

# Evaluation of Perchloroethylene Risks and the Philadelphia Dry-Cleaning Proposal

Adam M. Finkel, Sc.D., CIH

[Professor of Environmental and Occupational Health, UMDNJ School of Public Health *and*  
Executive Director, Penn Program on Regulation, Univ. of Pennsylvania Law School]

*\*Personal views only—affiliations provided for purposes of identification\**

Report to the Philadelphia Department of Public Health,  
Air Management Services

January 13, 2010

## Executive Summary

The city of Philadelphia has proposed to accelerate the federal (EPA) phase-down of perchloroethylene (Perc) use in dry-cleaning establishments that are co-located with day care centers, residences, and commercial establishments. Among other provisions, the proposed regulations would require operators of dry-cleaning establishments to take corrective action if the Perc concentration in co-located living or working spaces exceeds 40 parts per billion (ppb). The key questions for gauging the wisdom of the regulations include: how dangerous to human health is Perc, especially above the 40 ppb benchmark level? Can the 40 ppb level be achieved with existing and affordable technologies? What are the possible side-effects (“risk-risk tradeoffs”) of the proposal? As an expert in quantitative risk assessment, and a former chief of the regulatory divisions at the U.S. Occupational Safety and Health Administration (OSHA), I was asked to help shed light on these and other questions.

Uncertainty is the hallmark of all risk assessments. Statements that a substance like Perc can cause a particular human health effect should always be interpreted to mean “how confident can we be that the substance can or cannot cause the effect?” Similarly, quantitative statements that “the risk at this concentration of the substance is one chance in X of suffering the ill effect” are surrounded (whether acknowledged or not) by uncertainty. And when expressed in the opposite terms (pronouncements that a given concentration of the substance is “safe”), it is important to understand that the best science can say is that such a concentration is associated with a low (and uncertain) probability of causing harm. I have reviewed the most recent and definitive reviews of the non-cancer toxicity and the cancer potency of Perc, along with more recent peer-reviewed articles that post-date these reviews, and have reached various conclusions about the degree of risk and the uncertainties therein:

With respect to whether Perc can cause human health effects other than cancer (qualitative risk):

Perc unequivocally can cause various toxic effects, notably neurologic ones, in both experimental animals and humans.

*With respect to whether Perc can cause cancer in humans (qualitative risk):*

Perc is associated with a spectrum of malignant tumors in at least ten different laboratory bioassays (including both male and female rats and mice). Unequivocal positive evidence of excess cancer rates in humans exposed to Perc is limited, but of concern due to the multiplicity of suggestive findings and the similarities in some cases between the tumor types observed in rodents and in humans.

Compared to the dozens of other substances classified as “likely” human carcinogens by EPA and “reasonably anticipated to be human carcinogens” by the U.S. National Toxicology Program (i.e., substances with clear evidence in animals and limited evidence in people), I would place Perc near the top of this spectrum towards “known” status, because of the breadth and consistency of the responses.

Various mechanism-of-action theories have been advanced to cast doubt on the relevance to humans of some of the categories of animal tumors; however, each of these explanations at present fails to completely explain the observed positive findings, and is contradicted by some information that proponents of the theory fail to acknowledge.

Various objections have been raised to some of the suggestive human epidemiologic studies. These studies do suffer from many of the limitations common to retrospective analyses of common diseases with long latency periods. However, the existence of “negative” human studies does not in itself cast doubt on the various positive findings; in particular, the 2006 Nordic study (Lynge et al.) is not a powerful negative study, and has its own flaws and limitations.

*With respect to quantitative cancer risk:*

EPA estimates that the cancer potency of Perc falls between  $1.4 \times 10^{-5}$  and  $1.4 \times 10^{-4}$  per ppb; in other words, that an average human exposed to (say) 40 ppb continuously over a lifetime would have between 6 chances in 10,000 (that is, 40 times ( $1.4 \times 10^{-5}$ )) and 6 chances in 1,000 of developing cancer, above and beyond the background risk in the population. This range is between 600 and 6000 times higher than the 1-in-1-million benchmark of *de minimus* risk commonly sought as a goal by EPA.

I conclude, for reasons explained in detail in the text, that EPA’s estimates are not especially “conservative” (that is, they may overstate risk but not by a large amount), and that there are reasons to be concerned that they may understate true risk for some portions of the human population. However, I offer two caveats to this statement: (1) the *upper* end of the range may be slightly less plausible in light of the Nordic human study (in other words, if Perc was this potent a carcinogen, that study might have been powerful enough to show an excess of tumors at the levels to which the subjects were exposed); and (2) there is some logic in this case (though some pitfalls as well) in presenting an alternative exposure estimate, such as 8 hours/day for 10 years, rather than continuous lifetime exposure—that adjustment would lower both ends of the risk range by a factor of 21.

However, even using the middle of the EPA range (a potency estimate very close to the one calculated by the California EPA) and the alternative exposure assumption, the excess risk at 40 ppb would be approximately *1 in 10,000*, which is at or far above the highest benchmarks of “acceptable risk” used in all EPA programs.

A forthcoming new study raises concern that residential exposure to emissions from dry cleaners may be associated with an increase in kidney cancer (one of the tumor types found in the bioassay data).

*With respect to quantitative non-cancer risk (“safety” of the 40 ppb level):*

EPA's "Reference Concentration" (RfC) is 2.4 ppb, and is based on a study of neurological deficits (increased reaction time, etc.) in a group of residents exposed to Perc for an average of 10 years. The Agency for Toxic Substances and Disease Registry (ATSDR) has set a "Minimal Risk Level" (MRL) of 40 ppb based on a study of similar neurological deficits in a group of workers exposed to an equivalent of roughly 4 ppm Perc on a continuous basis.

For many reasons detailed in the text, one should not assume that the RfC or MRL are "safe" levels; rather, they are intended to approximate the levels that may produce a small effect (one that is not detectable among a small group of human subjects), adjusted so that they apply to other humans who are more susceptible than the average. Either the RfC or the MRL, therefore, may produce toxic effects in a small minority of persons so exposed.

The 40 ppb benchmark proposed by Philadelphia is the same as the ATSDR's MRL, and 17 times higher than EPA's RfC. There are scientific reasons to prefer the EPA estimate, which comes from a residential rather than a worker investigation (for example, continuous lower-level exposure *may* be more dangerous to the neurologic system than an equal total exposure accruing only during the workday and workweek).

New peer-reviewed articles, published after the EPA finished its internal review of Perc, raise new concerns about other and more dire neurological effects possibly associated with Perc, and about possible effects on the human immune system.

I have also reached various conclusions about the effectiveness of controls, the risk of "side-effects," and the benefits of considering a few modifications to the draft regulations, as follows:

- installation of engineering controls can successfully reduce Perc concentrations in co-residential settings on average by about a factor of 10, from well above the 40 ppb benchmark to reliably (slightly) below it;
- establishments that continue to use Perc will find it difficult to ensure co-residential concentrations at the lower EPA benchmark (2.4 ppb);
- concentrations at adjacent workplaces can be similar, if not somewhat higher, than at adjacent residences;
- some "risk-risk tradeoffs" mentioned as detrimental effects of the proposed regulations, notably the assertion that vehicle miles traveled (and carbon monoxide emissions) will spike upward as a result, are too indirect and speculative to significantly reduce the expected health benefits of the proposed rulemaking. However, it is easy to imagine a direct and highly unfortunate offsetting risk, if cleaners substitute a more toxic solvent for Perc. Of the available alternatives, one in particular—*n*-propyl bromide (a.k.a. 1-bromopropane)—is clearly more dangerous than Perc (according to a new cancer bioassay and various human case reports of irreversible neurological damage), and should be banned as a substitute.

- I recommend that Philadelphia consider a “hybrid” regulatory design, in which establishments would be given the choice of trying to achieve the 40 ppb benchmark via the specific measures spelled out in the regulation, or via any alternative technologies or practices each establishment believes will be equally or more effective at (presumably) lower cost. If, however, an establishment chooses to innovate, the monitoring requirements must be more frequent.
- I urge the city to consider establishing equally-stringent requirements for co-commercial and co-residential settings, on the grounds that workers may be exposed for more hours per day than residents, and that each co-located business will likely contain more inadvertently-exposed citizens than a residence would.
- Employers should be aware that in analogous cases, it is possible that controls designed to lower concentrations in adjacent establishments will have the unfortunate effect of *raising* levels within the cleaners and further increasing risks to dry-cleaning workers.

## **1. ISSUES:**

The Department of Public Health requested an analysis of the following issues:

1. The possible cancer risks from perchloroethylene (Perc)—including the weight of evidence that Perc is a human carcinogen, and data leading to quantitative estimates of its cancer potency and corresponding estimates of cancer risk at various exposure levels;
2. Selected non-cancer risks from Perc, and an appraisal of the “safety” of the two most prominent non-cancer benchmark exposure levels – the “Reference Concentration” (RfC) of 2.4 parts per billion (ppb) set by the U.S. Environmental Protection Agency (EPA) and the “Minimal Risk Level” (MRL) of 40 ppb set by the Agency for Toxic Substances and Disease Registry (ATSDR);
3. Evidence from various interventions by state and local agencies that bear on whether a given concentration can be readily achieved in residences and workplaces co-located with dry cleaners;
4. Concerns about new risks that might arise from the act of regulating current uses of Perc in dry cleaning; and
5. Information about alternative regulatory designs that might reduce risks in more cost-effective and flexible ways.

AMS specifically asked that the quantitative risk analysis focus on the appropriateness of using 40 ppb as a level of regulatory concern in co-located residences and establishments.

This report will review each of the five issues above in turn, with special attention to three “cross-cutting themes” (NAS 1994, Chapters 6-11) that are fundamental to interpreting risk data in ways that are not misleading: (1) uncertainty in the implicit or explicit assumptions used to estimate risk (with emphasis on “default” assumptions that may be inferior to more specific inferences developed for Perc); (2) uncertainty in the estimates of key inputs to risk

assessment models, which determines whether point estimates of risk are overly “conservative” (i.e., are likely to exaggerate the true risk) or are insufficiently precautionary to protect human health to a reasonable level of confidence: and (3) variability in risk across the human population, which determines whether estimates of risk under-predict or over-predict true risk to specific individuals or to persons with defined susceptibility or other characteristics that modify their risk.<sup>1</sup>

## **2. BACKGROUND:**

In addition to having worked for the past 25 years developing and improving methods of quantitative risk assessment (QRA) and cost-benefit analysis, with several books and over 50 articles in these fields, I have some specific experiences relevant to this topic. I supervised the development of OSHA’s 1997 regulation on methylene chloride (MC), which contained one of the most sophisticated risk assessments (incorporating detailed information on the pharmacokinetics of MC metabolism in rodents and humans) yet undertaken by a federal health agency. As OSHA’s Director of Health Standards Programs from 1995-2000, I developed public-private partnerships in lieu of regulation with major industries that manufacture and/or use certain toxic materials (e.g., fiberglass insulation, styrene, refractory ceramic fibers), and had begun to develop a similar partnership with the former International Fabricare Institute during the last year of the Clinton Administration. I was the Department of Labor’s voting representatives on the National Toxicology Program Executive Committee during this period, and participated in decisions listing, delisting, and modifying various entries in the NTP’s Report on Carcinogens.

## **3. CANCER HAZARD AND RISK OF PERC:**

---

<sup>1</sup> Risks are inescapably “three-dimensional”: the common shorthand that “the risk at this exposure level is one chance in X” hides the fact that at best, all science can say is that “we are Y% confident that the risk is less than one chance in X for Z% of the exposed population.” Similarly, “the exposure is below a level of concern” should be understood to mean “we are Y% confident that Z% of the population will be safe at this level.” For example, the statement that “one aspirin tablet per day is probably safe for persons over 18 to consume on a regular basis” expresses both some of the uncertainty (“probably”) and the variability (children being at special risk of developing Reye syndrome and other side-effects) common to all risk estimates but often obscured by overly-simplistic conclusions.

A. Hazard Identification (qualitative risk assessment):

(i) *Animal Studies:*

Perc has been listed for the past 20 years as “reasonably anticipated to be a human carcinogen” by the National Toxicology Program. EPA lists Perc as “likely to be carcinogenic in humans,” the second most definitive category of the five classes it uses. These listings are based on a very extensive set of findings in experimental animals that amply qualify as “sufficient evidence of carcinogenicity” in rodents. Perc administered by inhalation caused statistically significant increases in malignant (and, in some cases, benign)<sup>2</sup> tumors at the following sites (“NTP” denotes the 1986 bioassays in both rats and mice, and “JISA” denotes the 1993 bioassay in rats conducted by the Japan Industrial Safety Association):

- Liver carcinomas (and adenomas) in both male mice and female mice (NTP and JISA);
- Liver and spleen hemangiosarcomas in male mice (JISA);
- Mononuclear cell leukemia in male and female rats (NTP and JISA);
- Kidney tubular cell carcinomas (and adenomas) in male rats (JISA); *and*
- Brain gliomas in male rats (NTP)

Although (see subsection (iii) below) there exist alternative mechanism-of-action theories that might explain how some of these findings may not be relevant to human risk, if taken at face value this spectrum of carcinogenic to effects is noteworthy both for its breadth (many chemicals stringently regulated by EPA and OSHA only cause rodent tumors in one or perhaps two sites) and for the “low” doses at which widespread malignancies occurred in the laboratory. In some cases, more than half of the rodents exposed to 200 ppm for only six hours per day, five days per week (equivalent to 36 ppm of continuous exposure) developed malignancies—so the archaic OSHA Permissible Exposure Limit (PEL) of 100

---

<sup>2</sup> EPA’s Guidelines for Carcinogen Risk Assessment have recommended for decades that risk assessors consider both malignant tumors and other tumors known to progress to malignancy, on the grounds that these latter lesions may manifest themselves in humans as malignancies.

ppm for Perc therefore involves no “extrapolation” to lower doses whatsoever, but is within the actual dose range found in the laboratory (this is in my experience an unusual and very worrisome state of affairs).

