Author’s response to reviews

Title: Evaluation of contaminated drinking water and preterm birth, small for gestational age, and birth weight at Marine Corps Base Camp Lejeune, North Carolina: A cross-sectional study

Authors:

Perri Z Ruckart (afp4@cdc.gov)
Frank J Bove (fbove@cdc.gov)
Morris Maslia (mmaslia@cdc.gov)

Version: 2 Date: 24 October 2014

Author's response to reviews: see over
October 24, 2014

Dear EH staff,

Thank you for considering the manuscript “Evaluation of contaminated drinking water and preterm birth, small for gestational age, and birth weight at Marine Corps Base Camp Lejeune, North Carolina: A cross-sectional study” for publication in Environmental Health. We appreciate the peer reviewers’ comments and have revised the manuscript accordingly as well as provided a response below to each comment received.

We look forward to your decision.

Sincerely,
Perri Ruckart
Epidemiologist
Agency for Toxic Substances and Disease Registry

---

**Evaluation of contaminated drinking water and preterm birth, small for gestational age, and birth weight at Marine Corps Base Camp Lejeune, North Carolina: A cross-sectional study**

**Reviewer: Ann Aschengrau**

**Reviewer's report:**

This is a very thorough re-analysis of adverse birth outcomes in relation to contaminated drinking water at Camp Lejeune, North Carolina with many improvements since the previous 1998 study. My comments are very minor and can be left to the discretion of the authors.

A. Abstract:

1. **Comment:** Revise sentence: “We analyzed the entire pregnancy and individual trimesters.” with something like “We analyzed exposure data for the entire pregnancy and individual trimesters.”

   **Response:** Abstract was revised as reviewer suggested.

2. **Comment:** Add the interpretation of the results in words.

   **Response:** Given that the abstract is already at the word limit (350 words) and we don’t think there is anything in the abstract that can be deleted, we do not have space to add additional details.

B. Methods

1. **Comment:** State what was done when the LMP was missing, nonsensical or inexact.

   **Response:** We added the following sentence to the first paragraph of the results section: We excluded 113 births with missing or incomplete information on LMP.
2. **Comment:** Give more information on the amount of missing risk factor data.

**Response:** We added this information to the first paragraph of the results section.

3. **Comment:** I am unsure about the backwards elimination procedure. You state that you removed potential confounders “with the value closest to the null.” I assume that this value is for the association between the confounder and the outcome. Please clarify.

**Response:** Yes, that is correct. We have revised the sentence in the Data Analysis section to “Order of the elimination was determined by removing the potential confounder with the value closest to the null for the association between the confounder and the outcome and continuing until no factor could be removed without changing the estimate for the drinking water exposure by >10%.”

C. Discussion

1. **Comment:** You may want to add some biological support for the findings of increased risks for second trimester exposure.

**Response:** We agree and added two references to the 2nd paragraph of the Discussion that support the findings of increased risk for adverse birth outcomes and 2nd trimester exposures to hazardous substances.

D. Limitations

1. **Comment:** It does not seem to me that labor and delivery complications should be considered confounders. Depending on the type of complication, they could be intermediate steps in the causal pathway and, for that reason, don’t need to be controlled. More elaboration would be useful.

**Response:** We agree that these variables do not need to be controlled for and deleted mention of them from the Limitations section.

E. Additional File 2

1. **Comment:** Figure 5: the "odds ratio" label for the y axis in this figure is incorrect.

**Response:** We have corrected the y axis.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

**Reviewer:** Lucas Salas

**Reviewer’s report:**

The paper by Perri Z. Ruckart, Frank J. Bove, & Morris Maslia entitled “Evaluation of contaminated drinking water and preterm birth, small for gestational age, and birth weight at Marine Corps Base Camp Lejeune, North Carolina: A cross-sectional study” is a continuation of their previous work in evaluating a
historical leaking of volatile organic compounds (mainly TCE, PCE and benzene) into the groundwater provided to an US Marine Corps Base in North Carolina between years 1968 and 1985. They have previously evaluated other birth outcomes in the same setting.

The following are my comments for the authors, that I I hope will help to improve the paper quality:

• Major Compulsory Revisions

Comment 1: I disagree about the design stated by the authors. A cross-sectional design assumes that prevalence of the outcomes and the exposures occurred at the same timeframe just like a snapshot (cross sectional). In your case, your dataset is a historical registry based cohort and the exposure assessment provides information of a timing of exposure which is assumed to be previous to the outcome appearance. In my opinion that corresponds to a historical (retrospective) cohort not to a cross sectional design. Please correct accordingly or if you disagree provide specific references to support the design.

