RESPONSES TO PEER REVIEW COMMENTS

Summaries of the reviewers' comments are in bold, and Office of Environmental Health Hazard Assessment (OEHHA) responses and discussion of revisions are unbolded.

Stephen H. LaFranchi, M.D.
Oregon Health and Science University

Preterm infants (<37 weeks gestation) are the subgroup most at risk for impairment in thyroid hormone production with any reduction in uptake of iodine.
OEHHA agrees and has now listed preterm infants as one of the groups that may be particularly susceptible to perchlorate (PHG document page 119).

A re-analysis of Brechner *et al.,* 2000 by Lamm 2003 concluded that the difference in TSH levels between exposed and unexposed communities may have been due to other differences between Yuma and Flagstaff, including a higher Hispanic ethnicity in Yuma vs. Flagstaff (59% vs. 13%) and altitude differences between Yuma and Flagstaff (138 ft. vs. 7,000 ft elevation, respectively).
The Brechner *et al.* (2000) results were stratified into Hispanic and non-Hispanic groups and very similar results were seen in both groups. This is strong evidence that the positive effects seen in this study were not due to differences in ethnicity between the two cities. Lamm (2003) raises the possibility that differences in altitude may have caused the TSH differences seen between the two cities, but provides no references or evidence that higher altitudes can substantially lower TSH levels. In fact, several studies suggest the opposite: that high altitude actually decrease thyroxine levels and have little to no effect on TSH levels (Kotchen *et al.,* 1973; Sawhney and Malhotra, 1991; Richalet *et al.,* 2010). This is now stated and the references are provided in the discussion of the Brechner *et al.* (2000) study (page 45).

Crump et al. (2000) and Tellez Tellez (2005) did not report differences in neonatal TSH or T4 levels. However, a re-analysis found a 45% higher mean TSH in infants born in Taltal.
The 45% higher TSH in Taltal was reported in Crump *et al.* (2000), not in Tellez Tellez *et al.* (2005). Crump *et al.* (2000) found only small differences in TSH values between Taltal and the other cities when blood samples for TSH measurements were collected more than two days after birth. But in samples collected on the first two days after birth, the median TSH value was 45% higher in Taltal. This was not based on a statistical re-analysis of the data, but rather on a simple observation of the data presented in Table 8 of the Crump *et al.* (2000) study. This is now explained in the part of the Public Health Goal (PHG) document describing the Crump *et al*. (2000) study (page 41).

None of the studies of perchlorate exposure on thyroid function in infants were designed prospectively. In general, perchlorate was not measured in the study subjects (i.e., mothers or newborns), but rather data on perchlorate levels in drinking water samples was used as a surrogate for estimates of perchlorate intake.

Using single or ecologic assessments of perchlorate and thyroid hormones, rather than individual prospective assessments, could cause misclassifications of true long-term perchlorate and thyroid hormone status. Since all studies assessed perchlorate and thyroid hormone levels independently, this misclassification would be non-differential and most likely bias results towards the null. It would not create false effects. This issue is already discussed on pages 52-54 of the PHG document in the sections on misclassification of exposure and outcome.

None of the studies of perchlorate and thyroid hormones in infants was able to control for all the variables known to affect thyroid function tests in newborns. Examples of variables that influence neonatal thyroid function tests include birth weight and gestational age, postnatal acute illness (e.g., respiratory distress), certain drugs, and age that the newborn screening specimen is obtained.
We now provide a much more thorough and quantitative discussion of confounding (beginning on page 54). As discussed, it is very unlikely that confounding was completely responsible for all of the highly consistent positive results identified in these studies.

While it makes sense that results in the first 24 hours reflect maternal and therefore fetal/neonatal iodine uptake and thyroid test levels, if some factor, e.g., perchlorate exposure, causes an elevated TSH or decreased T4 level, I would expect most likely it would persist and so be found in specimens obtained after 24 hours of age. Infants with high TSH or low T4 levels tend to track higher or lower, respectively, recovering to the reference range in a week or two after birth. I think these are some of the reasons that "experts" in this field have not reached a consensus on this issue, and that considerable difference of opinion remains on whether there is a "cause and effect" relationship between perchlorate levels in drinking water and neonatal thyroid function.
As we discuss on pages 38-39, if large numbers of infants consume formula with perchlorate concentrations that differ from the perchlorate concentration in the drinking water used by the mother during pregnancy, then associations based on the mother's prenatal drinking water perchlorate concentration would not be expected to persist, at least not at the levels seen soon after birth. In Steinmaus *et al.* (2010), the increased odds ratios for having a high TSH comparing neonates from perchlorate-exposed and unexposed communities did persist for up to 60 hours after birth (Steinmaus *et al.,* 2010). After this time, the odds ratios fluctuated dramatically. This wide fluctuation however does not mean that effects disappeared completely. Since the numbers of children who had elevated TSH levels after 60 hours was relatively small, the study did not have adequate statistical power to examine more subtle effects. In addition, in California, where this study was done, and in many other mandatory neonatal screening programs, almost all healthy children will have their TSH measurements collected just before they leave the hospital, usually within 48 hours of birth. As such, those having measurements collected after this time likely includes a large number of children who have birth complications or other medical conditions and remain in the hospital for more than 2-3 days. Some of these conditions could affect their thyroid hormone levels and, therefore, diminish the ability to identify perchlorate-related effects in these children.

This is another reason why greater emphasis should be placed on those children who have had their TSH levels collected soon (e.g., within 48 hours) after birth in these studies.

It is worth noting that the "critical" study by Greer *et al.* 2002 found an effect on decreased uptake of I-123 starting at a dose of 7 ug/kg-day (though not statistically significant until 20 ug/kg-day), but they did not find an effect on serum thyroid function tests except at a dose of 500 ug/kg-day (=3,500 ug/day in a 70 kg adult) – and this effect was a lower TSH level. Thus, it seems a stretch that intake of perchlorate in drinking water as low as 5-20 ug/L (or 5-20 ppb) in a mother would impact neonatal thyroid test results.
As discussed in the PHG document, healthy adults have several weeks of stored thyroid hormone. As such, it is possible that the 14-day dosing period in Greer *et al.* (2002) was too short to deplete these stores. In addition, for the reasons discussed throughout the PHG document, pregnant women, neonates, and several other groups (those with low iodine intake, high thiocyanate intake, thyroid disease, preterm infants) may be much more susceptible to perchlorate than the normal healthy adults used in the Greer *et al.* (2002) study. The findings from several studies provide evidence that perchlorate exposures that are much lower than those used in the Greer *et al.* (2002) study may impact thyroid function, especially in susceptible populations. These include the studies of newborns in Table 13 as well as the Blount *et al.* (2006) and Crump *et al.* (2000) studies. As a whole, there is a large and varied body of evidence that several important groups may be much more susceptible to the thyroidal effects of perchlorate than the healthy adults in the Greer *et al.* study.