(ii) *Human Studies:*

Human epidemiologic studies of Perc reveal several areas of concern for cancer, although no agency or expert body has as yet pronounced the totality of this evidence “sufficient.” A good summary of the various kinds of studies (cohort and case/control, occupational and community, inhalation and drinking water ingestion, etc.) and their results can be found in Appendix B to Chapter 4 of EPA (2008). Basically, several cohort studies have shown roughly a doubling of certain cancer risks (notably, esophageal cancer and Hodgkin’s disease) among dry-cleaning workers and among workers in degreasing exposed to Perc. More limited evidence of an association has been found with respect to cervical, bladder, kidney, and lung cancer as well. As discussed in the next section, epidemiology can be vague, “noisy,” and imprecise for diseases like cancer with long latency periods and multiple causative factors, especially so when the kinds of tumors that appear to be elevated among exposed persons are commonly found among unexposed persons as well. Many substances other than Perc, particularly those known to cause relatively rare tumors, have more definitive results from human studies, results that can less readily be explained as due to random chance, confounding, or misclassification. Nevertheless, in my experience the number of independent human studies showing one or more troubling results, especially when combined with the varied, significant, and in some cases related positive effects in laboratory animals, places Perc near the upper (relatively more definitive) end among the continuum of those substances regarded as “reasonably anticipated” carcinogens because of strong animal evidence and suggestive human evidence.

Large, well-conducted epidemiologic studies yielding “negative” results can certainly cast doubt on one or more positive, yet not definitive, studies. Although in general terms, credible positives should outweigh credible negatives (truly benign substances never cause cancer, whereas a carcinogenic substance can fail to show an effect due to chance or other factors), it is also true that enough negative evidence can lead to specific explanations that invalidate or overturn what would otherwise remain troubling signals of harm. *But the much-discussed “negative” study of Lynge et al. (2006) does not repudiate any or all of the*

*other epidemiologic evidence yielding associations with Perc exposures.* I will discuss the Lynge study quantitatively in Section 3C below. In qualitative terms, although it is a very large and ambitious undertaking, and although the authors performed a very valuable service by controlling for smoking and socioeconomic status, is important to appreciate four major limitations of this study:

- It does not examine the relationship between exposure to Perc and cancer, but like all case-control studies, examines the relationship between cancer and exposure to Perc. This is not a semantic distinction: a “negative” case-control study is one that concludes that cancer victims and disease-free subjects are equally likely to have been exposed to the agent in question, evidence that bears upon causation but is “after the fact.” If you put 100 drunk drivers and 100 sober drivers in their cars, you expect to see more crashes among the former group (this would be a cohort study of a true risk factor) – but if you examine 100 crash victims and compare them to 100 drivers who didn’t crash, and you don’t find more drunkards in the first group, it doesn’t necessarily mean that drinking is safe.
- More than just after-the-fact design, this particular study really compares the odds that a cancer victim will have been *heavily* exposed to Perc with the odds that s/he will have sustained less, but not zero, exposure (the “unexposed” subjects also worked in dry-cleaning establishments but were assumed to have been doing tasks not in close proximity to Perc sources). If the “unexposed” were in fact “less exposed,” then the already-difficult task (see Section 3C below) of detecting small increases in odds would be exacerbated, and the “negative” findings would say at most that “working right next to Perc is not obviously worse than working around Perc.”
- Most importantly, even at face value there are positive well as “non-negative” results in the study. Table 5 in Lynge shows that two (of the eight total) tumor sites, bladder and pancreas,<sup>3</sup> showed increased odds of being associated with exposure versus “non-exposure” – and in the case of bladder cancer, that

---

<sup>3</sup> These sites happen to be the two with the largest total numbers of cases, and also happen to be two diseases that are notoriously difficult to treat clinically.

increase is statistically significant.<sup>4</sup> For a third case, non-Hodgkin’s lymphoma (NHL), the authors themselves say that “our results are in line” with previous studies that found 22 total cases of NHL when only 18.8 were expected (i.e., a positive, though perhaps not statistically significant, result).<sup>5</sup> In fact, if instead of disaggregating the eight tumor sites into five “negative,” two “weakly positive,” and one significantly positive subset, one adds together *all* the cases and all the controls, it turns out that the relative risk of exposure with respect to cancer is 1.06 – a very weak positive result, but not at all the same as a “negative” one.

- Finally, it is always crucial to think about the power of any observational study to detect a signal of roughly the strength one would expect to find in the first place. To oversimplify, one would think that alcohol doesn’t affect driving performance if all we had were studies of people who drove after consuming one ounce of light beer—the degree of impairment at this level would not cause enough of an increase in crashes among the “exposed” to be visible unless tens or hundreds of thousands of subjects were followed. *The Lynge study suffers from the same problem.* Figure 2 in Lynge shows how *low* the Perc exposures were in the Nordic countries during the study period—a median concentration of about 30 mg/m<sup>3</sup>, or about 4 ppm. Furthermore, these subjects were only exposed for roughly 8 hr/day, and for an average of about 10 years<sup>6</sup>. And, although the article does not provide information on the age of the subjects at the end of the study, I notice that there were only 1,616 cases of the eight kinds of cancer among the 46,768 subjects—a total of 3.5 percent—whereas the *lifetime* risk (at least in the U.S., according to American Cancer Society 2009) of

---

<sup>4</sup> The authors say this increase doesn’t rise in proportion to years of employment, but duration is of course only a proxy for exposure (two years’ employment at a leaky machine can give much more exposure than four years’ employment around a state-of-the-art machine).

<sup>5</sup> This admission is also evident from the study results themselves: Table 5 shows 42 cases of NHL in the exposed group out of a total of 187 HNL cases, for a slightly negative relative risk (RR=0.95). But if even three of the 145 “unexposed” cases in fact should have been coded as exposed, the relative risk would instead be slightly positive. In other words, even though this is a large study, there is a great deal of uncertainty even about the sign of the relationship once the subjects are divided into so many categories. Note that 3/145 is only about 2 percent, whereas Lynge et al. note that they could not ascertain job title information for 25 percent of all the cases and 19 percent of all the controls—so a misclassification rate of a few percent is probably less than the actual amount of error already present in the study.

<sup>6</sup> I took all the length-of-employment data in Table 6 of Lynge and calculated a weighted average assuming that the “> 10 years” category represented an average of 15 years’ employ—the weighted average across all subjects would only go up to about 15 years if one assumes this highest category represented 25 years.

these eight cancers combined is roughly 12 percent. So it follows that these subjects had not been followed for nearly long enough to develop all the cancers that might manifest themselves (cancer being predominantly a disease of old age). Leaving out this last adjustment entirely, and starting from the EPA range of  $1.4 \times 10^{-2}$  to  $1.4 \times 10^{-1}$  risk per ppm, I estimate that the subjects in the Lynge study faced an exposure of roughly  $[4 \times (10/70) \times (8/24) \text{ ppm}]$ , or *0.2 ppm* continuous equivalent, which would according to the EPA estimate confer an excess cancer risk of between  $2.6 \times 10^{-3}$  and  $2.6 \times 10^{-2}$ . As I will explain below, *I do not believe the Lynge study could possibly have detected an excess risk at the low end of this range.*

*(iii) Perspectives on Hazard Identification:*

It is important to note that NTP and EPA (as well as the International Agency for Research on Cancer and other bodies that classify carcinogens) use a classification system that reserves the “known human carcinogen” designation solely for substances shown via epidemiology to cause cancer in humans. Even when animal evidence is sufficient, when the mechanism of action is understood, and when human *in vitro* data show that humans develop the same mutations and/or produce the same metabolites as the affected animals, epidemiology is the “gold standard” without which a substance can’t be “known.” Relying so heavily on epidemiology in turn has a very important implication: only substances that cause rare human tumors, or increase the incidence of common tumors by an enormous amount, will yield statistically significant positive results, befitting the blunt instrument that observational epidemiology is for cancer and other long-latency diseases. Asbestos, benzene, vinyl chloride (VC) and many other “known” carcinogens are only “known” because they happen to cause rare tumors: if, for example, VC happened to cause colon cancer in humans instead of the rare neoplasm (liver hemangiosarcoma) caused only by VC and a few other substances, it might never have been listed as “known” or regulated.<sup>7</sup> This

---

<sup>7</sup> Suppose a substance causes one extra tumor in every 200 persons exposed (a huge excess risk), and suppose that a “rare” tumor is defined as one that naturally occurs once in every 1000 people, while a “common” tumor occurs once in every 20 people (the actual lifetime background risk of colorectal cancer (ACS, 2009)). If Substance X causes a rare tumor, a (fairly large) epidemiologic study of 1000 exposed and 1000 unexposed persons would be expected to find 6 tumors in the former group (5 caused by X and one “background” tumor) and only one in the latter—a significant finding with a 6:1 relative risk. If, however, X causes a common tumor,

is not to argue that Perc and the other substances with sufficient animal evidence and limited human epidemiology are “known carcinogens we fail to notice as known,” merely to emphasize that *some* of the substances in this category (perhaps Perc; we just don’t know...) certainly are misclassified as “likely” as opposed to “known” due to the limitations of epidemiology.

On the other hand, it is conceivable that Perc is incorrectly classified as “reasonably anticipated” because EPA, NTP, and many other institutions have failed to appreciate that (simultaneously) all of the human cancer evidence amounts to a false positive, and that all of the rodent cancer evidence signals harm for lab animals but not for people. There are, in fact, theories that could explain away the positive findings for many of the individual epidemiologic studies and many of the positive rodent tumor responses—basically, arguing that the human studies are flawed and that the rodent studies, though valid at face value, should not be assumed to have relevance for humans because of fundamental biological or biochemical differences between rodents and people.

First, it is important to realize that when EPA occasionally gives a cancer risk estimate and says “this is an upper bound and the true risk may be as low as zero,” it is not (or at least not correctly) implying that the official estimate is precarious and might be completely wrong. This boilerplate language is a remnant of an earlier confusion on EPA’s part, when the Agency was using the linear term of the multistage model as a proxy for cancer potency, and sometimes found that the statistical lower bound on this term was zero. But it was never the case that a zero value for the linear term meant zero potency—it only meant that the dose-response function was non-linear (convex upward), and that risks at very low doses were *smaller* than a straight-line assumption would have predicted. The most recent National Academy of Sciences study of EPA risk assessment urged the Agency to abandon this misleading boilerplate (NAS 2009, page 206).

On the epidemiologic evidence, skeptics of Perc’s human carcinogenicity (see, e.g., Dugard 2008) accuse EPA and others of a “one-sided” attitude towards the evidence (weighing positive signals more highly than negative ones). But this is exactly the attitude I believe scientists *should* take with respect to chronic disease epidemiology—false negatives can easily arise due to lack of power to detect, whereas false positives cry out for specific

---

the study would be expected to find 55 tumors in the exposed group as against 50 in the control group, and that elevation would not be statistically significant (RR=1.1) given the “noise” of the background process.

explanation and refutation. To argue, as skeptics often do, that “EPA ignores the various studies that did not find associations” is akin to telling astronomers they should ignore instances when they think they are detecting signals from other civilizations because of “all the times when you listen and don’t hear anything.” If the signal is a false one, there will be an explanation.

There are some theories to explain some of the positive human associations as spurious—it is possible, for example, that the observed increase in cervical cancer among dry cleaners may in fact be a manifestation of low socioeconomic status, or that (see Section 3E below) co-exposures to other chlorinated solvents may account for some of the lymphomas observed in Germany. Ultimately, of course, a compelling and varied set of positive bioassays does not need corroboration by the blunter instrument of epidemiology in order to elevate concern about a toxic chemical—but in the case of Perc, there are many different suggestive findings in humans, and this makes more difficult the job of coming up with independent plausible refutations of each one separately.