Response: It is true that some researchers consider this type of study to be a retrospective cohort study because the study population is a cohort (i.e., a “birth cohort”), the outcome is known at the time the cohort is identified, and the relevant exposure period occurred prior to birth. However, each of the adverse birth outcomes evaluated in this study is a prevalence outcome because it is an “existing state” at the time of birth (Rothman et al, Modern Epidemiology, 3rd edition, 2008, page 46). According to the textbook, Research Methods in Occupational Epidemiology (Checkoway et al, 2nd edition, 2004, p. 76): “The key feature in a cross-sectional study is that we study disease prevalence rather than incidence. Exposure may be measured at the same time point when disease status is determined, or may involve historical data.”

The textbook Modern Epidemiology, 3rd edition has a lengthy discussion of the terms “prospective” and “retrospective” (pp 95-97) and discusses the various ways these terms have been interpreted by researchers. However, this textbook recommends the use of the term “retrospective” if the “disease could influence the exposure information in the study.” This would not be the case in our study since the exposure is based on: (1) the mother’s residence during pregnancy identified from family housing records and (2) the modeled levels of contaminants in the drinking water serving that residence during her pregnancy.

The term “retrospective” has also been used when the accumulated follow-up time occurs prior to the conduction of the study. However, there is no follow-up time in our study. A fetus is not included in the study unless it survives to become a live birth. At birth (“time-zero”), the child enters the study and its status (weight and gestational age) is measured. There is no follow-up beyond birth.

In summary, because the study subjects (live births) constitute a survivor population that is assembled at birth and is not followed beyond birth, and because each outcome is a “snapshot” of the birth’s status (weight, gestational age) at birth, we think it is appropriate to call this a cross-sectional study.

Methods:
Comment 2: At study population subtitle paragraph (page 6) “All singletons 28-47 weeks of gestation weighting ≤500 grams”. This period is non-plausible, as human pregnancy is normally 39 to 40 weeks. Maximum postterm delivery is 42 weeks under very tightly controlled conditions. At this sentence also weighting has the “t” missing. Please correct accordingly.

Response: We have corrected the misspelling. Some studies make no exclusions at all for implausible gestational ages or birth weights. Some studies are more restrictive for the range of eligible gestational ages (e.g., 32-44 weeks gestation in Bell et al. Ambient air pollution and low birth weight in CT and MA. Environ Health Perspect 2007;115:1118-1124). Other studies use a wide range for eligible gestational ages (e.g., 140-320 days in Ha et al. The effects of air pollution on adverse birth outcomes. Environ Research 2014;134:198-204). Some studies have no upper limit to eligible gestational ages (e.g., ≥22 weeks in Yorifuji et al. Residential proximity to major roads and preterm births. Epidemiology 2011;22:74-80 and Groom et al. Small-for-gestational-age infants classified by customized or population birthweight centiles: impact of gestational age at delivery. Am J Obstet Gynecol 2007;197:239e1-239e5). One study has an upper limit of 50 weeks gestation in its analysis of preterm birth (Forand et al. Adverse birth outcomes and maternal exposure to trichloroethylene and tetrachloroethylene through soil vapor intrusion in New York State. Environ Health Perspect 2012;120:616-621.

We decided on the range of 28-47 weeks. Our study used the date of last menstrual period (LMP) to determine gestational age at birth. This is routinely used in studies of birth weight and preterm birth. This method likely overestimates the length of time from conception to birth by at least 2 weeks. However, it is well known that use of the LMP to estimate gestational age “can result in a range of values that extends beyond what is generally viewed as biologically plausible...” (Alexander et al. Gestational age reporting and preterm delivery. Public Health Reports 1990;105:267-275). Some factors leading to implausible estimates include digit preference, variations in the pre-ovulatory interval and misidentification of the actual LMP due to sporadic bleeding or unrecognized miscarriage. Because gestational ages can be overestimated when based on LMP, births with gestational ages greater than 42 weeks are usually eligible for inclusion in studies.

Comment 3: You have used unconditional logistic regression to estimate your relative risks-RR using odds ratios to approximate your RR (Miettinen, 1982).