The 95th percentile estimated water consumption rate for an infant 0-6 mo is reported at 0.234 L/kg-day, or 234 mL/kg daily (U.S. EPA, 2008b). An intake of 234 mL/kg-day might be necessary with excessive losses, e.g., with an acute illness, excessive heat and sweating, etc. It is hard to imagine ongoing, continuing intake at 234 mL/kg-day in an otherwise healthy infant. Such a continuing intake in a healthy infant might be dangerous, e.g., lead to "water intoxication." Ongoing intake of 234 ml/kg-day seems excessive. The 50th percentile estimated direct and indirect total water ingestion for infants <1 mo to 6 mo is 75-89 mL/kg-day (90% CI 64-114) (Kahn & Stralka 2009). I would think using the upper 90% CI for this estimate (e.g. 114 mL/kg-day) would be more realistic.
The reason OEHHA has chosen to use the 95th percentile drinking water intake rate rather than the median drinking water intake is that the California Safe Drinking Water Act of 1996 (Health and Safety Code, Section 116365) requires OEHHA to take into consideration the existence of groups in the population that are more susceptible to the adverse effects of contaminants than a normal healthy adult when establishing its PHGs. At a given chemical concentration in water, people who drink more water on a body weight basis (like infants) will have an overall increased intake of that chemical contaminant on a per kilogram body weight basis, and thus may be at greater risk from any harmful effects caused by that chemical. OEHHA's decision to use the 95th percentile drinking water intake rate rather than the 50th percentile (or its 90% confidence interval) is designed to help protect those people who drink more water on a

body weight basis than the normal healthy “median” adult or even the normal healthy “median” infant. This is consistent with OEHHA’s mandate to consider groups in the population (heavy water drinkers, in this case) who may be at greater risk than the average or median person. The 90% confidence interval is merely a statistical value meant to express the precision of the central tendency. It is not meant as an estimate of the drinking water intake rates of those who are at the higher end of the distribution of this variable.

Furthermore, OEHHA is now using the 95th percentile of drinking water intake rates from the Air Toxics Hot Spots Program Risk Assessment Guidelines Technical Support Document for Exposure Assessment and Stochastic Analysis (OEHHA, 2012) in its PHG calculations. This results in a slightly higher drinking water intake rate of 237 ml/kg-day. The reason for this is two-fold. First, the previous drinking water intake rate of 234 ml/kg-day (U.S. EPA, 2004) was based on all water sources, which would include water not obtained from a public water system, such as water from a household well or cistern, or water from unknown sources. Since the PHG is considered in setting MCLs, which are regulatory levels for public water systems, inclusion of water sources other than public water systems (community water) would not be appropriate. The new intake rates are derived from consumer-only, community water data. Second, an OEHHA analysis of the CSFII (Continuing Survey of Food Intake of Individuals) 1994-1996, 1998 and NHANES (National Health and Nutrition Examination Survey) 1999-2004 datasets found that the percentages of formula-fed infants ages 0 to <6 months consuming reconstituted formula (requiring the addition of water before it can be fed to the infant) were 71 percent and 87 percent, respectively (OEHHA, 2012). Additionally, a study of drinking water intake by 2-month-old infants in rural Canada noted that among 642 participants, 278 (43%) consumed dry formula reconstituted with water, and among these infants, 167 (60%) had their formula reconstituted with tap water (Levallois *et al*., 2007). Thus, it appears that a significant proportion of infants may have a higher risk of exposure to drinking water contaminants because they drink more water on a body weight basis than older children or adults. For deriving infant drinking water intake rates, OEHHA used data from the CSFII 1994-1996, 1998 dataset (the same dataset used by Kahn and Stralka, 2009) for infants consuming reconstituted formula. The drinking water intake rate estimated for infants ages 0 to <6 months is 237 ml/kg-day instead of the previously used 234 ml/kg-day. Details on how these rates were derived are now provided in the PHG document (page 120).

There are several reasons why OEHHA believes the 95th percentile water intake rates used in the PHG calculations represents an accurate estimate of the true 95th percentile intake of children ages 0-6 months in the U.S. population and is not consistent with water intoxication.
First, the data used to estimate the 95th percentile intake were derived from the CSFII, a large multistage probability sample involving over 20,000 individuals from throughout the U.S. run by the U.S. Department of Agriculture. The very large sample size of this survey helps to ensure the precision of its results, including those results at the tails of the distribution. It also helps ensure that a small number of outlying values, although they might affect the mean, would not substantially affect the 95th percentile.

Second, drinking water intake data in the CSFII was collected from each subject on two non-consecutive days, 3-10 days apart. Because multiple samples were collected from each person, "regression to the mean" should be less of a problem in this dataset than in other surveys that assessed drinking water intake on only a single day. OEHHA evaluated the possible magnitude of the impact of collecting drinking water data on two days versus only one day by comparing the distributions of the drinking water intake measured on day one of the CSFII to those of the two-day averages. If "regression to the mean" is a major problem one would expect the 95th percentile of the two-day average to be closer to the mean value than the 95th percentile of the day one-only measurements. Table 1 below shows the results of this analysis of the CSFII data for various age groups. As seen, the two-day average 95th percentiles were closer to the mean than the 95th percentiles from the day one-only measurements, but this effect was relatively small: 9.7 percent for all age groups combined and essentially zero for children < age 1. These results suggest that regression to the mean may not be a major source of bias in this dataset. The reason for this is unknown although it is likely related to the fact that unlike intake of many individual foods, a consistent intake of water is necessary for life and good health.

