl chloride, methylene chloride, etc., making it virtually impossible to reliably tease specific effects of TCE out of this chemical mélange.

V. Carcinogenicity

While acute risks of developmental toxicity are characterized by EPA as of the greatest concern, the Work Plan Assessment also concludes that all but one of the degreaser exposure scenarios exceeded all the target cancer levels. The discussion of carcinogenicity in the Work Plan Assessment suffers from slavish reliance on EPA's IRIS Assessment. The IRIS Assessment classifies TCE as "Carcinogenic to Humans." It fails to discuss (or even to recognize) that such classification is inconsistent with a definitive report by the National Academy of Sciences.⁴² Thus, it ensures that the public will continue to be confused as to the potential cancer risk posed by TCE. We briefly address below how the epidemiological data on TCE do not meet the threshold for classification as "Carcinogenic to Humans."

A. Guidelines for Carcinogen Risk Assessment

EPA's 2005 Guidelines for Carcinogen Risk Assessment provide the following

⁴² National Research Council, Contaminated Water Supplies at Camp Lejeune: Assessing Potential Health Effects (2009) (hereinafter "Camp Lejeune report").

descriptors as to the weight of evidence for carcinogenicity:

- Carcinogenic to humans,
- Likely to be carcinogenic to humans,
- Suggestive evidence of carcinogenicity,
- Inadequate information to assess carcinogenic potential, and
- Not likely to be carcinogenic to humans.⁴³

According to the Guidelines, “carcinogenic to humans” means the following:

“This descriptor indicates strong evidence of human carcinogenicity. It covers different combinations of evidence.

- “This descriptor is appropriate when there is convincing epidemiologic evidence of a causal association between human exposure and cancer.
- “Exceptionally, this descriptor may be equally appropriate with a lesser weight of epidemiologic evidence that is strengthened by other lines of evidence. It can be used when *all* of the following conditions are met: (a) There is strong evidence of an association between human exposure and either cancer or the key precursor events of the agent's mode of action but not enough for a causal association, *and* (b) there is extensive evidence of carcinogenicity in animals, *and* (c) the mode(s) of carcinogenic action and associated key precursor events have been identified in animals, *and* (d) there is strong evidence that the key precursor events that precede the cancer response in animals are anticipated to occur in humans and progress to tumors, based on available biological information. In this case, the narrative includes a summary of both the experimental and epidemiologic information on mode of action and also an indication of the relative weight that each source of information carries, *e.g.*, based on human information, based on limited human and extensive animal experiments.”

⁴³ 70 Fed. Reg. 17766-817 (April 7, 2005).

According to the Guidelines, the descriptor “likely to be carcinogenic to humans”:

“is appropriate when the weight of the evidence is adequate to demonstrate carcinogenic potential to humans but does not reach the weight of evidence for the descriptor ‘Carcinogenic to Humans.’ Adequate evidence consistent with this descriptor covers a broad spectrum. . . . Supporting data for this descriptor may include:

“An agent demonstrating a plausible (but not definitively causal) association between human exposure and cancer;

- “An agent that has tested positive in animal experiments in more than one species, sex, strain, site or exposure route, with or without evidence of carcinogenicity in humans;
- “A positive tumor study that raises additional biological concerns beyond that of a statistically significant result, for example, a high degree of malignancy or an early age at onset;
- “A rare animal tumor response in a single experiment that is assumed to be relevant to humans; or
- “A positive tumor study that is strengthened by other lines of evidence.”

According to the Guidelines, the descriptor “suggestive evidence of carcinogenicity”:

“is appropriate when the weight of evidence is suggestive of carcinogenicity; a concern for potential carcinogenic effects in humans is raised, but the data are judged not sufficient for a stronger conclusion. This descriptor covers a spectrum of evidence associated with varying levels of concern for carcinogenicity, ranging from a positive cancer result in the only study on an agent to a single positive cancer result in an extensive database that includes negative studies in other species. Depending on the extent of the database, additional studies may or may not provide further insights. Some examples include:

- “A small, and possibly not statistically significant, increase in tumor incidence observed in a single animal or human study that does not reach the weight of evidence for the descriptor ‘Likely to Be Carcinogenic to Humans;’

- “A small increase in a tumor with a high background rate in that sex and strain, when there is some but insufficient evidence that the observed tumors may be due to intrinsic factors that cause background tumors and not due to the agent being assessed;
- “Evidence of a positive response in a study whose power, design, or conduct limits the ability to draw a confident conclusion (but does not make the study fatally flawed), but where the carcinogenic potential is strengthened by other lines of evidence; or
- “A statistically significant increase at one dose only, but no significant response at the other doses and no overall trend.”