You are dealing with subjects in a historical dynamic cohort whose you censor once the delivery occurred (the subject was excluded from further prenatal exposure). This class of design usually uses other regression approaches to deal with this censoring to estimate the RR (usually a Poisson or a semiparametric Cox survival analysis). However, you decide to use of OR and then you should be careful. When you use OR to estimate your risks, it requires the “rare disease” assumption for being an adequate proxy of the RR in this kind of cohorts, otherwise your estimations will be inflated given than both the exposed and unexposed subjects had higher than 10% of “cases”. As stated in table 1A SGA and preterm births do not hold the classic “rare diseases’ label (as both are higher than 10%). In consequence your estimates may be inflated for these outcomes. Have you considered this potential inflation on your risk estimates? Have you used an alternative regression to compare the results? Please state this issue clearly on the paper.
Response: Two different issues were raised by this comment. First, there is the issue of whether the odds ratio (OR) or the risk ratio (RR) should be calculated in this study. Since the outcomes under evaluation are prevalence outcomes, one can use logistic regression to estimate the OR, or one can estimate the RR using either poisson regression or binomial regression (or negative binomial regression). For SGA and preterm birth, it makes little difference whether an OR or RR is calculated. For example, the percent of SGA in the highest exposure level of TCE is 15.8%, yet the OR and RR are very similar, 1.49 and 1.41, respectively. Smaller differences between the OR and RR occur for the analyses of preterm birth.

Second, there is the issue of censoring since preterm birth and SGA are “time-dependent”. Although these outcomes could be analyzed using survival methods, it is not necessary in this study. First, because we included only births with gestational ages between 28 and 47 weeks, all the births had the potential for exposure in each trimester. Second, we did not evaluate duration of exposure or cumulative exposure, both of which would be vulnerable to bias due to censoring. Third, we evaluated average exposure during each trimester and during the entire pregnancy, and we accounted for gestational age when assigning average exposure. For example, if a birth had a gestational age of 28 weeks, then the average exposure during the third trimester would be the exposure during week 28. For a birth with a gestational age of 29, the average exposure during the third trimester would be the average exposure over weeks 28 and 29, and so on. This approach avoided biases due to censoring.

Comment 4: In your case you do not have information about specific risk factors during pregnancy. Neither have you had available population exposure to other environmental or lifestyle factors in a very long time frame (1968 to 1985). On the other hand, your design includes the year of delivery and the mother age, both of them influencing the birthweight. In my opinion, this makes your data suitable for an age-period-cohort analysis given that other unmeasured factors in the period of pregnancy may be related to the outcome. A recent approach for attaining this has been recently published (Margerison-Zilko C. The contribution of maternal birth cohort to term small for gestational age in the United States 1989-2010: an age, period, and cohort analysis. Paediatr Perinat Epidemiol. 2014 Jul;28(4):312-21. doi: 10.1111/ppe.12127. Epub 2014 May 7.), which is accompanied by an editorial comment (Keyes KM, Ananth CV. Age, period, and cohort effects in perinatal epidemiology: implications and considerations. Paediatr Perinat Epidemiol. 2014 Jul;28(4):277-9. doi: 10.1111/ppe.12129.). Please support your approach and state the age-period (cohort) effect that may affect your estimations.

Response:

We agree that age-period-cohort approaches are useful for identifying trends in a particular outcome such as preterm birth, SGA and MBW. However, the purpose of this study was to evaluate whether the drinking water contaminants increased the risk for adverse birth outcomes, not to identify trends in the adverse birth outcomes. From the birth certificate, we had information on several risk factors including mother’s race, prenatal care, age of mother and father, parity, educational level of mother and father, sex of child, and if the mother had a previous fetal death. In addition, we had information on military rank of the parent. These were the risk factors evaluated in the previous study of adverse birth outcomes at Camp Lejeune (ATSDR 1998; Sonnenfeld et al 2001), and were the risk factors selected a
priori for evaluation in the current study. Mother’s birth cohort and the birth period were not evaluated in the previous study and were not among our a priori list of potential risk factors for the current study. We therefore are making no changes to the text based on this comment by the reviewer.