Third, OEHHA also notes that while the 95th percentile water intake rate for children 0-6 months is more than 2-fold higher than the median intake, a very similar pattern is seen in all other age groups (see Table 2 below). It seems unlikely that these 95th percentiles would be associated with water intoxication in every age group. This similarity across age groups provides additional evidence that the 95th percentile intakes measured in infants is valid and is not consistent with water intoxication.

Fourth, other sources of infant water intake data suggest that a water intake of 237 ml/kg-day is a reasonable estimate of the true 95th percentile water intake in infants. In an OEHHA analysis of two studies which longitudinally assessed infant breast milk intake, the 95th percentile intake was estimated to be 167 ml/kg-day (OEHHA, 2012). Although this is below 237 ml/kg-day, it seems unlikely that the 70 ml/kg-day difference would be enough to lead to water intoxication and death. The OEHHA estimate is also consistent with data from the large evaluation of U.S. water intake rates by Ershow and Cantor (1989). This evaluation used data from the 1977-78 Nationwide Food Consumption Survey (NFCS), a stratified random sample of over 30,000 people designed to represent the non-institutionalized U.S. population. Based on dietary and water intake records for three days, the 95th percentile intakes for ages 0-6 months in this survey were 353 ml/kg-day for total water intake and 155.6 ml/kg-day for tap water intake. OEHHA's estimate of 237 ml/kg-day (direct and indirect water in community water consumers) is within the range of these values.

Table 1. Comparing CSFII Drinking Water Intake Rates (ml/kg-day): Day 1 Intake Versus Two Day Average

	Day One			Two Day Average			Difference[b]	
	Mean	95th perc[a]	Diff[a]	Mean	95th perc	Diff[a]	Absolute	%
All	23.0	56.0	33.0	21.8	51.6	29.8	3.20	9.7%
Age < 1 year	86.0	200.1	114.1	84.4	198.8	114.4	-0.30	-0.3%
Age 1-6	33.5	84.7	51.2	31.4	74.4	43.0	8.20	16.0%
20+ males	20.5	46.8	26.3	19.4	43.3	23.9	2.40	9.1%
20+ females	22.2	49.5	27.3	21.4	45.4	24.0	3.30	12.1%
Pregnant	22.4	45.4	23.0	20.7	44.5	23.8	-0.80	-3.5%
Age 55+	21.4	43.5	22.1	21	40.7	19.7	2.40	10.9%

[a]Diff, Difference (95th percentile minus the mean); Perc, percentile
[b]Day One difference minus Two Day Average difference

Table 2. Ratio of 95th Percentile to Median Drinking Water Intake Rate by Age Groups

Age group	Median	95th percentile	Ratio: 95th percentile to median
0-6 mo.	123[a]	237[a]	1.93
1-3 years	20	68	3.40
4-6 years	18	63	3.50
7-10 years	13	40	3.08
11-14 years	10	36	3.60
15-19 years	9	32	3.56
20+ years	13	39	3.00
20-24 years	11	39	3.55
25-54 years	13	40	3.08
55-64 years	14	38	2.71
65+ years	16	37	2.31
All ages	13	44	3.38

[a] Values are from OEHHA, 2012; remainder of drinking water intake rates are from U.S. EPA, 2004

Finally, OEHHA reviewed the published literature on water intoxication and found little evidence that overall water intake of 237 ml/kg-day would cause water intoxication. Most reports of water intoxication appear to involve total water intakes of well over 300 ml/kg-day (Corneli *et al.,* 1985; David *et al.,* 1981; Keating *et al.,* 1991; Medani, 1987; Rodriguez-Soriano *et al.,* 1981). For example, in a case series of 34 infants (average age = 4.2 ± 2.1 months) treated at the St. Louis (Missouri) Children's Hospital between 1975 and 1990, Keating *et al.* (1991) estimated a water intake rate of 7.5 L/m^2 prior to hospitalization. In a 6 kg four-month-old child, this corresponds to a water intake rate of approximately 390 ml/kg-day, well above the 237 ml/kg-day value used in OEHHA's calculations. Using data on maximal free water clearance by the kidneys in infants provided by Rodriquez-Soriano *et al.* (1981), Medani (1987) estimated that children with normal filtration and diluting capacity should be able to excrete more than four liters of free water per day. In a 6 kg child, this would correspond to an intake rate of more than 600 ml/kg-day, again, much higher than 237 ml/kg-day. In conclusion, OEHHA found little evidence that a drinking water rate of 237 ml/kg-day would cause water intoxication in most children. Based on this finding, combined with the other evidence presented above, OEHHA concludes that a drinking water intake rate of 237 ml/kg-day provides a valid representation of the true 95th percentile intake in California infants.

Braverman 2006: The 2006 publication did not find an effect of long-term (6 mo) ingestion of two doses of perchlorate (0.5 mg/day and 3.0 mg/day) on RAI uptake or serum thyroid function tests. These results are difficult to reconcile with the reduced RAI uptake found by Greer 2002. One possible explanation is that chronic exposure to perchlorate may "up-regulate" the NIS, overcoming the effect of perchlorate. If correct, based on thyroid physiology, I would expect this compensation to work best in healthy adults, perhaps somewhat less well in young children, and perhaps not as well in infants.
The negative effect at the 0.5 mg/day dose in Braverman *et al.* (2006) is consistent with the very small, essentially negative effect seen at the lowest dose of 0.007 mg/kg-day in Greer *et al.* (2002) (In a 70 kg person, a dose of 0.007 mg/kg-day would be 0.49 mg/day). But, OEHHA agrees that the Braverman *et al.* (2006) results are inconsistent with the Greer *et al.* (2002) results at higher dose levels. The authors present several possible reasons for this and these are presented in the final paragraph of the section describing this study (page 69). Given that several other physiologic and metabolic systems have been shown to work less well in developing children than in adults, OEHHA agrees that if this upregulation occurs, it may work less well in young children than in adults. This is also now mentioned in this section.