1. Application of the Guidelines to TCE

In considering the data in the context of applying the “Carcinogenic to Humans” descriptor, one first considers the weight of the epidemiological evidence. We judge the epidemiologic evidence to be neither “convincing” nor “strong,” two key terms in the Guidelines. This judgment is based on four recent reviews and meta-analyses of occupational TCE exposures and cancer as well as other reviews of this literature.⁴⁴ The recent review and meta-analysis by Kelsh *et al.* focuses on occupational TCE exposure and kidney cancer, and includes the Charbotel *et al.* study that is emphasized in the EPA assessment.⁴⁵ Both the EPA meta-analysis and the Kelsh *et al.* meta-analysis of the TCE kidney cancer epidemiologic literature produced similar summary results. However in Kelsh *et al.* the limitations of this body of research, namely exposure assessment limitations, potential unmeasured confounding, potential selection biases, and inconsistent findings across groups of studies, did not allow for a conclusion that there is sufficient evidence of a causal association, despite a modest overall association.

There are reasonably well-designed and well-conducted epidemiologic studies that report

⁴⁴ Alexander, D, *et al.*, A meta-analysis of occupational trichloroethylene exposure and multiple myeloma or leukaemia, *Occup Med (Lond)* 56:485–493 (2006); Alexander, D, *et al.*, A meta-analysis of occupational trichloroethylene exposure and liver cancer, *Int Arch Occup Environ Health* 81(2):127–43 (2007); Mandel, J, *et al.*, Occupational trichloroethylene exposure and non-Hodgkin's lymphoma: a meta-analysis and review, *Occup Environ Med* 63:597–607 (2006); Kelsh, M, *et al.*, Occupational trichloroethylene exposure and kidney cancer: a meta-analysis, *Epidemiology* 21(1): 95-102 (January 2010).

⁴⁵ Charbotel, B, *et al.*, Case-control study on renal cell cancer and occupational exposure to trichloroethylene, Part II: Epidemiological aspects, *Ann Occup Hyg* 50(8):777–787 (2006).

no association between TCE and cancer, some reasonably well-designed and conducted studies that did report associations between TCE and cancer, and finally some relatively poorly designed studies reporting both positive and negative findings. Overall, the summary relative risks or odds ratios in the meta-analysis studies (EPA or published meta-analyses) generally ranged between 1.2 and 1.4. The draft assessment refers to these associations as “small,” a term not typically consistent with “convincing” and “strong.” Weak or small associations may be more likely to be influenced by or be the result of confounding or bias. Smoking and body mass index are well-established risk factors for kidney cancer, and smoking and alcohol are risk factors for liver cancer, yet the potential impact of these factors on the meta-analysis associations was not fully considered. There were suggestions that these factors may have impacted findings (*e.g.*, in the large Danish cohort study of TCE exposed workers, the researchers noted that smoking was more prevalent among the TCE exposed populations, however little empirical data were provided). In addition, co-linearity of occupational exposures (*i.e.*, TCE exposure correlated with chemical and/or other exposures) may make it difficult to isolate potential effects of TCE from those of other exposures within a given study, and hinder interpretation across studies. For example, although Charbotel *et al.* reported potential exposure response trends, while controlling for many confounders of concern (which strengthens the weight of evidence), they also reported attenuated associations for cumulative TCE exposure after adjustment for exposure to cutting fluids and other petroleum oils (weakening the weight of the evidence). This study is also limited due to other potential study design considerations such as selection bias, self report of work histories, and residual confounding.

When examining the data for TCE and non-Hodgkin lymphoma, kidney cancer, and liver cancer, associations were inconsistent across occupational groups (summary results differed between aerospace/aircraft worker cohorts compared with workers from other industries), study design, location of the study, quality of exposure assessment (*e.g.*, evaluating studies that relied upon biomonitoring to estimate exposure *vs.* semi-quantitative estimates *vs.* self-report, etc.), and by incidence *vs.* mortality endpoints. Although EPA examined high dose categories, it did not evaluate any potential dose-response relationships across the epidemiologic studies (except for Charbotel *et al.*). Reviews of the epidemiologic data reported in various studies for different exposure levels (*e.g.*, cumulative exposure and duration of exposure metrics) did not find consistent dose-response associations between TCE and the three cancer sites under review.⁴⁶ An established dose-response trend is one of the more important factors when making assessments of causation in epidemiologic literature. Thus, based on an overall weight of evidence analysis of the

⁴⁶ Mandel, J, *et al.*, Occupational trichloroethylene exposure and non-Hodgkin's lymphoma: a meta-analysis and review, *Occup Environ Med* 63:597–607 (2006); Alexander, D, *et al.*, A meta-analysis of occupational trichloroethylene exposure and liver cancer, *Int Arch Occup Environ Health* 81(2):127–43 (2007); Kelsh, M, *et al.*, Occupational trichloroethylene exposure and kidney cancer: a meta-analysis, *Epidemiology* 21(1): 95-102 (January 2010).

epidemiologic research, these data do not support the conclusion that there is “strong” or “convincing” evidence of a causal association between human exposure and cancer.