However, although mother’s birth cohort and the birth period were not selected a priori for evaluation as potential confounders, we were prompted by the comment to determine whether these factors confound and/or modify the effect of the contaminant-outcome relationships in this study, and we report on our findings in this response. A tertile categorization of mother’s birth cohort did not confound the contaminants’ relationships with the adverse birth outcomes evaluated in this study.

The prevalence of SGA declined among more recent births: 15.5%, 13.1% and 11.2% for births occurring in 1968-1973, 1974-1979, and 1980-1985, respectively. Mean birth weight among term births increased among more recent births: 3,374 g, 3,426 g, and 3,433 g for births occurring in 1968-1973, 1974-1979, and 1980-1985, respectively. Although the child’s period of birth did not confound the contaminants’ relationships with preterm birth or term low birth weight, confounding did occur for SGA. After adjustment for period of birth, the ORs for PCE and SGA from low to high exposure were 1.2, 1.2, 1.2, and 1.3, indicating a monotonic trend. This differs from the finding for PCE in Table 2 where no association is evident. On the other hand, the OR for TCE and SGA at the highest exposure category is reduced from 1.5 to 1.3 after adjustment for period of birth, and the OR for benzene at the highest exposure category is reduced from 1.2 to 1.0. There was also evidence of effect modification by child’s period of birth, with SGA ORs in the higher exposure categories for TCE of about 1.7 in the more recent period (1980-1985), compared to ORs of about 1.1 for the earlier periods. A similar pattern was seen for SGA and PCE with ORs of about 1.7 in the higher exposure categories during 1980-1985. Child’s period of birth also confounded the association between TCE and mean birth weight, reducing the mean birth weight difference in the highest exposure category from -78 to -30 g. However, effect modification in a similar pattern observed for SGA was also evident, with the strongest effect occurring during 1980-1985 with mean birth weight deficit of about -70 g in the highest exposure category. One explanation for the effect modification may be that contamination levels in the Tarawa Terrace and Hadnot Point water systems were highest during the period 1980-1985. The confounding effects of period of birth are difficult to explain given that period of birth was not a strong risk factor for SGA or mean birth weight deficit.

- Minor Essential Revisions

Tables and figures:

Comment 5: Usually when presenting cubic splines or GAMs you add a scatter of the number of subjects just above the x axis (exposure variable). This guides the reader about potential flat areas related to insufficient sample size. Other graphical approaches are also acceptable, but the estimates alone are difficult to interpret without this information. Please correct accordingly.

Response: Although we agree with the comment, the SAS macro that we are using does not produce such a graph. The splines are provided as supplementary material and are not a primary focus of the
analyses. We do present the confidence intervals on the spline graphs which provide information concerning the precision of the graph (and indirectly, the sample size) at various levels of exposure.

**Comment 6:** On table 3, the star marks PCE as it was adjusted by mother race. If I am right the star should mark the OR column, not the row, or it could be interpreted as if the percentages were adjusted and not the risk estimates. Please correct accordingly.

**Response:** We revised as suggested by the reviewer.

• Discretionary Revisions

**Comment 7:** When defining the three outcomes: Have you compared your term low birth weight (TLBW) with your term Small for gestational age (SGA)? I guess they agree but just out of curiosity given the different populations, how much they agree?

**Response:** Term SGA was not an a priori outcome. These are two very different outcomes. Slightly over 2% of the term births are low birth weight but over 13% are SGA. Therefore one would expect different findings. For TCE, the ORs are higher for term SGA than for term low birth weight, especially in the lowest and highest exposure group: 1.3, 1.3, 1.2 and 1.6. On the other hand, for benzene, the ORs are lower for term SGA: 1.0, 1.2. The findings for PCE were similar for term low birth weight and term SGA.

**Comment 8:** Have you any estimation of the number of pregnant women who had their delivery off-base? If so, this information may be interesting to add.

**Response:** There is no available data to answer this question.

**Comment 9:** On table 2, given the TCE findings a linear p-trend would be interesting to be added. The same comment for Table 5.

**Response:** We report p-values only for information purposes so that readers who expect to see p-values will not be disappointed. We do not base our interpretation of findings on p-values or trend test p-values. Additionally, the trend test is not a good indicator of a monotonic trend since p-values could be <0.05 when the trend is not monotonic and >0.05 when the trend is monotonic. (For example, the p-value for a test of trend is 0.01 for TCE and SGA even though the trend is not monotonic.)

Level of interest: An article of outstanding merit and interest in its field

Quality of written English: Acceptable