Blount 2006: Using creatinine-adjusted urinary iodine levels, Lamm did not find a negative correlation of urinary perchlorate with serum T4 in women of childbearing age and low urinary iodine (Lamm 2007). Arguments can be made pro and con as to which results carry more credence; in general, creatinine-adjusted measurements are thought to correct for day-to-day variations in measured analytes.
See page 76-77 for full discussion of the weaknesses of dividing urinary analytes, like iodine, by urine creatinine. Several authors suggest that this method is not appropriate for the reasons discussed in this section. Summaries from these authors include:

- "Thus the use of µg iodine/ g creatinine ratio as an individual measure is doubtful. In fact, there is a poor correlation between µg iodine/ g creatinine ratio and 24-hour iodine excretion" (Vejbjerg *et al.,* 2009).
- "If these factors (body size, age, gender, diet, genetic polymorphisms) are not considered, adjustments of variations in urine dilution by creatinine may seriously invalidate the interpretation of biomarker data" (Nermel *et al.,* 2008).
- "It was found that the uncertainty associated with creatinine standardization (19-35%) was higher than the uncertainty related to volume standardization (up to 10%)" (Garde *et al.,* 2004).
- "We conclude that iodine-creatinine ratio in casual urine samples is an unsuitable indicator for evaluating iodine status in areas where large inter- and intraindividual variations in urinary creatinine excretion exist" (Furnee *at al.,* 1994).
- "The results for iodine also give no support for expressing iodine as the iodide-creatinine ratio" (Thomson *et al.,* 1996).
- "The urinary concentrations of these chemicals are often reported on a weight/volume basis and a creatinine-adjusted basis. However, urinary creatinine

> concentrations differ dramatically among different demographic groups; thus, biomonitoring studies using creatinine concentrations to adjust the concentrations of environmental and occupational chemical concentrations should seriously consider the impact these findings will have on the data." (Barr *et al.,* 2005).

As we conclude in our discussion of this topic, "Because urine creatinine concentration is dependent on all of these factors (age, sex, genetics, physical activity, muscle mass, and diet…) using it to adjust for urine dilution may introduce a degree of misclassification of true iodine status which could overwhelm any improved accuracy that results from correcting for urine dilution" (page 77).

In summary, several studies published since the 2004 perchlorate PHG offer evidence in support of the effect of perchlorate on RAI uptake, which in turn could affect thyroid function. At the same time, it needs to be disclosed that some of the studies do not, including Braverman 2006 which did not find an effect of chronic perchlorate intake on RAI uptake and the re-analysis by Lamm 2007 which appears to refute some of the findings of Blount 2006.
We state, **"**Lamm *et al*. (2007) analyzed the NHANES 2001-2 data using the iodine:creatinine (I/Cr) ratio rather than iodine concentration and *found no association* between T4 and perchlorate in women with low I/Cr values" (page 76). We also discuss the possible reason why the Lamm *et al.* (2007) and Blount *et al.* (2006) results differ (pages 76-77). In our discussion of Braverman *et al.* (2006) we state, "There were *no significant changes* in RAIU, T4, free T4 index (FTI), or TSH during or after the dosing period" (page 69).

I think it is important to bear in mind that the Greer study did not find a statistically significant decrease in iodide uptake at 0.007 mg/kg-day (the lowest dose used in the study).
RAIU was decreased 1.8 percent from baseline at the 0.007 mg/kg-day dose, but this estimate is very unstable due to the small sample size (n = 7 in this group). When we plotted the data for all of the dose groups using the Hill model (as was done in our benchmark dose calculations), it was estimated that a perchlorate intake of 0.007 µg/kg-day would cause about a 5 percent decrease in RAIU from baseline. Given the very small sample size and the large standard error reported for the 0.007 mg/kg-day finding, it is very possible that the difference between the predicted 5.0 percent and the observed 1.8 percent is solely due to chance.

David Q. Rich, Sc.D.
University of Rochester School of Medicine and Dentistry

Drinking water perchlorate exposure is an imperfect proxy as there may be other sources of perchlorate exposure during pregnancy, and thus bias may result from the use of this proxy as well. This should be discussed.

A discussion of exposure misclassification is provided on page 52, and in the discussions of many of the individual human studies on perchlorate and thyroid hormones throughout the document.

There is no discussion of outcome misclassification.
This has now been added (page 54).

Many of the studies (Crump *et al,.* 2000; Kelsh *et al*., 2003) control minimally for potential confounders between cities. A discussion of these confounders (both related to city of residence as the exposure and related to actual perchlorate level in subject's drinking water) should be provided (similar to Table 22).
In response to this reviewer's many concerns about confounding we have significantly expanded our discussion of confounding on pages 54-61 and provided many new quantitative analyses of confounding.

Confidence intervals should be provided in Table 13 and elsewhere.
This has been done.

In Table 13, "mean TSH = 27% higher" does not say that subjects from the exposed city had a higher TSH level than subjects from the unexposed study. Make it clear.
The top of this column states that these results are for the exposed group compared to the unexposed group.

A column with the statistical method used and confounding variables adjusted for in each study should be added.
The potential confounding variables that were assessed in each study using either adjustments, exclusions, or stratified analyses have now been added to Table 13. Statistical methods are provided in the descriptions of the individual studies.

Some of the effect estimates presented in Table 13 are not the main results of the paper, but instead were calculated using data in the tables of the paper by the authors of this draft document. Therefore, these are unadjusted estimates. This should be clearly marked in this table as well.
We now state in Table 13 that the Kelsh *et al.* (2003) results are a re-analysis, and the Li *et al.* (2000a) results are from a figure in their article. As mentioned above, a column with the statistical adjustments has also been added.

I assume the main result of the Steinmaus *et al*., 2010 paper uses the same data as the Kelsh *et al.* (2003) study (i.e. both of these OR's=1.57 are the same effect estimate on the same data). If not, then make it clear how they are different.
These studies are different. As stated in the first sentence of the description of Steinmaus *et al*. (2010) in the PHG document (page 48), this study used data from Buffler *et al*. (2006), not Kelsh *et al.* (2003). The Steinmaus *et al.* (2010) odds ratio is 1.53, not 1.57, and this typographical error in Table 13 has been corrected.