EPA’s Guidelines also state that a chemical may be described as “Carcinogenic to Humans” with a lesser weight of epidemiologic evidence that is strengthened by other lines of evidence, all of which must be met. One of these lines of evidence is “extensive evidence of carcinogenicity in animals.” Therefore, we must briefly evaluate the animal data.

The criteria that have to be met for animal data to support a “carcinogenic to humans” classification are stated in a sequential manner with an emphasized requirement that all criteria have to be met. Since the Guidelines consider this to be an “exceptional” route to a “carcinogenic to humans” classification, we would expect rigor to have been applied in assessing animal data against the criteria. This simply was not done.

Of the four primary tissues that EPA evaluated for carcinogenicity, only one or perhaps two rise to the level of biological significance. Discussion of the remaining tumor types appears to presuppose that TCE is carcinogenic. The resulting discussion appears then to overly discount negative data, of which there are many, and to highlight marginal findings. The text does not appear to be a dispassionate rendering of the available data. Specifically, EPA’s conclusion that kidney cancer is evident in rats rests on *one* statistically significant finding in over 70 dose/tumor endpoint comparisons and references to exceedances of historical control values.⁴⁷ Using a 0.05 p-value for statistical significance, a frequency of 1 or even several statistically or biologically significant events is expected in such a large number of dosed/tumor groups. EPA’s overall conclusion based on these flawed studies cannot be that TCE is a known kidney tumorigen. The best that can be said is that the data are inconsistent. Certainly they do not meet the criterion of “extensive evidence of carcinogenicity in animals.” Several marginal findings do not constitute “extensive evidence.”

For all these reasons, EPA’s classification of TCE as “Carcinogenic to Humans” is not supported by the evidence and cannot be justified under the 2005 Guidelines.

2. Contrast between EPA Position of ‘Convincing Evidence’ and NAS Conclusion of ‘Limited or Suggestive Evidence’

⁴⁷ And that bioassay is from a laboratory whose studies EPA is reviewing and has placed on hold several ongoing IRIS assessments as a result.

The IRIS Assessment states that "TCE is characterized as 'carcinogenic to humans' by all routes of exposure. This conclusion is based on convincing evidence of a causal association between TCE exposure in humans and kidney cancer."

Box 2 of the Academy's Camp Lejeune report, enclosed, categorizes every cancer outcome reviewed in relation to exposure to TCE, the dry cleaning solvent perchloroethylene, or a mixture of the two. The categories are taken directly from a respected Institute of Medicine (IOM) report.⁴⁸ These categories are "sufficient evidence of a causal relationship," "sufficient evidence of an association," "limited or suggestive evidence of an association," "inadequate evidence to determine an association," and "limited or suggestive evidence of no association," all as defined in Box 1, also attached.

Looking at Box 2, evidence considered by EPA to be "convincing evidence of a causal association between TCE exposure in humans and kidney cancer" would seem to be considered "sufficient evidence of a causal relationship." Yet the Academy found no outcomes in that category. It would at least be "sufficient evidence of an association." Again, the Academy found no outcomes in that category. Only in the third category, "limited or suggestive evidence of an association," does one find kidney or any other cancer outcome associated with TCE.

"Limited evidence of an association" is far from "convincing evidence of causation." One would expect at the least a detailed explanation of EPA's very different conclusion. Although the 2009 Camp Lejeune study was already published, and indeed is cited in the references, there is no mention of it in the text of the IRIS Assessment, even though the previous draft had just been the subject of a multi-year review by the Academy.

The Camp Lejeune committee began with a comprehensive review of the epidemiology studies of the two solvents by the IOM for its Gulf War Report. They then identified new studies published from 2003 to 2008 and considered whether these changed the conclusions in the IOM report. In the case of TCE and kidney cancer, this was the case. The Camp Lejeune committee considered six new cohort studies and two case-control studies (including Charbotel *et al.*). They concluded that several of these studies reported an increased risk of kidney cancer, but observed that the results were often based on a relatively small number of exposed persons and varied quality of exposure data and methodology. Given these data, the committee raised the classification for TCE to match the IOM conclusion of "limited" evidence for perchloroethylene.

⁴⁸ Institute of Medicine, Gulf War and Health, Vol. 2, Insecticides and Solvents (National Academies Press) (2003).

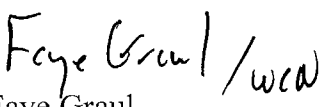
EPA, on the other hand, offered the summary conclusion of convincing human evidence, based on the "consistency" of increased kidney cancer across the different studies. The authors of these studies, however, do not agree with EPA's characterization of them. For example, the authors of Charbotel *et al.*, the study EPA finds most compelling, state that the "study suggests an association between exposures to high levels of TCE and increased risk of [renal cell carcinoma]. Further epidemiological studies are necessary to analyze the effect of lower levels of exposure."