With respect to the theories that might explain away each of the bioassay results as irrelevant to humans, I believe the proper context has been provided by the recent NAS *Science and Decisions* report (Chapter 6). This NAS committee agreed that EPA should be receptive to alternative theories that contradict one or more of its established “default” models or assumptions, but agreed that it should do so only if the alternative is “clearly superior” to the default. Placing this sort of burden on proponents of new theories or assumptions is not intended to give the defaults special weight because of convenience, but is intended to fairly value the defaults in light of the decades of theoretical underpinning and empirical corroboration each of them has. For example, the default that rodent tumors are presumed relevant to humans *in the absence* of a well-developed explanation to the contrary gives appropriate weight to the many similarities between the two kinds of mammals and to the many cases (in therapeutics as well as in toxicology) where rodent models have performed well qualitatively and quantitatively as surrogates for humans. EPA has not developed a specific set of scientific criteria by which one might judge whether an explanation for species-specificity (irrelevance of rodent tumor response) is compelling, but the NAS committee felt such criteria should probably be roughly of the same level of detail as OSHA developed in 1997 when it agreed with HSIA and others that a “physiologically-based pharmacokinetic” (PBPK) model was clearly superior to body-weight scaling in

evaluating the risks of methylene chloride (OSHA's 11 criteria for helping judge the validity of a PBPK model are presented in Table 1 below).

So are any of the various theories currently being advanced to explain away positive Perc bioassays "clearly superior" to the long-standing default assumptions that, for example, rodent tumors are relevant to people, and that effects seen at high (but not toxic) doses in the laboratory are relevant to what might be observed at lower doses if the sample size of the bioassay was large enough? HSIA (see Dugard 2008) has posited different explanations for some of the major rodent tumor responses, as follows:

- Perc could activate the "peroxisome proliferator activated receptor" known as PPAR- $\alpha$ , causing oxidative stress to the liver resulting in cell proliferation, which could in turn increase cancer risk. Thus, the rodent liver tumors could appear only at high doses of Perc, or only in rodents or other species susceptible to this mechanism of action;
- The kidney tumors could arise via accumulation of the protein  $\alpha$ -2 $\mu$ -globulin, leading to irritation, thence to compensatory cell proliferation, and ultimately to greater risk of tumor formation—again, a phenomenon that could conceivably be limited to high doses and/or to rodents as opposed to humans;
- The strain of rat that NTP and JISA used could be predisposed to develop mononuclear cell leukemia with advancing age, and humans do not develop this exact disease anyway.

Although it is possible that the NAS Tetrachloroethylene committee will shed a different light on these controversies, EPA clearly believes that not only do each of these three explanations fail to completely explain the observed bioassay findings, they are contradicted by information the proponents fail to acknowledge. For example:

- "Several recent studies have expanded the scientific understanding of the PPAR- $\alpha$  mode of action... including the demonstration that PPAR- $\alpha$  activation in hepatocytes induces peroxisome proliferation but not liver tumors" (it is also noted that human cells do respond to agents that activate PPAR- $\alpha$ );

- “The fact that renal tumors have been observed at doses lower than the ones shown to cause the  $\alpha$ -2 $\mu$ -globulin response is inconsistent with this phenomenon being responsible for tumorigenesis.”
- Generic “susceptibility” to MCL does not explain the greater incidence of MCL in exposed rats versus controls, the dose-related increase in the severity of the cancer in individual animals, or the earlier appearance of MCL with higher doses of Perc.

It appears that none of these alternative explanations is fully-developed, and it is important to emphasize that *all* of them would have to be accepted (along with others) in order for Perc to pose a substantially lower (let alone a zero) cancer risk than EPA currently estimates (see Figure 4 and its discussion in subsection [C] below to appreciate the multiplicity of separate bioassay responses that would have to be discarded, and the quantitative effect of basing potency estimates on the site(s) that would remain if one or more of the spectrum of responses was to be deemed irrelevant). EPA is *not* generally recalcitrant to alternative explanations—in addition to having discarded several generic types of rodent responses in light of well-developed alternatives (e.g., certain chemicals causing thyroid follicular tumors in rodents), the Agency has exonerated various substances based on the  $\alpha$ -2 $\mu$ -globulin response (e.g., *d*-limonene, trimethyl pentane)—it simply does not believe that this case has been made for Perc with similar validity and rigor.

In my personal experience, trade associations with commercial interests in a particular chemical have a tendency to trumpet preliminary hypotheses, associated with isolated “findings,” as definitive when in fact they only begin to raise important (or not-so-important) questions about what the true mechanism of action might be. When OSHA was about to conclude its regulation of methylene chloride in 1997, HSIA issued a press release and presented a series of in-press articles, announcing “completed research” purporting to show that “there are no foreseeable conditions under which the carcinogenic effects seen in mice would be expected to occur in man.” When OSHA received extensive public comments on these articles from many independent scientists, typical reactions included: “I have serious concerns about this [DNA single-strand breaks] assay. It is well known that this assay is extraordinarily difficult to standardize and is sensitive only to very high doses of genotoxic compounds... This data, therefore, is certainly not compelling; persuading any independent scientist of its relevance to humans would be difficult” and “This interpretation

of mRNA distribution is profoundly in error and contradicts some of the most well established and fundamental principles of molecular biology.”

In summary, each of these alternatives *at present* is far from “clearly superior” to the standard model of toxicologic risk assessment, and in fact each raises more questions (and introduces more inconsistencies) than it resolves.

B. Cancer Potency (Quantitative Dose-Response Assessment):

In its most recent toxicological profile of Perc, EPA (2008) has conducted one of the most thorough evaluations of cancer dose-response, and the uncertainties therein, I have seen from a regulatory agency.<sup>8</sup> EPA concluded the excess lifetime human cancer risk from continuous exposure to X ppb of Perc falls between  $1.4 \times 10^{-5}$  and  $1.4 \times 10^{-4}$  per ppb. *So if a national or local government wanted to ensure (say) that its citizens faced excess risks of 1 in 100,000 or less from Perc, this result would correspond to an exposure level of between 0.07 and 0.7 ppb, well below the 40 ppb benchmark Philadelphia proposes.* EPA arrived at this range for potency via the following multi-step process:

1. Of the roughly 13 different combinations of (bioassay, male/female, tumor type) that yielded excess tumors in the laboratory, EPA chose the JISA results from MCL in male rats as the most sensitive result, to guard against the possibility that humans exposed to Perc might respond as strongly to it as these animals did (see the end of this section for a discussion of the “conservatism” in this and the other science-policy assumptions EPA made). EPA converted the experimental conditions in this bioassay (0, 50, 200, and 600 ppm) to units of continuous exposure, and thence to exposures to the rats in units of (mg Perc metabolized per kilogram of body weight per day). EPA then fit the linearized multistage cancer dose-response model<sup>9</sup> to the

---

<sup>8</sup> This document was circulated for external review in 2008; although it may change in response to comments, it is the product of many years of effort by EPA and its contractors, and has undergone extensive peer review within the Agency. In my experience, EPA often leaves the “draft” designation on its review documents for years after changes are no longer made to them. EPA has asked the National Academy of Sciences to review the toxicological profile (unreferenced EPA document, 2009), and as of this writing, the final NAS report is expected to be released within the next month.

<sup>9</sup> The LMS model basically fits an upward-curving (convex) function to the data, but one dominated at very low doses by its linear term – it is therefore “linear at low doses” although not necessarily a straight line in the experimental range.

JISA MCL data, and found that the best-fitting function had a linear term (“potency”) of 0.078 per (mg/kg/day).

2. In order to account for the random error in the small groups of animals exposed, EPA then calculated a statistical upper 95<sup>th</sup> percentile confidence limit on the linear term. Although in practice the dose-response software estimates the slope of a function that fits the observed tumor data acceptably well, the theory behind this adjustment can be seen in a different way in Figure 1. The blue diamonds represent the observed tumor data; at 50 ppm (equivalent to 1.4 mg/kg/day) 14 of the 50 rats exposed developed MCL, and so on. But if each rat faced a risk of exactly (14/50), one could run the experiment again and observe somewhat fewer or somewhat more tumors occurring by chance (just as you might flip 50 coins and see 25 heads, or 22, or 28). If only two additional tumors (16 in all) had occurred in this dose group instead of 14, the low-dose slope of the dose-response function would have been estimated as 0.13 rather than 0.078, a factor of 1.6 higher than the face-value estimate. EPA calls the 0.13 value the “ $q_1$ ” or the upper bound on the dose-response slope at low doses.<sup>10</sup>
3. Since EPA estimates an excess cancer risk of 0.13 at an exposure of 1 mg/kg/day, that is the same as saying there is a 10% (0.1) excess risk at roughly 0.8 mg/kg/day. EPA calls this 10% risk level the “ED<sub>10</sub>,” and assumes that risk below this “point of departure” exposure level decreases linearly as exposure decreases. EPA assumes that humans and rodents are equally sensitive to an equivalent average tissue concentration of the active metabolite of Perc, which necessitates that the rodent exposure in mg/kg/day be scaled via the relative body weights of the two species to the (3/4) power—this “allometric adjustment” yields a human-equivalent dose at the point of departure of 0.23 mg/kg/day.

---

<sup>10</sup> I note here that in recent years, EPA has begun to move away from the concept of carcinogenic potency as the low-dose slope found via the linearized multistage model, and instead has begun to use the LMS or similar models to estimate a “point of departure” where the excess risk is roughly 10 percent (see paragraph 3 immediately following this note). However, EPA then assumes a straight line from the “p.o.d.” to the origin, which effectively makes the low-dose slope follow the same continuous function as the LMS model, particularly when the underlying tumor data (as is true in this case) basically fall in a straight-line pattern anyway. But to be completely precise, EPA currently doesn’t estimate the upper bound on the slope, but instead estimates the *lower* bound on the p.o.d. (often calling this the “LED<sub>10</sub>,” or the lower bound on the 10-percent risk-specific dose). The LED<sub>10</sub> in this case falls approximately at the 10-percent risk level using the upper-bound parameterization of the LMS, so the two procedures are really mirror-images of each other.

4. Rather than simply assuming, however, that rats and humans convert Perc into its active metabolite in proportion to their body weights or to some other simple adjustment, EPA has analyzed various different PBPK models that purport to explain in great biological and biochemical detail the speed and extent of Perc metabolism in both species. Figure 2 (original found on page 3-51 of EPA 2008) shows the three PBPK models that EPA compared; each model estimates how many mg/kg/day of Perc metabolites the average human would produce if exposed to a given concentration in ppm. The red diamonds in Figure 2 show the 0.23 mg/kg/day metabolized dose (the “point of departure”) for the most conservative of the three models at low doses (that of Bois et al.) and for the least conservative one (that of Rao and Brown); the former model estimates that this amount of metabolized dose would be produced at a concentration of roughly 0.7 ppm (see the horizontal axis), and the latter model at roughly 7 ppm—a factor of 10 span of “PBPK uncertainty.”<sup>11</sup>
5. Assuming linearity below the ED<sub>10</sub>, if 7 ppm (Rao and Brown) corresponds to a 10 percent risk, then the low-dose potency is roughly (0.1/8) per ppm, or  $1.4 \times 10^{-2}$  per ppm; using the Bois et al. model, the potency is roughly (0.1/0.7), or  $1.4 \times 10^{-1}$  per ppm. Converting from ppm to ppb, this gives the range of from  $1.4 \times 10^{-5}$  to  $1.4 \times 10^{-4}$  per ppb, as mentioned above. At the 40 ppb benchmark, the lifetime excess human cancer risk is derived by multiplying the (risk per ppb) by 40—yielding a range of between  $6 \times 10^{-4}$  and  $6 \times 10^{-3}$ , as seen in Figure 3 (the yellow or upper band). For purposes of comparison, the California EPA (2009) has estimated Perc’s cancer potency using the liver tumor data from the NTP bioassay, and a simplified PBPK adjustment, and arrived at an estimate of  $4 \times 10^{-5}$  per ppb, which happens to fall squarely in the middle of EPA’s order-of-magnitude range (see the gold ⊗ symbol in Figure 3). The Colorado Department of Public Health, using an older EPA potency value but applying an upward adjustment to account for the slightly increased risk of exposures sustained as a child, estimates a potency of  $2 \times 10^{-5}$  per ppb (see the red ⊗ symbol in Figure 3). Finally, EPA (2008) asserts that using a 1993 epidemiologic study of laryngeal cancer, one can compute a potency estimate of  $5.4 \times 10^{-5}$  per ppb, which is very similar to the CalEPA estimate (not shown in Figure 3 for readability)—so there are at least three other point estimates consistent with this

---

<sup>11</sup> Note that EPA has access to the underlying computer codes for all three models, whereas I am only “eyeballing” the relationship from the Figure.

10-fold uncertainty range that EPA asserts. Note especially that all three PBPK models (Figure 2) predict the metabolism *saturates* above roughly 1 ppm (Bois) or roughly 50 ppm (Rao and Brown) – in other words, that it would be incorrect to say (using the Bois estimate, for example) “if 14 percent of those exposed to 1 ppm are supposed to develop cancer, then everyone (“10 times 14 percent”) of those exposed to 10 ppm should have cancer” – since metabolism can’t increase much above 1 ppm according to these models, neither can cancer incidence increase much above 15 percent or so.

### C. How “Conservative” are the EPA Potency Estimate(s)?

One way to assess how “conservative” the estimates are is to consider each of the common complaints about EPA’s purported exaggeration of risk in turn, and explain both the absolute amount of bias each assumption or method might introduce, but also explore whether plausible alternative assumptions might suggest that EPA’s estimates could instead *understate* the true risk.