Related to the above request for confidence intervals is the use of the p-value in the text. A p-value is interpreted as reflecting the probability the findings were due to random chance. This is incorrect and should be removed. The p-value, by definition, combines both the strength of the association observed as well as the precision of that estimate (i.e. sample size). It is used to make a qualitative judgment on whether an effect is different from a hypothesized value, not a quantitative estimation of effect. By definition, the p-value is the probability, under the assumption that the null hypothesis is true (e.g. OR = 1), that you would have observed this result or one more extreme. The p-value is not the probability the null hypothesis is true, and thus is not the probability the result observed was due to chance. It should just be viewed as a relative measure of consistency of the study data with a null hypothesized OR=1 (Rothman 2002). The confidence interval should be provided for each study and used to assess the role of random chance in the study. I suggest removing all of these statements from the document.

Confidence intervals have now been reported throughout. We agree that we oversimplified the interpretation of the p-value and now state, "These very low p-values provide evidence that the elevated odds ratios identified in these studies are unlikely due to chance" (page 61).

Kelsh *et al,.* 2003. Although this result is suggestive of an effect, this is just an unadjusted estimate from a study with several limitations (imperfect outcome assessment, not complete TSH screening). The statement that confounding by age at collection, ethnicity, sex, birth weight, or birth year would not likely explain the effect observed, does not address the role of other potential confounders (e.g. other factors more closely linked to socioeconomic status, other environmental toxicant exposures, etc.). I would suggest you do not make such strong inference from them.

Since the study involved a mandatory state screening program, important bias from incomplete screening is likely very small. The likely impacts from imperfect outcome assessment are discussed above. The role of confounding in this and other studies is also discussed above. In addition, as we specifically mention for this study, "it is unlikely that adjusting for age at collection, ethnicity, sex, birth weight, or birth year would have any major impact on this odds ratio since adjusting for these factors had little impact on the TSH odds ratios provided by the authors." Importantly, in order for a variable to cause important confounding it must be related to both the outcome and the exposure of interest, and these associations must be fairly strong. The unadjusted odds ratios OEHHA calculated classified exposure and used the exact same perchlorate exposed and unexposed regions as those calculated by Kelsh *et al.* (2003). The fact that adjustments for age, ethnicity, sex, birth weight, and birth year had little effect on the Kelsh *et al.* (2003) results provides good evidence that these factors were not strongly related to perchlorate exposure in this study and therefore were not strong confounders either in their results or in OEHHA's results. This is important since these factors include some of the important determinants of thyroid hormone levels on a population basis. OEHHA does not make strong inference based on this single study alone. Instead, inference is based on the consistency of findings, detailed evaluations of

other important aspects of causal inference, and a weight of evidence approach that incorporates the findings and evaluations of many different studies.

Lamm and Doemland 1999. The document states "Results were not adjusted for several important variables". If they did not adjust for several important variables (I assume confounders), state what you think those variables were and whether not adjusting for these factors could explain the null finding.
See our notes regarding confounding above. We have now removed the statement regarding "important" confounders in this section since although confounding is a possibility we have no evidence that important confounding occurred in this study. The much more likely and important potential bias in this study is exposure misclassification and this potential bias is now emphasized.

Lamm and Doemland 1999. The statement about significant misclassification of exposure here needs to be clarified. There are two sources of exposure misclassification: 1) errors in the county to which subjects were assigned (were subjects incorrectly assigned to a county of residence?), and 2) using a county as a proxy for subject's perchlorate exposure instead of measuring it directly. Clarify which is the source of the bias. Discuss whether it's differential versus non-differential, and then whether it could explain the result they observed.
We now state that the bias is likely non-differential and towards the null. The second source of misclassification was already mentioned. The first source is also now mentioned although there is no evidence that it caused substantial bias, and given the very large size of these exposed counties (e.g. San Bernardino and Clark Counties have about 2 million people each) important bias from this source seems very unlikely (page 43).

Li et al, 2000a. **This is clearly a null study, but this lack of any adverse effect may be due to exposure and outcome misclassification, and/or confounding. However, none of these are discussed as potential reasons for this null finding. Further, the authors use Figure 3 of the paper to conclude that the mean T4 level in Las Vegas was 22 µg/mL lower than in Reno. I suggest you add a cautionary statement here that again, this is just observation from a figure and is not adjusted for potential confounders.**
Exposure and outcome misclassification and confounding are now mentioned (page 44). These issues are also reviewed in our overall summary section on these studies. It is now clearly stated in this section and in Table 13 that this T4 finding is from a figure provided in the article**.**

Li et al 2000b. **It would be useful to provide the mean and standard deviation TSH levels in Reno and Las Vegas, rather than just the p-value for the t-test. Further, although I agree weaknesses of the study are residual confounding by age, and uncontrolled confounding by birth weight and ethnic origin, it is not clear if these factors could completely explain the lack of an effect.**
These TSH means and standard deviations are now provided (page 44). We agree that it is not clear that these factors by themselves could completely explain the lack of an

effect, so we now say they could have caused at least *some* bias (page 44). Other factors are also now mentioned and we now state that misclassification of exposure and outcome, and the use of TSH levels after 24 hours of age were likely the more important biases.

Brechner et al. 2000. **The authors do a reasonable job of documenting the deficiencies of the study. However, they present Table 10 that provides median TSH levels in Flagstaff and Yuma for each day of measurement, and the difference in them. The authors then place a statement in Table 13 that the Yuma values are 27% higher than those in Flagstaff (6.4/24.0) as evidence that perchlorate is associated with increased TSH. However, this is an unadjusted effect estimate that may be due to confounding. Too much causal interpretation is put on a difference in TSH levels associated with living in one city versus another, which is not adjusted for potential confounders by factors such as city differences in SES, other environmental toxicants that may affect TSH levels (e.g. nitrate and thiocyanate), etc. This limitation needs to be made clearer in Table 13.**
The potential effects of socioeconomic status (SES), other environmental toxicants, and other potential confounders are discussed above and we have added a number of new analyses addressing the reviewer's concerns regarding confounding. Table 13 is meant to provide a summary of the findings. The unadjusted estimates are identified, but it is not possible to include a full discussion of confounding in this table. For this reason, we have chosen to address the reviewer's numerous concerns regarding confounding by adding a significant amount of new text about confounding to the document (pages 54-61). As mentioned above, we do not make inference based on any single study, but rather on all studies, likely biases and their likely magnitude and direction, and the totality of the evidence.