Given the flaws in the IRIS Assessment, and the very different conclusion reached by the Academy in its Camp Lejeune report on the same body of data, the Work Plan Assessment should not rely on the IRIS Assessment's classification of TCE as "Carcinogenic to Humans."

VI. Conclusion

For the reasons set forth above, EPA should revise the Work Plan Assessment to ensure that it meets data quality requirements before relying on it as the basis for regulation of TCE. In particular, EPA should take into account the critical commentary offered by its own peer reviewers on the Work Plan Assessment.

Respectfully submitted,


Faye Graul
Executive Director

Enclosure

**Contaminated Water Supplies at Camp Lejeune,
Assessing Potential Health Effects
National Research Council of the National Academy of Sciences (2009)**

BOX 1 Five Categories Used by IOM to Classify Associations

Sufficient Evidence of a Causal Relationship

Evidence from available studies is sufficient to conclude that a causal relationship exists between exposure to a specific agent and a specific health outcome in humans, and the evidence is supported by experimental data. The evidence fulfills the guidelines for sufficient evidence of an association (below) and satisfies several of the guidelines used to assess causality: strength of association, dose-response relationship, consistency of association, biologic plausibility, and a temporal relationship.

Sufficient Evidence of an Association

Evidence from available studies is sufficient to conclude that there is a positive association. A consistent positive association has been observed between exposure to a specific agent and a specific health outcome in human studies in which chance and bias, including confounding, could be ruled out with reasonable confidence. For example, several high-quality studies report consistent positive associations, and the studies are sufficiently free of bias, including adequate control for confounding.

Limited/Suggestive Evidence of an Association

Evidence from available studies suggests an association between exposure to a specific agent and a specific health outcome in human studies, but the body of evidence is limited. . . .

Inadequate/Insufficient Evidence to Determine Whether an Association Exists

Evidence from available studies is of insufficient quantity, quality, or consistency to permit a conclusion regarding the existence of an association between exposure to a specific agent and a specific health outcome in humans.

Limited/Suggestive Evidence of No Association

Evidence from well-conducted studies is consistent in not showing a positive association between exposure to a specific agent and a specific health outcome after exposure of any magnitude. . . .

Source: IOM (Institute of Medicine). 2003. Gulf War and Health, Vol. 2, Insecticides and Solvents. Washington, DC: National Academies Press.

**Contaminated Water Supplies at Camp Lejeune,
Assessing Potential Health Effects
National Research Council of the National Academy of Sciences (2009)**

BOX 2 Categorization of Health Outcomes^a Reviewed in Relation to TCE, PCE, or Solvent Mixtures

Sufficient Evidence of a Causal Relationship

- No outcomes

Sufficient Evidence of an Association

- No outcomes

Limited/Suggestive Evidence of an Association

- Kidney cancer
- Adult leukemia (solvent mixtures)
- Multiple myeloma (solvent mixtures)
- Myelodysplastic syndromes (solvent mixtures)
- Scleroderma (solvent mixtures)
- Neurobehavioral effects (solvent mixtures)

Inadequate/Insufficient Evidence to Determine Whether an Association Exists

- Oral/pharyngeal cancer
- Nasal cancer
- Laryngeal cancer
- Esophageal cancer (TCE)
- Stomach cancer
- Colon cancer
- Rectal cancer
- Pancreatic cancer
- Hepatobiliary cancer
- Lung cancer (TCE)
- Bone cancer
- Soft tissue sarcoma
- Melanoma
- Non-melanoma skin cancer
- Breast cancer (TCE)
- Cervical cancer
- Ovarian/uterine cancer
- Prostate cancer
- Bladder cancer (TCE)
- Cancer of the brain or central nervous system
- Non-Hodgkin lymphoma
- Hodgkin disease
- Multiple myeloma
- Adult leukemia
- Myelodysplastic syndromes
- Childhood leukemia
- Childhood neuroblastoma
- Childhood brain cancer
- Aplastic anemia
- Congenital malformations
- Male infertility
- Female infertility (after exposure cessation)
- Miscarriage, preterm birth, or fetal growth restriction (from maternal preconception exposure or paternal exposure)
- Preterm birth or fetal growth restriction (from exposure during pregnancy)
- Cardiovascular effects
- Liver function or risk of cirrhosis
- Gastrointestinal effects
- Renal toxicity
- Amyotrophic lateral sclerosis
- Parkinson disease
- Multiple sclerosis
- Alzheimer disease
- Long-term reduction in color discrimination
- Long-term hearing loss
- Long-term reduction in olfactory function

Limited/Suggestive Evidence of No Association

- No outcomes

^aOutcomes for TCE and PCE unless otherwise specified*

* PCE-only outcomes omitted