- *EPA purportedly uses the most sensitive result from all the bioassay experiments.* This is not strictly true, as Table 2 indicates: the JISA male rat MCL potency value of 0.13 is actually the second-largest of the 11 sex-strain-site combinations from the NTP and JISA inhalation studies. Moreover, of the nine estimates lower than 0.13, the least “conservative” of them is a factor of 15 smaller, and the geometric mean of all 11 estimates is  $3.54 \times 10^{-2}$ , or a factor of 3.7 smaller than the EPA estimate – so a less precautionary central estimate of potency would not change the level of concern about 40 ppb by nearly an order of magnitude. As Figure 4 (from EPA 2008) demonstrates, the male rat MCL potency estimates are higher than the other estimates shown in this Figure (which omits the NTP male-rat-MCL data); choosing any other site would still yield risk estimates within an order of magnitude of EPA’s official estimates.<sup>12</sup> On the other hand, the act itself of

---

<sup>12</sup> Having generated the potency estimates in Table 1 myself, I believe there is a typographical error in Figure 4 (originally Fig. 6-4 in EPA 2008): where the second entry reads “male mouse, hepatocellular tumors and hemangiosarcomas,” it should read “hepatocellular carcinomas and adenomas.” EPA apparently combined the malignant and benign tumors of primary liver cells (as is its default), and did *not* combine two different tumor

choosing a single tumor site is arguably an “anti-conservative” step. Since we have no basis for expecting Perc to produce only a single kind of tumor in humans, many expert bodies (see, e.g., NAS 1994, Chapter 11, “Aggregation”) have urged EPA to combine potencies across some or all tumor sites (not by simply adding slope factors, but either by using the total number of control *versus* exposed animals with one or more tumors, or by combining potencies using Monte Carlo simulation). I was unable to locate the individual-animal tumor data from the JISA study, but the NTP technical report does present that information. Among female rats, for example, 30 of the 50 control animals had a tumor in at least one site, whereas 39 animals in the 200 ppm group and 38 animals in the 400 ppm group had at least one type of tumor: this yields a  $q_1^*$  value of 0.16 per (mg/kg/day), slightly higher than the EPA estimate of 0.13. A Monte Carlo analysis to aggregate site-specific potency values would certainly yield an even higher estimate.

- *EPA uses the statistical upper bound on the dose-response slope.* This is certainly the case; however, the amount of “conservatism” inherent in this technique is itself routinely exaggerated. Even when the dose-response data are highly non-linear, it is statistically impossible for the 95<sup>th</sup> percentile UCL on the linear slope at a given low dose to be more than roughly four times the expected-value estimate for the slope at that dose (Hattis and Goble 1991). For the JISA MCL data, which follow a generally linear pattern (see Figure 1), the amount of “conservatism” is far less than this—the EPA estimate of 0.13 per (mg/kg/day) is only about 60% higher than the maximum likelihood estimate (face-value curve-fitting) of 0.078 (see above). On the other hand, the 99<sup>th</sup> percentile potency estimate, which I calculated (using the linearized multistage curve-fitting software “MSTAGE,” courtesy Edmund Crouch) as 0.33, is nearly three times higher than the “official” EPA estimate. The choice among an estimate that has a 50 percent chance of underestimating the true value, one that has a 5 percent chance of doing so, or one that has a 1 percent chance of doing so is a matter of policy – none of

---

types affecting the liver, as the Figure suggests. (the correct hemangiosarcoma estimates are given at the bottom of Figure 6-4).

these estimates is incorrect given the statistical uncertainty caused by the small sample sizes in these bioassays.

- *EPA assumes persons are exposed to Perc 24 hours per day for 70 years.* This is also an accurate criticism, as this is certainly a “100<sup>th</sup> percent” upper bound on daily exposure and nearly that for length of exposure. Various expert bodies (see, e.g., Chapter 10 of NAS 1994) have strongly cautioned against relaxing the exposure-duration assumption, however, for one simple point of logic: although U.S. citizens move from one house to another on average about every 10 years, if EPA changed all its risk estimates to reflect 10 years’ duration (and thus redefining concentrations seven times higher as “acceptable”), each person’s lifetime risk would likely increase by that factor of 7, as s/he moved from place to place and encountered the new, higher concentrations everywhere. However, I do see some justification for estimating risk from co-located dry cleaners in Philadelphia using both the 70-year and alternative assumptions – some people will in fact relocate out of the jurisdiction, and although progress has been slow in the U.S. towards converting to the kind of modern low-emission equipment that Europeans have been using for decades, perhaps 70 years is too long a time horizon for estimating the risks without a regulation like the one Philadelphia is contemplating.<sup>13</sup> The hours-per-day assumption could also be relaxed to explore alternative risk estimates, although some co-located residents may well spend 16 hours or more per day exposed to Perc, and some workers may face consecutive exposures from the same or from different facilities as they spend time at their jobs and at their homes, totaling 16 hours or more per day. Nevertheless, in Figure 3 I have constructed an alternative risk range using EPA’s 10-fold range for potency and assuming only 8 hours/day and 10 years of exposure: under this scenario, the risk would range from 3 in 100,000 to 3 in 10,000, which is between 3 and 30 times higher than the 10<sup>-5</sup> minimal-risk benchmark, or 30 to 300 times the 1-in-one-million

---

<sup>13</sup> Of course, many of the residents and workers who would benefit from tighter controls on Perc have *already* been exposed for decades, but still, others will have begun their exposures more recently and may not be exposed for many more decades whether or not this regulation takes effect.

standard for acceptable risk in the federal Clean Air Act, Safe Drinking Water Act, and other pronouncements about risk.

- *EPA estimates risk using susceptible strains of rodents.* This argument is doubly suspect: for one thing, bioassays are purposely designed so that the test subjects are not exposed to any other toxic or carcinogenic stimuli, which of course is the opposite of the human condition, where whatever damage is caused by the substance of interest becomes incremental to the background of environmental exposures we all face. Moreover, the only careful studies comparing the potencies EPA estimated from bioassay data alone to the potencies one can estimate from unequivocally positive epidemiologic data (Allen, Crump, and Shipp 1988) have shown that the rodent strains actually overpredict *and* underpredict human potency for roughly an equal number of chemicals. The real concern about extrapolating from inbred strains of laboratory animals to people is the *underestimation* of risk that must follow for all humans who are more susceptible to cancer than the average person. The most recent NAS report on risk assessment (2009) recommends definitively that EPA correct the asymmetry between cancer and non-cancer methods—EPA incorporates an upward factor of 10 in non-cancer assessment to account for human interindividual variability (see below), but has always incorrectly assumed that we are all equally susceptible to carcinogenesis, which a vast array of literature shows is clearly not the case. NAS (2009) urged EPA to incorporate an upward adjustment of between **10 and 50-fold** to make its cancer risk estimates relevant to persons of above-average susceptibility. If EPA followed that recommendation for Perc, its potency estimate might increase to roughly  $3 \times 10^{-3}$  per ppb, which would make continuous exposure to 40 ppb carry nearly a 1 percent excess risk of cancer.

#### D. Quantitative Summary:

Can anything meaningful, therefore, be said to gauge how “conservative” the EPA potency estimate of [ $1.4 \times 10^{-4}$  to  $1.4 \times 10^{-5}$ ] per ppb Perc is? One approach would be to tally up the magnitudes of the individual upward and downward biases. It is tempting to argue that these quite possibly cancel each other out quantitatively – a factor of 4 at most coming

from the choice of tumor site, and a factor of 2 at most coming from the upper-bound slope, versus an estimate that fails to take account of any humans with above-average susceptibility, leading to an underestimate of 10-fold or more for many exposed persons. Unless (see above) much more evidence is amassed to cast doubt upon all of the animal tumor data, I conclude that there is no strong basis to believe that EPA, CalEPA, and others have interpreted the bioassay data in “conservative” ways at all.

However, it is also possible to cross-check the animal-based potency estimates against epidemiologic data, to see if any inconsistencies arise. Here the Lynge study may be of some use. I estimated above that the subjects in Lynge et al. would have faced an excess cancer risk of roughly  $2.6 \times 10^{-2}$  (0.026) if the *highest* EPA potency estimate was correct. Again, the Lynge study doesn't estimate risk at all (only the odds of having been exposed or exposed), but it is possible to infer whether an excess risk of this size could have been detected. The background lifetime risk of cancer is roughly 1 in 4 (0.25), so the relative risk of (exposed/unexposed) would be  $[(0.25 + 0.026) / 0.25]$ , or 1.1, *if* the highest EPA potency estimate was correct. It is well-known (see, e.g., Crombie 1981) that it is extremely difficult for case-control studies to detect relative risks below about 1.5, so it would seem that the Lynge study couldn't have detected this risk anyway. But to put the power of this study in the most favorable possible light, suppose that all the excess cancer risk accrued at only one tumor site. The lifetime risk of bladder cancer is about 4 percent (0.04), and theoretically a large study *should* have been able to see a relative risk of  $[(0.04 + 0.026) / 0.04]$ , or 1.65. So the Lynge study might provide some evidence that the risk might not have been as high as 0.026 (at the Perc concentrations the subjects faced). By the same token, though, the *lower* estimate of the EPA risk range (0.0026 at these concentrations) could certainly *not* have been detected by the study – the relative risk even if all the excess applied to one tumor site would have been no more than an undetectable  $RR=1.065$ .

#### E. Newest Findings:

Although it is well-known that individual substances may cause tumors in both rodents and humans, but in different organ systems and/or different cell types, there exists suggestive evidence, some of it too recent for EPA to have considered in its 2008 review, that some of the animal tumor types may also be occurring in humans exposed to Perc.

Seidler et al. (2007) reported an odds ratio of 3.4 in a case-control study of German citizens with malignant lymphoma who were exposed to more than about 80 ppm-years (e.g., 10 ppm for 8 years) Perc, although the authors noted that many of those who were exposed to Perc were also exposed to trichloroethylene. But although (see above) humans do not develop the exact analog to rat mononuclear cell leukemia, this and other findings (see the discussion of human lymphoma in connection with the Lynge study above) suggest that Perc may be capable of causing cancer in the blood-forming system in both species.

A new article that has just been accepted for publication (personal communication from Judith Schreiber, New York Attorney General's Office, Jan. 2010) provides suggestive evidence that Perc may be capable of causing kidney cancer in humans, which if borne out would reinforce the findings in male rats (and vice versa). In a presentation made to the NAS Tetrachloroethylene committee in 2009, Schreiber reported on an earlier draft of this study, which was conducted by researchers at the State University of New York at Albany and looked at the possible relationship in the 164 Zip code regions in metro NYC between kidney cancer (as estimated by hospital discharge rates) and the density of dry cleaning establishments. Compared to the Zip codes with the lowest number of cleaners per square km (between 0 and 0.47), all four of the higher-density Zip code groups (ranging up to 2.7 cleaners/km<sup>2</sup> and above) had small but statistically-significant (rate ratios of roughly 1.3-fold) elevations of kidney cancer.

#### ***4. NON-CANCER HAZARDS AND RISKS OF PERC:***

##### **A. Key Studies and Effects:**

As described in great detail in EPA (2008), in addition to the toxic effects of acute exposure to higher levels (circa 100 ppm) of Perc, human exposure to roughly 15 ppm Perc over weeks or months can cause changes in liver enzymes and liver morphology, as well as changes in kidney enzymes suggestive of subtle damage to this organ system (these effects are generally reproducible in laboratory animals). As is the case for many solvents, however, it is believed that damage to the neurologic system is the critical effect of Perc (the effect that occurs at the lowest long-term concentrations). In addition to some evidence that chronic exposure to roughly 20 ppm Perc can cause dizziness and impairment of color

vision, two cross-sectional human studies form the basis for the non-cancer benchmark levels set by the EPA and the ATSDR.

EPA focuses on the study by Altmann et al. (1995). In that study, 14 residents of a West German city who lived above or next to a dry cleaning establishment for an average of 10.6 years were compared with 23 control subjects not exposed to Perc (the exposed and unexposed were said to have the same average levels of alcohol use, tobacco use, and body mass index, but the controls generally were better-educated than the exposed). Figure 7 (Altmann's Figure 1B) shows the Perc levels in the residences of each of the subjects (mean concentration of 4980  $\mu\text{g}/\text{m}^3$ , or 730 ppb). The 37 subjects were given a battery of neurophysiological and neurobehavioral tests; for example, the "continuous performance test" measured how vigilant subjects were to a series of letters displayed on a screen (how quickly they could react to the correct letter when it appeared). The average reaction time of the exposed group was 370 milliseconds (msec), whereas that of the control group was 324 msec (a statistically significant difference), as were several other measures of performance. The authors claimed that these deficits "correspond to the pattern of prenarctic CNS [central nervous system] depression characterizing many solvents." As discussed in detail below, EPA designated the 730 ppb level as the "Lowest Observed Adverse Effect Level" (LOAEL), and after making several adjustments (see the end of Section B below), concluded that a level of **2.4 ppb** was "likely to be without appreciable risk."

ATSDR instead focused on a 1992 study by Ferroni et al., in which 60 female dry cleaning workers (with an average of 10 years' exposure to a median concentration of 15 ppm Perc) were compared to 30 female controls with similar age and verbal IQ. The average reaction time in the former group was 259 msec, as compared to 235 msec in the control group, a statistically (highly) significant decrement. As discussed below, ATSDR took 15 ppm as a LOAEL, and arrived at a "minimal risk level" recommendation of **40 ppb**.