Buffler et al. 2006. **The authors of this document, using data in the paper, calculated an unadjusted odds ratio of having a TSH > 25 µU/mL regardless of age (both those <24 hours of age and those ≥ 24 hours) and present it in Table 11 (OR=1.59; 95% CI = 1.33-1.91). They then conclude that the effect does not appear to be due to confounding by age at sample collection. However, too much emphasis is placed on this simple unadjusted calculation that is likely confounded somewhat by age at sample collection (if in fact there is a difference in age at sample collection by high versus low perchlorate level) and potentially other confounders. This should be considered suggestive evidence at best, and this limitation should be noted in Table 13.**
The Buffler *et al.* (2006) data is not included in Table 13. In Steinmaus *et al.* (2010), which used the same data as Buffler *et al.* (2006), adjustments for collection age (and many other factors) had no important effects on the results.

Steinmaus *et al.* 2010. This study attempts to confirm the estimated concentration in Kelsh et al. (2003) and was done by an author of the draft document. Although the authors had the continuous TSH data, they chose to still dichotomize it into High vs. Low. A more convincing argument would have been to estimate the change in TSH level associated with an incremental increase in a

community's perchlorate concentration. Further, use of this continuous outcome would likely result in more statistical power than the dichotomous outcome.
As discussed above, this study is a re-analysis of Buffler *et al.* (2006), not Kelsh *et al.* (2003). OEHHA agrees that sometimes the use of continuous data can provide a more convincing argument regarding causality, and considered analyzing these data in this way. However, the use of continuous variables requires *a priori* assumptions or *a priori* hypotheses regarding the shape of the dose-response curve. If these assumptions or hypotheses are wrong, true associations could be misrepresented or missed. One could test many different dose-response models and ultimately select the one that best fits the data, but this type of multiple comparisons increases the possibility of Type 1 statistical error. Perhaps an even greater problem is that in researching the community water systems in California it became very clear that the water distribution systems in many parts of this state are incredibly complex. Because of the extensively overlapping and continually changing water sources in many parts of California, assigning a single perchlorate concentration to each individual would have introduced considerable misclassification. This could have introduced strong bias, particularly if those subjects in the upper ranges of community perchlorate concentration were misclassified. For these reasons, communities (and the subjects who lived in those communities) were divided into two groups based on whether it was more or less likely the sources of their residential drinking water tended to have perchlorate concentrations greater or less than 5 ppb. This method was not only chosen by the authors of Steinmaus *et al.* (2010), but also independently by the authors of Buffler *et al.* (2006). Some exposure misclassification was still likely with this method of classification, but this would have biased results towards the null. Importantly though, most results in Steinmaus *et al.* (2010) were not null. Rather, statistically significant elevated odds ratios were seen, providing evidence that the Steinmaus *et al.* (2010) method of classifying exposure and the resulting misclassification were not strong enough to mask this evidence of association. This is now explained in the PHG document (page 48). Finally, this study included almost 500,000 infants, so lack of statistical power was not an issue in most analyses.

Steinmaus *et al.* (2010). Also of note is that this analysis adjusted for many more factors (potential confounders: age at sample collection, gender, mother's age, income, race/ethnicity, birth weight, and feeding type) than the studies reviewed previously calling into question whether previous study findings were due to residual confounding. Last, when providing an OR, please provide the precision of that OR (95% confidence interval) rather than just the p-value.
We now state that adjustments for these factors in Steinmaus *et al.* (2010) caused little change in the results. This does not provide evidence that the findings from other positive studies were due to residual confounding. Instead, it provides evidence of the opposite: that while these factors may be related to thyroid hormone levels, they are probably not associated strongly enough with both thyroid hormone levels and perchlorate exposure to cause substantial confounding. Confidence intervals are now provided (page 48).

Steinmaus *et al*., 2010. This study provides evidence that perchlorate in drinking water is associated with a high TSH level in California newborns. It also provides a very nice summary of the possibility that nitrate, thiocyanate, and iodine could confound the observed association. The relationship of nitrate, thiocyanate, and iodine with perchlorate and how they may confound and/or interact with perchlorate in affecting fetal TSH levels needs to be discussed in this document. The conclusion of this paper even states "*Further research is needed on this issue, and needed to evaluate the possible role that iodine, thiocyanate, nitrate, and other thyroid-active agents may have played in these findings*." Therefore, it seems necessary to at minimum include a discussion of these other toxicants and if and how they may interact or confound the associations and inference described in the document.

Confounding is reviewed in many of the preceding comments. It should be noted that the importance of iodine, thiocyanate, and nitrate were raised in the conclusion of this paper not because of their role as confounders (which was deemed unlikely), but because of their role as agents that might act cumulatively with perchlorate. This is a much different issue than confounding. Regardless, whether or not these types of cumulative effects occur has no effect on the fact that perchlorate-TSH associations were identified in this study.

Steinmaus *et al.* (2010). It is clear that many of the arguments and second hand calculations presented in this document about previous studies were first done as part of this Steinmaus et al. (2000) paper. This discussion section very precisely describes the many limitation of the previous work. This is the first study, and in fact the only one of the five presented in Table 13, that provides an effect estimate that I am not overly concerned may be due to confounding or other study limitation.

None of the factors that were adjusted for in Steinmaus *et al*. (2010) had any important impacts on the results, either when adjusted for individually or in combination. This does not highlight the importance of these adjustments. Rather, it provides very strong evidence that these adjustments are not important.

Chang *et al*, 2003. Too few details are given for this study. To be consistent with the other studies presented, please describe the study population, how the exposure was defined and measured, how the outcome(s) was/were defined and measured, what statistical analyses were used, what confounders were and were not included in the analysis, and how the briefly mentioned study limitations could explain the results observed. Also, the reference was not provided in the reference section.

These details are now provided and the reference has been added (page 49)

The Amatai *et al.* (2007) study particularly points out that mothers with adequate intake of iodine may not be susceptible to the effects of perchlorate, again suggesting the need for a discussion of this relationship in the document.

The assessment of iodine in this study was minimal, and an assessment of high versus low iodine status was not done. Therefore, it is OEHHA's judgment that this study does not provide any valuable evidence regarding the effects of high or low iodine.

I suggest you include the Cao *et al.* (2010) study in this review, as it assessed the association of urinary perchlorate, nitrate, thiocyanate, and iodide with urinary T4 and TSH levels in infants.
We now describe this study (page 51). However, a potentially major bias (the use of urinary creatinine concentration on both sides of the regression equation) limits the interpretation of its results.