In Appendix C to its comments on EPA (2008), HSIA criticizes the Altmann study for two basic sets of reasons: (1) the controls were more highly educated than the exposed, and since education could plausibly lead to faster reaction time (directly, or indirectly through other variables such as lower alcohol use), the decrement in the exposed may have been due to bias rather than exposure (HSIA asserts that Altmann's regression analysis, in which the authors controlled for education, was inadequate to isolate the exposure effects); and

(2) occupational studies at higher exposure levels failed to show more dramatic neurologic effects, and some of these studies did not show an effect at all.

In response, I would emphasize that the decrement in reaction time apparently persisted after the “education effect” was controlled for—the dispute over whether the regression methods were adequate remains a subjective one dividing HSIA from EPA and others who rely in part on the Altmann results as presented. The second argument above is in some ways a *non sequitur*—the point of a LOAEL is that, of course, it’s the *lowest* level showing effects, and the existence of “higher LOAELs,” and/or studies that failed to show an effect for reasons that are not explored, doesn’t by itself cast doubt on the appropriateness of 730 ppb as a LOAEL.

The 40 ppb action level in the Philadelphia proposal is, of course, nearly 20 times *higher* than the benchmark level EPA derived from Altmann—I am unaware of detailed objections to the results or interpretations from the Ferroni study, which is the one that led to the 40 ppb ATSDR benchmark.

A deficiency with both Altmann and Ferroni that HSIA has not mentioned concerns me—neither study provides data on the neurological performance of *individual* subjects, along with their individual exposures. Hence, the group LOAEL estimates may be too “conservative” (if only persons with exposures greater than this caused the average reaction time to suffer significantly) or not “conservative” enough (if several people at the lower end of the exposure distribution were the ones influencing the average performance the most). I think the central estimates EPA and ATSDR chose are reasonable, but unavoidably uncertain (ideally, someone would conduct a study where individuals could serve as their own controls—measuring individual neurologic performance before and after several years’ exposure to Perc—but such a result would be definition be years away from being available to inform this rulemaking).

The other major uncertainty in these two studies—to what extent occupational exposure is inherently different from residential exposure and thus requires upward or downward adjustment to derive a benchmark level—was recently addressed by Benignus et al. (2009). Figure 7 shows six studies of neurobehavioral deficits, three in residential subjects and three in workers. After the authors tried to put all six studies on a common dose (concentration times duration) axis and a common response (severity of the deficit)

axis, they found that the slopes of the three residential studies were all far steeper than the slopes of the occupational studies (the inset at bottom right “magnifies” the far-left region of the main chart at upper left). Benignus et al. hypothesized this difference could be attributed to either characteristics of the study populations and/or to qualitative differences in the exposure pattern. The well-known “healthy worker effect,” in which occupational groups are systematically less susceptible than the general population (fitness to engage in manual labor rules out frail individuals) plagues occupational epidemiology, and partly explains the use of the 10-fold “intraspecies adjustment factor” (see below) even when the NOAEL or LOAEL comes from human (worker) studies. Unlike diseases such as cancer with long latencies, though, individual sensitivity to the more immediate CNS effects of a substance like Perc might further “strengthen” the occupational population as individuals who notice mild impairment (or experience other effects) might leave jobs that involve high exposures. In this way, workers with long-term Perc exposure might be doubly different from the general population. Benignus et al. also theorized that intermittent (8 hr/day, 5 day/wk) exposure to substances like Perc might have *less* of an effect on the CNS than an equivalent (in terms of concentration-times-time) amount of continuous (residential) exposure, if neurological deficits could be staved off by the daily “recovery” afforded by long periods of non-exposure. This new meta-analysis needs to be studied further, but it could lead to an interpretation of one of the major complaints about the Altmann study as a strength of the investigation (to set a safe level for the whole population, you need to study effects among the spectrum of sensitivities) rather than a weakness.

Figure 5 (from EPA 2008) shows that EPA did not choose the most “conservative” study to set its RfC: of the seven neurotoxicity investigations, the RfC derived from Altmann yields the third most “conservative” value.<sup>14</sup> Figure 6 (also from EPA 2008) shows that although the RfC derived from neurotoxicity is lower than what would be derived from three other non-cancer health hazards (kidney toxicity, reproductive toxicity, and liver toxicity), the RfCs from any of these three alternative datasets would be about 10-fold

---

<sup>14</sup> Note that the legend in the original figure is incorrect—the diagonal striping that is supposed to depict the uncertainty factor for human variation should be the “square grid” pattern that is the middle factor of 10 for Altmann and several of the other studies.

higher—that is, still below (about one-half of) the 40 ppb benchmark set by ATSDR and proposed by Philadelphia.<sup>15</sup>

#### B. Perspectives on Non-Cancer Risk Assessment :

Widespread confusion exists about the “safety” of exposures at the RfC or the MRL, which reflects a lack of understanding of the underlying theory of non-cancer risk assessment, of the key raw material for deriving an RfC or MRL (the “no-effect level” estimated from animal toxicology), and of the purpose and implications of the various “uncertainty factors” EPA or ATSDR applies. I will briefly discuss each of these controversial and misunderstood issues – with the object of explaining that the RfC/MRL should not be thought of as necessarily a “safe” level of exposure.

The very notion that there are “safe” levels of exposure to non-carcinogenic stimuli is gradually giving way to the view that while individual humans may have biological thresholds such that low exposures present no risk, two fundamental facts—the biological variability that distinguishes each person uniquely, and the wide differences in “background” (pre-existing) exposures to similar stimuli—tend to erase any such thresholds for the entire population. In other words, as exposure decreases, the probability that an individual will succumb also decreases, but some individuals (because of their susceptibility, their concurrent exposures to other toxins, or both) will face non-zero risks. For these and other reasons, the most recent National Academy of Sciences report on risk assessment (NAS 2009) strongly recommended that EPA and other agencies “unify” dose-response assessment for non-carcinogens and carcinogens, so that rather than striving to find “safe” levels for the former class of agents, they instead would estimate continuous functions relating exposure to risk for both categories. *The real question, therefore, about the two benchmarks of 40 ppb and 2.4 ppb for Perc is not they whether they are “safe” levels, but whether the population risk of adverse consequences at these levels is acceptably low.* Lacking a formal dose-response model for the neurologic and other non-cancer effects of Perc, one can still attempt to answer this question using EPA’s traditional approach to non-cancer risk assessment.

---

<sup>15</sup> Similarly, note that the legend in Figure 6 is incorrect—in addition to the correction in the previous footnote, the horizontal striping for “animal-to-human extrapolation” should instead denote the dark grey filling that is the top-most region in the two studies on the right of the figure (rat and mouse).

The most common misconception about the two-step process EPA and ATSDR use (first estimating a no-effect level in animals, and then dividing that level by several factors to derive an RfC or an MRL) is that it starts with a “safe” level and then adds margins of safety to it. *Neither of these inferences is correct.* EPA and ATSDR acknowledge this in a veiled way by defining their benchmark levels as “likely to be without appreciable risk” rather than ever calling them safe,” but this admission is not detailed or forthright enough.

First, the under-appreciated part of the NOAEL is the “O” part—these are levels not *observed* to cause elevated frequency of adverse effects in test animals, which is vastly different from a level that in fact did or would not cause harm, especially in the very small sample sizes in the typical non-cancer bioassay (10 to 50 animals per group). Suppose you are designing an underpass on a highway and wanted to make sure the clearance was sufficient for cars and trucks to past through it safely. If you measured 10 vehicles and they were all less than 10 feet high, it would be foolhardy to build in only a 10-foot clearance—the 11<sup>th</sup> vehicle might well be 11 feet high.<sup>16</sup> Roughly speaking, when you observe zero adverse effects in a group of N test subjects, you can only be 90 percent certain that the true risk to each of those subjects is less than  $(1 - e^{-3/N})$ —for example, with a sample size of 10, this upper limit equals 0.26, which is the same as saying that if each of the 10 individuals really had a 0.26 probability (a huge risk, more than one chance in four) of suffering the ill effect, there’s a small but significant chance you could “miss” seeing even one such effect in this small group. Therefore, even with a rather large group of test subjects (50 individuals), observing zero health effects really only tells us with confidence that the risk at that level is less than about  $(1 - e^{-3/50})$ , or about 5 percent. *An exposure that could harm 5 percent of the population is hardly “safe.”*

The situation may be even worse when the study does not yield a NOAEL, but rather a LOAEL, an exposure that causes an observed increase in harm. Here EPA and the other agencies apply a factor of 10 to “adjust” the LOAEL to where the NOAEL *might* have been found had the study included a (LOAEL/10) dose group – and that adjustment may “overshoot” the NOAEL or be insufficient to pin down the true NOAEL (which, again, is only an exposure that is likely to confer a risk of less than a few percent). A recent study

---

<sup>16</sup> This error would be much worse, of course, if you happened to do your research in a residential subdivision where no trucks were allowed – the face-value approach might lead you to build a 6-foot underpass.

(Alexeeff and Broadwin, 2004) suggests that where both a LOAEL and a NOAEL exist in a study for purposes of comparison, most of the time the LOAEL is seven times higher than the NOAEL or less, implying that an adjustment factor of 10 is usually sufficient, but again, the limitations of the NOAEL itself must be emphasized.

Most significantly, the factors EPA and ATSDR apply to the NOAEL are not “safety factors” at all. They are designed to help ensure that humans exposed at the RfC or MRL will face an exposure at or below their NOAEL, not necessarily with a safety margin below their NOAEL. The factor of 10 to adjust exposures (if only a subchronic study is available) to long-term exposure is self-explanatory, and clearly does not confer a margin of safety, but merely accounts for the fact that adverse effects can occur with chronic exposures that might well be missed in short-duration studies. The other two factors of 10, however, are often misconstrued as providing a margin of safety. The interspecies factor of 10 (which was *not* used in either setting the RfC or MRL for Perc, because human subjects were involved in both of the underlying studies) takes a NOAEL from test animals and attempts to convert it to a NOAEL for humans if it turns out that humans as a group are 10 times more sensitive to the given substance than the animals are. If humans are equally sensitive or less sensitive than the animals, this adjustment will in effect confer a margin of safety, but given that by definition the test animals sustain no “background” exposures, such a circumstance would be welcome but not expected. The *intraspecies* factor of 10, which both EPA and ATSDR did apply here, is needed so that persons who are 10 times more susceptible than the average person will not exceed *their* NOAEL—this is an adjustment for them, not a margin of safety for them. Various studies (see, e.g., the references in NAS 2009, Chapter 5) suggest that human-to-human variability is large enough that a factor of 10 will often fail to adjust the population NOAEL to a “no-observed-effect” level for a substantial minority of the human population, but the size of this error is less important than the fact that *any* intraspecies adjustment will leave some fraction of the population right at their own NOAEL for that health effect.

In summary, the traditional non-cancer risk assessment procedures take an exposure level that is associated with a reasonably low risk, and *at best* provide some fraction of the population with an exposure that leaves them at exactly this same low-risk level. Assuming that regulatory or voluntary measures can eliminate exposures above the

resulting RfC or MRL, this is a reasonable way to reduce risks, but it is not a way to assure “safety.”

Specifically with regard to the EPA and ATSDR levels:

- EPA took the mean exposure of subjects showing neurological deficits in Altmann et al. (1995), which was 730 ppb (= 4.98 mg/m<sup>3</sup>), and divided it by 10 to approximate a NOAEL from this LOAEL, by 3 to account for “database deficiencies,” and by 10 again to account for persons of above-average susceptibility. Therefore, **IF** 73 ppb was truly a no-effect level (and this seems doubtful given that Figure 1B in Altmann shows that 4 of the 13 subjects in the exposed group had exposures *at or below 73 ppb*—see Figure 8 in this report), and if the factor of 3 was sufficient to account for the database deficiencies, then 2.4 ppb would provide persons who are 10 times more sensitive than the average member of Altmann’s study group with an exposure exactly at *their* NOAEL—it would “overprotect” people less sensitive than this, but would pose *more risk* than this reasonably low (but not zero) risk at the population NOAEL for anyone who was more than 10 times as sensitive.
- ATSDR took the median exposure of subjects in Ferroni et al. (1992) who showed neurological deficits—15 ppm—and converted it from 8-hour exposure 5 days per week to a continuous exposure scenario by multiplying it by (8/24) and by (5/7), yielding 3.6 ppm, or 3600 ppb. ATSDR then divided 3600 by 10 to approximate a NOAEL from this LOAEL, and by 10 again to account for persons of above-average susceptibility. Therefore, **IF** 360 ppb was truly a no-effect level (and this seems doubtful given that Figure 1 in Ferroni shows that 3 of the 60 subjects in the exposed group had exposures *at or below 360 ppb*), then 40 ppb (note: 36 rounded up by ATSDR to 40) would provide persons who are 10 times more sensitive than the average member of Ferroni’s study group with an exposure exactly at *their* NOAEL—it would “overprotect people less sensitive than this, but would pose *more risk* than this reasonably low (but not zero) risk at the population NOAEL for anyone who was more than 10 times as sensitive.