The title of the next section should perhaps be changed. As I understand it, this document is attempting to summarize the information on whether fetal exposure to perchlorate (drinking water perchlorate) is associated with fetal TSH and T4 levels (first 24 hour TSH or T4 levels used as a proxy). If correct, the title should be changed to state this.
The reviewer's implication is correct that one of the important issues raised in this section is whether or not the positive effects identified in these studies represent effects in the fetus, and this issue is reviewed on page 63. But, because all of the studies being summarized in this section involve measuring thyroid hormone levels in the infant, not in the fetus, OEHHA feels it is most appropriate to title this section, "Summary of Studies of Perchlorate and Infant Thyroid Hormone Levels" (page 51).

The summary inference to be made from this section appears to be based on Table 13, which is based on only five studies. I agree with the authors that the Tèllez Tèllez et al (2005), Li et al (2000b), and Amatai et al (2007) studies should not be included as they do not assess this association in the time period the authors argue is most important. However, of these 5, only one uses an exposure measure other than 'city with high perchlorate levels in drinking water' versus 'city with low or no perchlorate in drinking water' (Steinmaus et al 2010).
The use of "city" as an exposure metric as the reviewer describes would cause a non-differential misclassification of exposure that would bias the results of these studies towards the null, not towards the associations identified. As such, the use of "city" as a proxy for perchlorate exposure does not by itself invalidate the fact that associations were identified. It is true that errors in exposure may affect the magnitude of the association, but these studies are not being used to define exact dose-response relationships. Instead, they are being used simply to evaluate whether or not an association exists. Since positive associations were found in each of these five studies, arguments about bias from non-differential misclassification of exposure in these studies are most likely moot. As summarized by Rothman and Greenland in their discussion of this topic, "in studies that describe a strong nonzero effect, preoccupation with nondifferential exposure and disease misclassification is rarely warranted…" (Rothman and Greenland, 1998).

Further, of these 5 studies, given the limitations in the studies noted, and the lack of confounder control present in most of these studies, I am not convinced that

these studies, by themselves, suggest a clear effect of maternal perchlorate levels in drinking water on adverse changes in fetal TSH and T4 levels. They are clearly suggestive of such an effect, but too many limitations in them could be argued to explain the effects. Ideally, a prospective study based on this summary document could be designed and done to more properly assess this association.
OEHHA agrees that a prospective study in a sufficiently large population of pregnant women and infants, in a population with a wide range and stable source of perchlorate exposure, with adequate collection of data on potential confounders and effect modifiers might provide valuable evidence of an association. However, it is important to keep in mind the major goal of reviewing these studies: to see if there is evidence in the current literature that some people may be more susceptible to perchlorate than the healthy adult volunteers assessed in the Greer *et al*. (2002) study. This information is helpful in assessing whether or not an uncertainty factor should be applied to the Greer *et al.* (2002) data to help account for the possibility that susceptible populations might exist. OEHHA feels that these five studies provide some evidence that infants or the fetus may be more susceptible to perchlorate than "normal" healthy adults. OEHHA also feels that these findings, *in combination* with a substantial amount of other evidence that susceptible subgroups exist (summarized on page 117-119), provides strong justification that an uncertainty factor needs to be applied to results from the normal healthy adults used in Greer *et al*. (2002).

Appropriately, the authors attempt to summarize the limitations of the studies used in this section. The argument for why many of these potential confounders could not likely explain the results observed is not convincing. I would suggest you provide quantitative examples of how much a confounding variable could alter an effect estimate (similar to Table 22 in the document). Perhaps date from the existing Steinmaus et al. (2010) study could provide this. Without it, the argument to use the second-hand, unadjusted effect estimates calculated for several studies using descriptive data summaries in paper tables, in Table 13, is not a strong one. Second, a discussion of how outcome misclassification could have biased the effect estimates presented is warranted.
Both of these have now been done and are discussed in previous comments

Braverman et al. 2005. This study was adequately described and limitations noted. However, in line 15, I suggest you present both the mean pre- and post-shift levels of the workers, rather than just the pre-shift level.
These data are now given (page 71).

Gibbs and Landingham, 2008. The actual effect estimates described in the text should be presented, not just that which is statistically significant. Further, do not just present a "b" which I assume is the parameter estimate. If the parameter estimate and its standard error are provided in the paper, take the parameter estimate and convert it to a meaningful effect estimate and 95% confidence interval , even if the authors of the paper do not.
We now state that the authors only provided coefficients for statistically significant associations (page 72). With regards to interpreting the coefficient: these data were

published only as a Letter to the Editor and not as a full report. Given the complexity of the analyses performed (e.g., fT4 was entered as 1/Sqrt(fT4); perchlorate-iodine interaction term expressed as perchlorate x iodine when these variables are hypothesized to have opposite effects of fT4) and the lack of detail on exactly how they were performed, a meaningful interpretation of the coefficient could not be derived. This is now stated (page 72).

Blount *et al.* 2006 and Steinmaus *et al.* 2007 are described together as they both used data from 2001-2002 NHANES survey. The study design features of Blount et al. (2006) are adequately described. However, the statements where associations are described as "statistically significant" without providing the direction of the effects observed is misleading. Also, how clinically important are effects these size?
The directions of the associations are now provided (page 73). The potential clinical relevance of these findings is also now discussed (page 73-74).

Andrea B. Kirk, Ph.D.
University of North Texas Health Sciences Center

Perchlorate concentrations are substantially higher in breast milk. Why was human milk not used as the basis for the public health concentration, since breastfed infants may be most vulnerable to perchlorate, given evidence of low iodine in many human milk samples? One strategy might have been to estimate what concentration of perchlorate in drinking water would result in a concentration of perchlorate in breast milk that would pose no risk to infants.
OEHHA agrees that perchlorate appears to concentrate in breast milk and this could increase the susceptibility of the infant to perchlorate (Valentín-Blasini *et al.,* 2011). However, there are few studies (e.g., Dasgupta *et al.,* 2008) and only small sample sizes on which to estimate the concentration of perchlorate in drinking water that would result in a concentration of perchlorate in breast milk that would pose no risk to infants. This lack of robust data would lead to significant uncertainty about the accuracy and generalizability of these estimates.

Is a relative source contribution of 73% water reflective of California, especially given the high likelihood of use of perchlorate-contaminated irrigation water relative to the rest of the country? Of course it is possible that foods are so widely distributed that local conditions do not matter, but it seems that this is an area of uncertainty. It is hoped that the conservative approach used by OEHHA would more than cover any errors or uncertainties in relative source contribution.
Data that is specific to perchlorate intakes from foods in California are not available. We agree that this is a source of uncertainty and that the uncertainty factor applied by OEHHA in its PHG calculations likely covers some of this uncertainty.

There are at least two studies which do not appear to have been included in the PHG (Cao *et al.* 2010 and Valentin-Blasini 2011) which would provide additional support for the argument that infants are a vulnerable subpopulation.

These studies have now been added (pages 11 and 51).

OEHHA has considered perchlorate exposures with potential to alter function or structure, and addressed the issue through selection of a level of exposure at which 5% inhibition of iodine uptake would occur. This is a reasonable point of departure, although it is not clear that this (5% inhibition) would be the same level for individuals across the board, regardless of the level of iodine intake.
We agree that it is possible that there is interindividual variability in the dose-response relationship between perchlorate and thyroid iodide uptake. OEHHA has applied an uncertainty factor of 10 to help account for this type of variability.

There is not currently a demonstrated safe dose-response threshold for this contaminant. While there are a number of epidemiology studies showing no associations between perchlorate exposure (or presumed exposure) and thyroid hormone parameters they are not sufficiently strong to justify their use in determining a safe dose. The Brechner study, showing a positive association, is limited in the same way. The Blount and Pearce studies, showing associations between direct measures of perchlorate exposure and direct measures of thyroid hormones, are quite interesting and highlight the need for continued research in this area.
We agree that further research could add important new information on the exact health risks of perchlorate in susceptible populations. Regardless, as discussed throughout the PHG document, we believe there is sufficient evidence that some people are more susceptible to perchlorate than others and some subgroups could be more susceptible than the healthy volunteers in the Greer *et al.* (2002) study. For this reason, OEHHA has used its standard practice of applying an uncertainty factor of 10 to account for this variability.

Heightened focus on iodine nutrition, especially among pregnant women and infants will do much to protect people from perchlorate and other iodine uptake inhibitors. More information on the vulnerability of fetuses, neonates and infants to TH-disruption would be key to developing a more precise estimation of a safe exposure level for perchlorate. In the interim, a drinking water concentration of 1 ppb is likely protective to the population.
Efforts to assure adequate iodine nutrition is beyond the scope of this PHG. OEHHA agrees that further research might help refine the risks in susceptible populations like those with low iodine. Until these data are available, OEHHA's standard approach is to use a 10-fold uncertainty factor to account for possible interindividual differences in perchlorate toxicity.

Concentrations of perchlorate in water used for irrigation should also be addressed.
OEHHA has been unable to identify sufficient information regarding perchlorate in irrigation water, and assessing the risks of perchlorate concentrations in irrigation water is beyond the scope of this PHG.

References:

Barr DB, Wilder LC, Caudill SP, Gonzalez AJ, Needham LL, Pirkle JL (2005). Urinary creatinine concentrations in the U.S. population: implications for urinary biologic monitoring measurements. Environ Health Perspect 113(2):192-200.

Corneli HM, Gormley CJ, Baker RC (1985). Hyponatremia and seizures presenting in the first two years of life. Pediatr Emerg Care 1(4):190-193.

David R, Ellis D, Gartner JC (1981). Water intoxication in normal infants: role of antidiuretic hormone in pathogenesis. Pediatrics 68(3):349-353.

Ershow AG and Cantor KP (1989). Total Water and Tap Water Intake in the United States: Population-Based Estimates of Quantities and Sources, Federation of American Societies for Experimental Biology, Bethesda, MD.

Furnee CA, van der Haar F, West CE, Hautvast JG (1994). A critical appraisal of goiter assessment and the ratio of urinary iodine to creatinine for evaluating iodine status. Am J Clin Nutr 59(6):1415-1417.

Garde AH, Hansen AM, Kristiansen J, Knudsen LE (2004). Comparison of uncertainties related to standardization of urine samples with volume and creatinine concentration. Ann Occup Hyg 48(2):171-179.

Keating JP, Schears GJ, Dodge PR (1991). Oral water intoxication in infants. An American epidemic. Am J Dis Child 145(9):985-990.

Levallois P, Gingras S, Caron M, Phaneuf D (2008). Drinking water intake by infants living in rural Quebec (Canada). Sci Total Environ 397(1-3):82-85.

Medani CR (1987). Seizures and hypothermia due to dietary water intoxication in infants. South Med J 80(4):421-425.

Nermell B, Lindberg A, Rahman M, Berglund M, Åke Persson L, El Arifeen S, Vahter M (2008). Urinary arsenic concentration adjustment factors and malnutrition. Environ Res 106 (2):212-218

OEHHA (2012). Air Toxics Hot Spots Program Risk Assessment Guidelines. Technical Support Document for Exposure Assessment and Stochastic Analysis. Chapter 8. Water Intake. Office of Environmental Health Hazard Assessment (OEHHA). California Environmental Protection Agency (CAL/EPA). April, 2012

Rodriguez-Soriano J, Vallo A, Castillo G, Oliveros R (1981). Renal handling of water and sodium in infancy and childhood: a study using clearance methods during hypotonic saline diuresis. Kidney Int 20(6):700-704.

Rothman and Greenland, 1998. Precision and Validity in Epidemiologic Studies. In: Modern Epidemiology, 2nd Edition, Rothman and Greenland (eds), Lippincott-Raven, Philadelphia, page 132.

Thomson CD, Smith TE, Butler KA, Packer MA (1996). An evaluation of urinary measures of iodine and selenium status. J Trace Elem Med Biol 10(4):214-222.

Vejbjerg P, Knudsen N, Perrild H, Laurberg P, Andersen S, Rasmussen LB, Ovesen L, Jørgensen T (2009). Estimation of iodine intake from various urinary iodine measurements in population studies. Thyroid 19(11):1281-1286.