### C. Newest Findings:

Five very recent studies, only one of which (Perrin et al. 2007) was considered by EPA (2008), add suggestive evidence to the possibility that Perc can cause a spectrum of neurological effects in humans that go far beyond the effects on vision and reaction time previously established, or (the last two studies summarized below) suggest a new set of possible adverse effects, on human immune-system function:

- Perrin et al. (2007) looked at nearly 100,000 birth records in Israel between 1964 and 1976, to explore the parental occupations of the roughly 637 children and young adults diagnosed with schizophrenia among this group. Of the 144 subjects with one or both parents working as dry cleaners when they were born, there were four cases of schizophrenia, which was 3.4 times the incidence rate in the rest of the cohort. According to the authors, this elevation was not appreciably altered after controlling for socioeconomic status, paternal age, and various other factors that could confound the association.
- Li et al. (2008) looked at over 40,000 persons hospitalized in Sweden between 1987-2004 for epilepsy, and examined 43 job titles held by these subjects. Of the 43 jobs, three were associated with a significantly lower incidence of epilepsy (nurses, farmers, and public-safety workers), and five with a significantly higher incidence, including “laundry and dry cleaners,” which was associated with the second-highest relative risk (1.27-fold).
- Fang et al. (2009) looked at 109 cases of amyotrophic lateral sclerosis (“Gehrig’s Disease”) in New England. Among 21 chemicals and groups of agents that the cases and controls self-reported occupational exposures to, “dry cleaning agents” (which could of course refer to solvents other than Perc) had the highest odds ratio (1.9), although this estimate was not significant at the  $p < 0.05$  level.
- Pralong et al. (2009) published a case report of an individual who died from diffuse systemic sclerosis (an autoimmune disease that has been associated with chemotherapy treatment and with occupational exposures to solvents), who had reported prolonged occupational exposure to Perc and to trichloroethylene.
- Emara et al. (2010) found elevated immunoglobulin E levels, increased white blood cell count, increases in natural killer cell counts, and other markers of a heightened immune response, among 40 Perc-exposed dry cleaning workers in

Egypt. The authors hypothesize that “tetrachloroethylene exposure... may lead to the augmentation of allergic diseases or appearance of autoimmune reaction.”

##### **5. THE EFFECT OF INTERVENTIONS ON EXPOSURES:**

The box-whisker plot in Figure 9 shows the medians (bottom of the “boxes”), means (tops of the “boxes”), and maxima and minima (ends of the “whiskers”) from two “before and after” snapshots of regulatory or enforcement interventions in dry cleaning, with the concentration-ranges in the Altmann and Ferroni groups presented for context. The “NYC” plots come from McDermott et al. (2005)—the “before” values are given in summary form in their Table 4, and I took the 65 individual values in their Table 2 and derived summary statistics from them. The “Day Care” plots come from data provided by Philadelphia AMS with respect to an intervention at the Pelham Plaza establishment. The two right-most plots come from a database of current concentrations AMS has measured, and from co-commercial samples collected in New York.

These data suggest the following four general conclusions: (1) installation of engineering controls (or, if these fail, switching to a non-Perc solvent) can successfully reduce Perc concentrations in co-residential settings on average by about a factor of 10, from well above the 40 ppb benchmark to reliably (slightly) below it; (2) establishments that continue to use Perc will find it difficult to ensure co-residential concentrations at the lower EPA benchmark (2.4 ppb); (3) at present, AMS has documented average Perc concentrations in co-located facilities well above 40 ppb, and found sporadic concentrations as high as almost 1000 ppb; and (4) concentrations at adjacent workplaces can be similar, if not somewhat higher, than at adjacent residences. *This last conclusion is a particular concern*—it stands to reason that Perc levels in the cleaners themselves will tend to be higher during normal business hours when machines are running more continuously, and these are the very times of day when co-commercial establishments are full of people (workers) and when adjacent residences *may* be empty or occupied by fewer people. *I therefore commend Philadelphia for considering controls on co-commercial settings, thereby correcting a glaring deficiency in the federal (NESHAP) rules.*

## 6. *RISK-RISK TRADEOFFS:*

Gradually, cutting-edge risk assessments produced for regulatory purposes have begun to attempt not only to estimate the benefits of measures that reduce the “target risk(s)” to health and the environment, but to estimate the *countervailing* risk(s) that could be increased by the very policies being contemplated (see, e.g., Sunstein 1996). Probably the most well-studied (and controversial) example of a possible “risk-risk tradeoff” is the claim that measures to increase automobile fuel economy will lead to increased traffic fatalities (if manufacturers achieve the mpg gains by making cars lighter in weight, and if these lighter cars are less crash-worthy) that may partially, totally, or more than totally offset the lives saved from cleaner air. Although the focus of Agency critics has been the failure to consider offsetting risks, there are, of course, major un-analyzed categories of offsetting *benefits* as well, including both additional risks that are co-controlled by policies not directly aimed at them (see, e.g., Rascoff and Revesz 2002) and economic benefits, as in the common situation where regulatory burdens on one industrial sector are partially or totally offset by new jobs, markets, and consumer surplus in pollution-control sectors and/or the production of substitute goods.

Analysis of offsetting risks and benefits is surrounded by at least as much estimation uncertainty as traditional risk and cost analysis, but it is also confounded by at least three crucial definitional problems (Finkel 2007): (1) one often has to take on faith that a risk tradeoff will even occur at all, because the behavioral response to regulation is usually completely in the control of the very industries who have a compelling interest in convincing government not to regulate them in the first place; (2) it is easy to concoct analyses of “new” risks that involve incorrect double-counting or that are “cherry-picked” to only consider harms and not benefits; and (3) it is entirely a matter of policy and judgment whether to consider a real offsetting risk as a reason to back away from a proposed regulation or as a “wake-up call” to regulate *both* the primary and the secondary risk(s).

Some of the risk-risk tradeoffs mentioned in response to the Philadelphia proposal are suspect along the first two lines above. For example, the claim has been made (Commissioners 2009) that health gains from reduced Perc exposures will be partially

offset by increased exposures to carbon monoxide (from more vehicle miles traveled) as dry cleaners move to remote suburban locations to avoid the regulation. In addition to the inappropriateness of this comparison (CO clearly has no carcinogenic (no-threshold) risk, and the highest concentration ever recorded in Philadelphia is well below the EPA air quality standard of 9 ppm, whereas some indoor Perc levels are currently far above the EPA and ATSDR benchmarks, and Perc reductions by definition are expected to lower cancer risks), the logic that regulation will lead to more car travel is not substantiated. Dry cleaners could respond to the regulation by adopting engineering controls, by changing to less-toxic solvents, by leasing spaces within the city that are not co-located, or by converting to “drop shops” and using Perc at free-standing or other locations outside the regulatory jurisdiction. Only the last of these behavioral responses might increase net vehicle traffic, and it would be in the form of daily truck traffic to and from the “drop shop,” not in the form of individual customers traveling longer distances.

Such a claim is reminiscent of claims HSIA and others made in the OSHA methylene chloride rulemaking (for example, that users of MC would be “forced” to switch to flammable substitutes, leading to increased fires and explosions, or that aircraft paint stripping would be less effective, leading to mechanical failures in flight). With 12 years’ hindsight, it appears the regulation caused few if any users to switch solvents, that those who did so chose substitutes that were less toxic and no more inflammable, and that in any event no serious fire/explosion events were reported in these sectors (and similarly, no reports of paint-stripping-related plane crashes). In general, the more indirect the purported tradeoff, and the more believing it depends on a dubious behavioral response, the less duty an Agency has, in my opinion, to take such a claim seriously.

However, the MC experience *does* warn about one particular kind of direct tradeoff, one with immediate relevance to the Perc proposal. After OSHA’s rule was promulgated, aggressive marketing by overseas producers (abetted by “tunnel vision” at one office within EPA) caused some users of MC to substitute *n*-propyl bromide (*n*-PB), also known as 1-bromopropane, at the time a known neurotoxin and suspected carcinogen. More recently, some dry cleaners in the region (CDC 2008) have also switched from Perc to *n*-PB, with dire consequences. In my opinion, *n*-PB is more dangerous than Perc at comparable concentrations, and new information continues to accrue to reinforce that view:

- The *European Journal of Endocrinology* (1998) reported on 16 (of 25) female Korean workers using 2-BP who developed primary ovarian failure, and 6 of 8 male workers who developed severe sperm-count reductions (
- CDC (2008) published a case report of a 43-year-old man in NJ who had recently begun dry cleaning with “DrySolv” and was hospitalized with headaches, fatigue, visual disturbances, twitching, and joint pain;
- Raymond et al. (2007) reported on 4 furniture workers using 1-BP glue (18 - 254 ppm in air) who developed inability to walk, pain, numbness, and vomiting—these symptoms in some cases persisting for 8 years after leaving workplace;
- Majersik et al (2007) reported that 6 workers exposed to roughly 100 ppm *n*-PB while gluing furniture developed chronic neuropathic pain, persisting for years after leaving workplaces.

Several months ago, the NTP released new data on a cancer bioassay of *n*-PB—the substance caused skin, lung, and intestinal malignancies in male and female mice and rats (NTP’s current opinion is that the evidence of carcinogenicity is “clear” in two of the four sex/species combinations). I fit the female mouse lung tumor data to the linearized multistage model, and I estimate an upper-bound potency value of  $1.95 \times 10^{-3}$  per ppm—using the same method (i.e., no adjustment for pharmacokinetics, in the absence of a quantitative PBPK model for *n*-PB at present), the potency of Perc would be  $1.61 \times 10^{-3}$  per ppm—so the substitute is a more potent carcinogen in addition to apparently being a more potent neurotoxin. *In my opinion, Philadelphia should add a provision to its draft regulation forbidding dry cleaners from switching from Perc to this more dangerous substitute.* Through outreach and training programs, the city should help users identify safer substitutes if they choose not to comply via engineering controls, and to minimize offsetting risks from whatever substitute they adopt (for example, even certain kinds of wet cleaning may have a downside, if cleaners are not cognizant of potential ergonomic hazards to workers handling heavy loads).

## **7. REGULATORY DESIGN:**

As I discussed in my public presentation on October 21, 2009, I recommend that Philadelphia consider a “hybrid” regulatory design, in which establishments would be given the choice of trying to achieve the 40 ppb benchmark via the specific measures spelled out in the regulation, or via any alternative technologies or practices each establishment

believes will be equally or more effective at (presumably) lower cost. If, however, an establishment chooses to innovate, the monitoring requirements must be more frequent.

I also urge the city to consider establishing equally-stringent requirements for co-commercial and co-residential settings, on the grounds that workers may be exposed for more hours per day than residents and that each co-located business will likely contain more inadvertently-exposed citizens than a residence would (on the other hand, the evidence that workers as a group may be less susceptible than the general population might justify the dichotomy in the proposed rule). *In any event, Perc monitoring should take place in the establishment itself as well as in the co-located facilities*, for two reinforcing reasons: levels in the cleaners will surely be higher than in the adjacent establishments, and it is possible that controls designed to lower concentrations in adjacent establishments will have the unfortunate effect of *raising* levels within the cleaners (EPA 1999), and this side-effect must be watched for. Finally, I recommend that Philadelphia consider the definition of “co-located” to encompass situations (if any) where apartments or workplaces share the same ventilation system with a dry cleaners, even if they do not share a wall/floor/ceiling in common.

## **8. CONCLUSIONS:**

- Perc is a dangerous substance, known to cause neurologic damage in humans, and likely to be a human carcinogen.
- Although explanations are being offered as to why regulatory agencies should ignore certain experimental results and certain epidemiologic findings, the number of independent “false positives” that would have to be asserted, each arising from a different misinterpretation of the data, is much greater for Perc than for most other regulated substances—skeptics have an overwhelming variety of concerns to explain, and so far they have not sufficiently explained away any single one of the many findings.
- On balance (considering factors that tend towards overestimation and ones that tend towards underestimation), the upper end of the EPA cancer risk range for

Perc (between  $1.4 \times 10^{-5}$  and  $1.4 \times 10^{-4}$  per ppb) may be somewhat “conservative,” but there is no compelling evidence to suggest that the lower end of this range is conservative at all.

- The EPA RfC of 2.4 ppb may also be somewhat “conservative,” but evidence also exists to suggest it might not be “conservative” enough.
- Therefore, the 40 ppb benchmark proposed by Philadelphia is **not** a “safe” level of continuous exposure. It is likely to exceed the true no-effect level for a significant minority of citizens, and it is likely to pose an excess cancer risk approaching 1 in 1000 for continuous exposure, or approaching 1 in 10,000 for 8 hr/day, 10-year exposure. It is, however, a reasonable science-policy “bright line” to set a level, achievable in other jurisdictions, that will greatly reduce current risks to Philadelphians.
- Although most of the substitutes for Perc are safer, some—especially *n*-propyl bromide, are clearly even riskier and should be controlled as well.

TABLE 1  
(reproduced from Finkel and Ryan, 2007)

Risk in the Workplace

223

EXHIBIT 9.1. OSHA'S SCIENTIFIC CRITERIA  
FOR ACCEPTING A PBPK MODEL.

1. The predominant and all relevant minor metabolic pathways must be well described in several species, including humans. (Two metabolic pathways are responsible for the metabolism of MC in humans, mice, rats, and hamsters).
2. The metabolism must be adequately modeled. (Only two pathways are responsible for the metabolism of MC as compared to several potential routes of metabolism for other compounds, such as benzene and the dioxins. This simplified the resulting PBPK models).
3. There must be strong empirical support for the putative mechanism of carcinogenesis (e.g., genotoxicity), and the proposed mechanism must be plausible.
4. The kinetics for the putative carcinogenic metabolic pathway must have been measured in test animals *in vivo* and *in vitro* and in corresponding human tissues (lung and liver) at least *in vitro*, although *in vivo* human data would be the most definitive.
5. The putative carcinogenic metabolic pathway must contain metabolites that are plausible proximate carcinogens (e.g., reactive compounds such as formaldehyde or S-chloromethylglutathione).
6. The contribution to carcinogenesis via other pathways must be adequately modeled or ruled out as a factor. For example, there must be a reasonable analysis of why reactive metabolites formed in a second pathway would not contribute to carcinogenesis (e.g., formyl chloride produced via the MFO pathway is likely to be too short-lived to be important in MC carcinogenesis).
7. The dose surrogate in target tissues (lung and liver in the case of MC) used in PBPK modeling must correlate with tumor responses experienced by test animals (mice, rats, and hamsters).
8. All biochemical parameters specific to the compound, such as blood:air partition coefficients, must have been experimentally and reproducibly measured. This must be true especially for those parameters to which the PBPK model is most sensitive.
9. The model must adequately describe experimentally measured physiological and biochemical phenomena.
10. The PBPK models must have been validated with data (including human data) that were not used to construct the models.
11. There must be sufficient data, especially data from a broadly representative sample of humans, to assess uncertainty and variability in the PBPK modeling.

Source: Occupational Safety and Health Administration (1997, pp. 1533–1534).

TABLE 2

Upper-Bound ("q<sub>1</sub><sup>\*</sup>") Potency Estimates, Per mg/kg/day, For Each of the Positive Bioassay Results for Perc (Calculated Using "MSTAGE" Emulator of the EPA Linearized Multistage Dose-Response Software)

TUMOR SITE	LABORATORY/SEX/SPECIES	POTENCY
<b>Liver Adenomas/Carcinomas</b>		
	NTP, male mice	3.79x10 <sup>-2</sup> *
	NTP, female mice	1.53x10 <sup>-2</sup> **
	JISA, male mice	2.91x10 <sup>-2</sup>
	JISA, female mice	1.04x10 <sup>-2</sup>
<b>Hemangiosarcomas</b>		
	JISA, male mice	8.47x10 <sup>-3</sup>
<b>Mononuclear Cell Leukemia</b>		
	NTP, male rats	2.19x10 <sup>-1</sup>
	NTP, female rats	1.16x10 <sup>-1</sup>
	JISA, male rats	<b>1.30x10<sup>-1</sup> ***</b>
	JISA, female rats	5.19x10 <sup>-2</sup>
<b>Kidney Tubular Cell Adenomas/Carcinomas</b>		
	NTP, male rats	2.70x10 <sup>-2</sup>
<b>Brain Gliomas</b>		
	MTP, male rats	1.58x10 <sup>-2</sup>

*Notes:*

\* for liver carcinomas alone, value was 2.25x10<sup>-2</sup>

\*\* for liver carcinomas alone, value was 8.13x10<sup>-3</sup>

\*\*\* EPA's official potency estimate

FIGURE 1

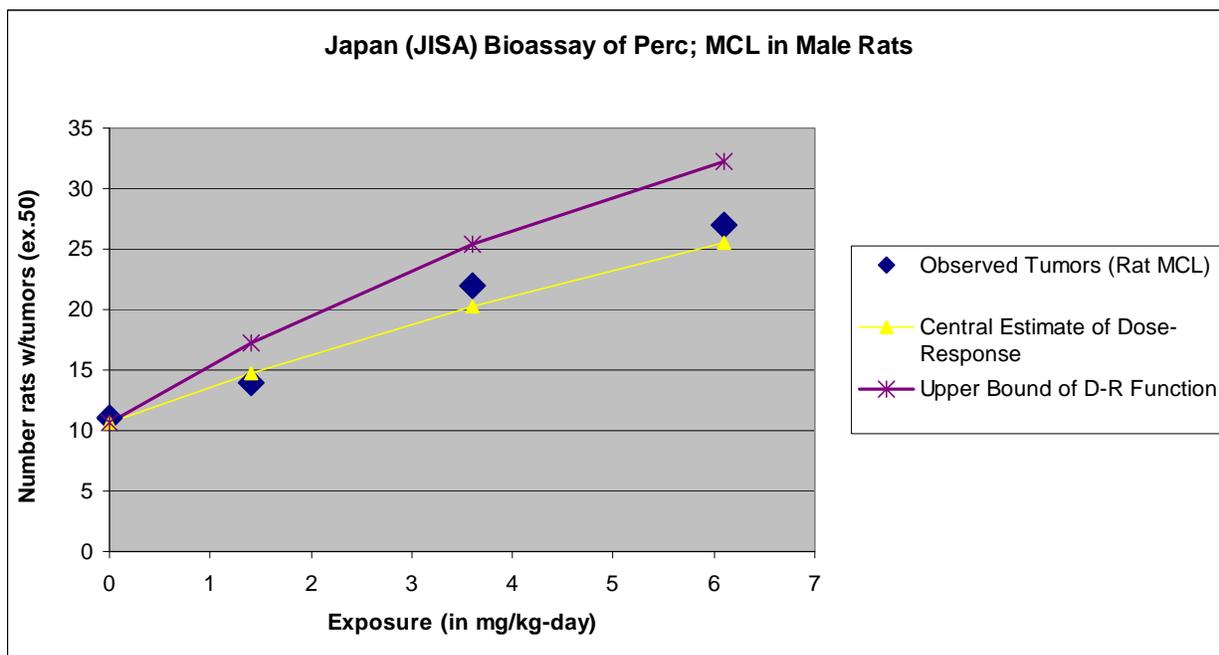


FIGURE 2

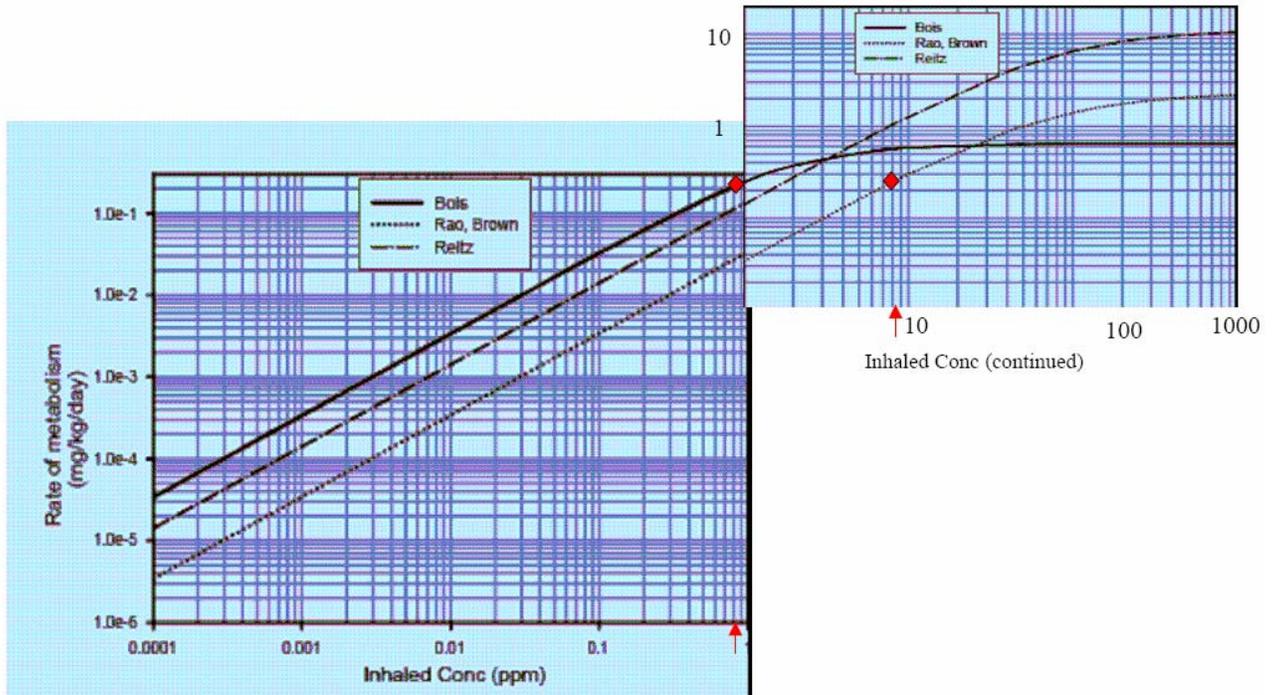


FIGURE 3

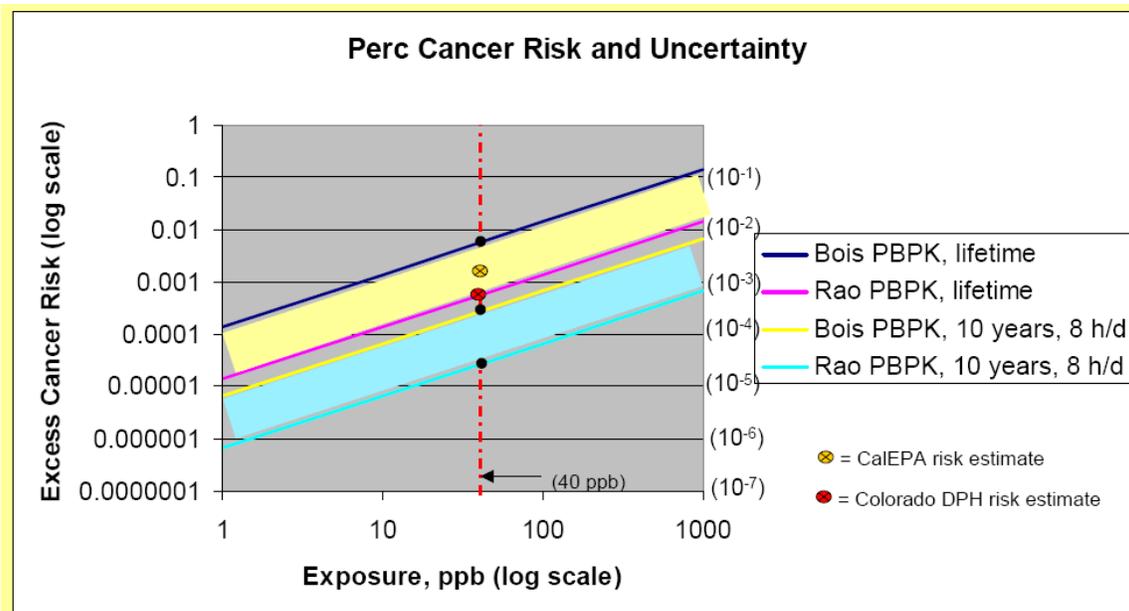
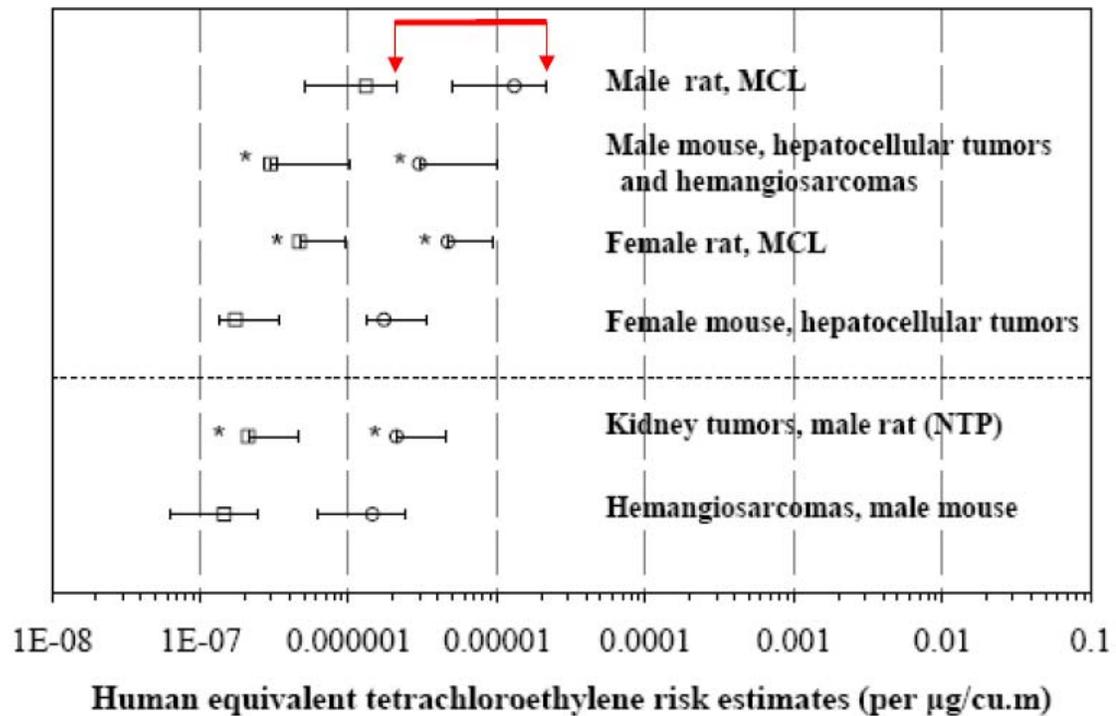


FIGURE 4



**Figure 6-4. Cancer risk estimates for tumor sites associated with tetrachloroethylene exposure in rodent bioassays, using the multistage model.** The four gender/species data sets are provided in the upper section of the graph while two tumor types observed in single bioassays are provided in the lower section. The symbols denote the slopes (to background risk) from the mean estimate of exposure corresponding to 10% extra risk, using the Rao and Brown (1993) PBPK model ( $\square$ ) and the Bois et al. (1996) model ( $\circ$ ) to extrapolate to human equivalent exposures. The bars indicate the slopes from the lower and upper bounds on the mean estimates. \* indicates lower bounds that could not be estimated.

FIGURE 5

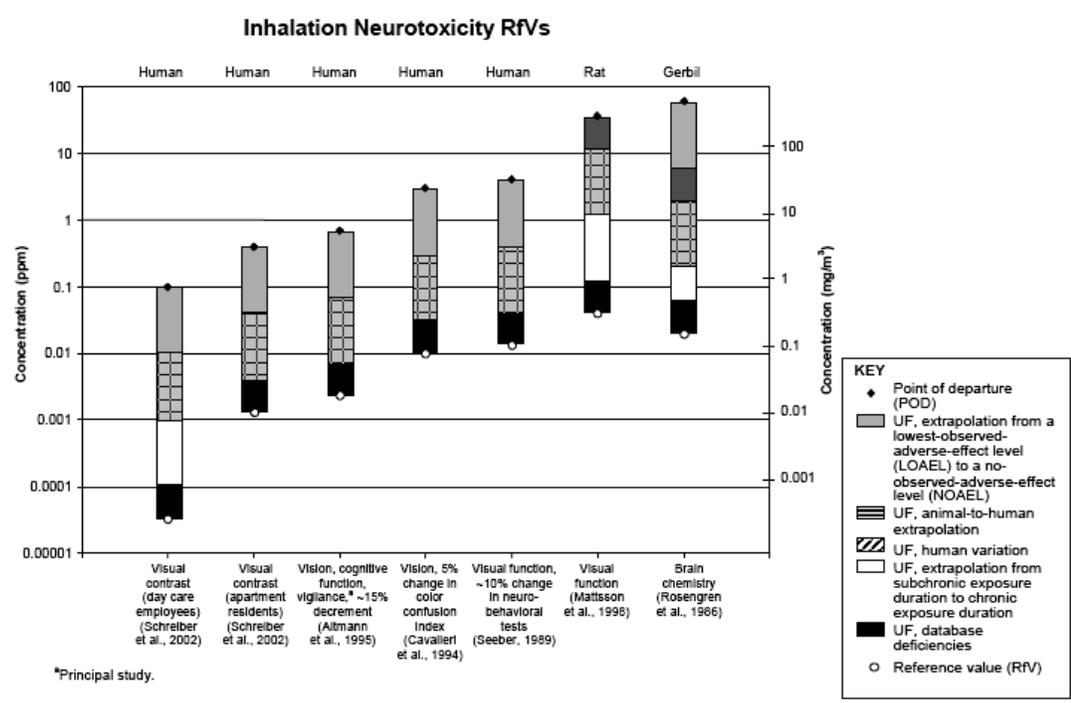


Figure 6-1. Array of PODs and reference values for a subset of neurotoxic effects in inhalation studies.

FIGURE 6

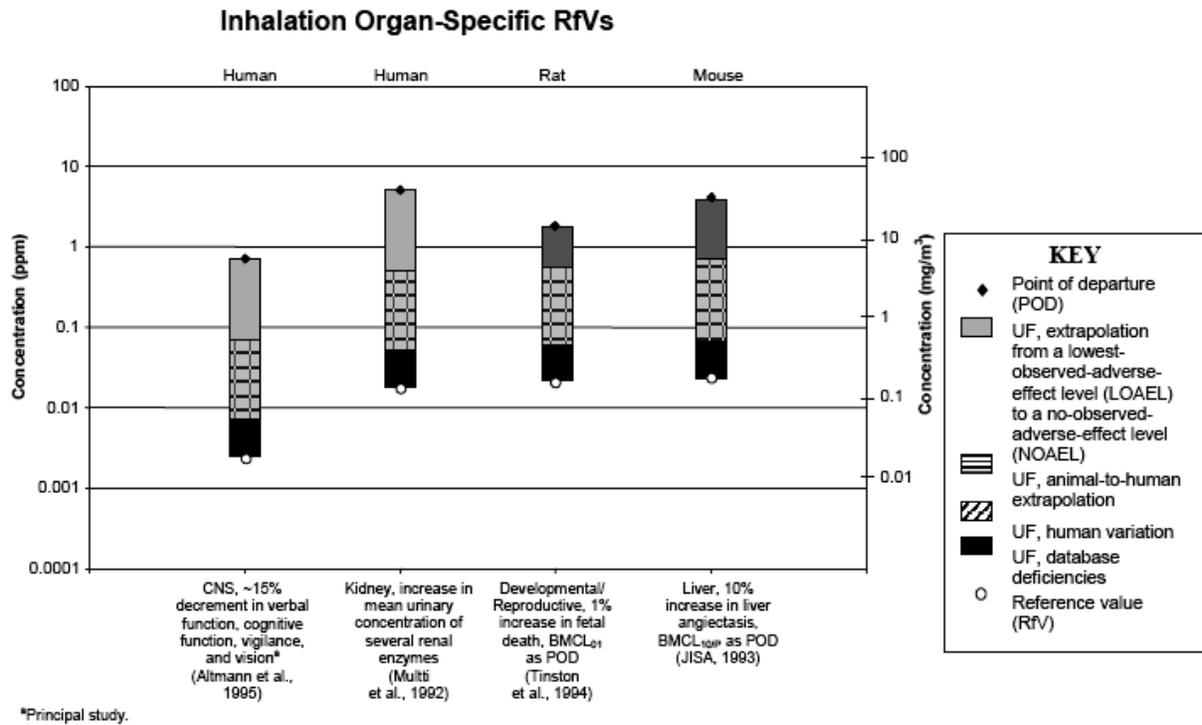
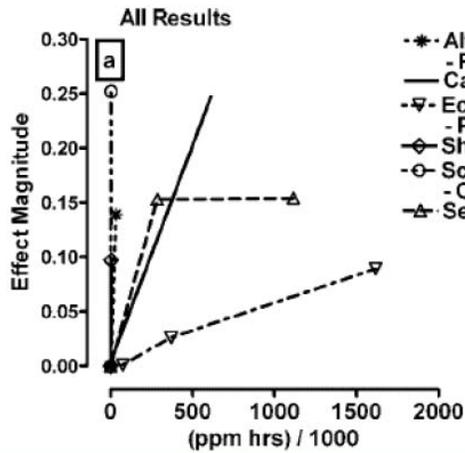


Figure 6-2. Organ-specific RfVs for inhalation exposure to tetrachloroethylene.

FIGURE 7



*Journal of Toxicology and Environmental Health, Part A, 75:425-431, 2003*  
 This article is an U.S. Government work and, as such, is in the public domain in the United States of America.

**Long-Term Perchloroethylene Exposure: A Meta-Analysis of Neurobehavioral Deficits in Occupationally and Residentially Exposed Groups**

Vernon A. Benignus<sup>1,4</sup>, William K. Boyes<sup>2</sup>, Andrew M. Geller<sup>3</sup>, and Philip J. Bushnell<sup>1</sup>

<sup>1</sup>Human Studies Division, <sup>2</sup>Neurotoxicology Division, and <sup>3</sup>Research Planning and Coordination Staff, National Health and Environmental Effects Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, and <sup>4</sup>Department of Psychology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

FIG. 5. Comparison of all of the long-term exposure data: (a) all data plotted on same scale; (b) only the high-exposure, low-sensitivity data; and (c) only the low-exposure, high-sensitivity data. Note that the low-sensitivity data are all from workplace exposure and all of the high-sensitivity data are from residential exposures.

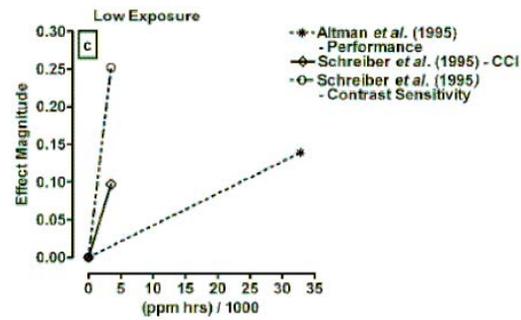


FIGURE 8

Individual Exposures in Altmann et al. (1995)

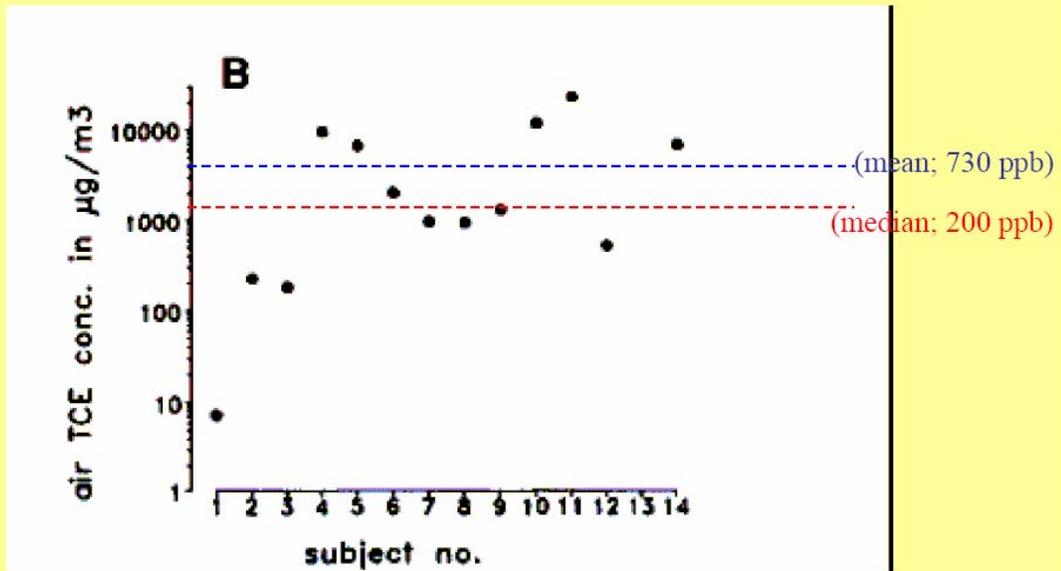
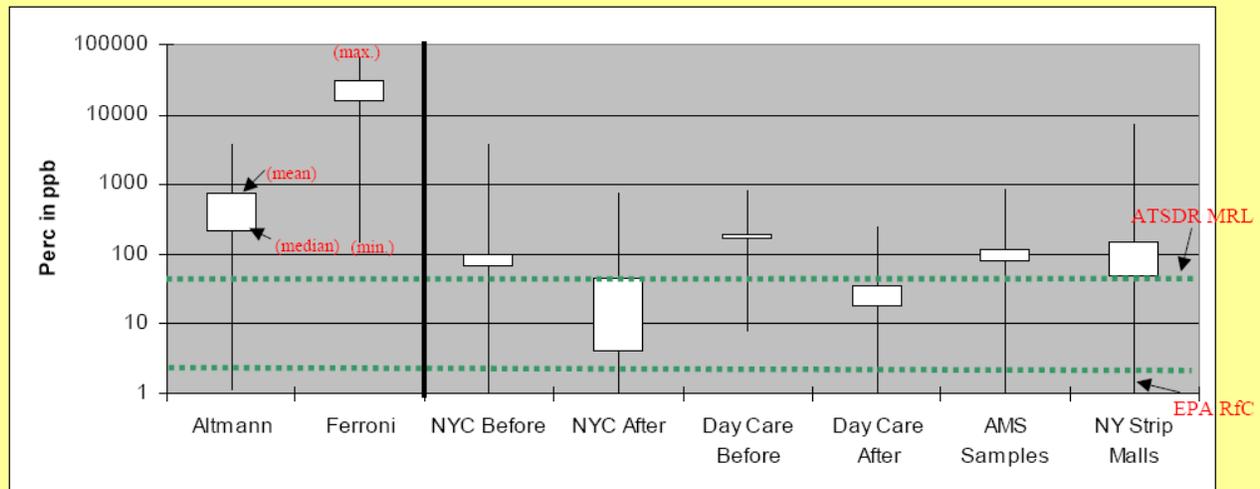


FIGURE 9

Ranges and Central Estimates of Key Perchloroethylene Concentrations Measured:



Note that: (1) regulations and site-specific intervention can lower concentrations by roughly a factor of 10; and (2) co-commercial concentrations are similar, if not slightly higher, than co-residential concentrations

## REFERENCES CONSULTED

- Aggazzotti, Gabriella, et al. (1994). "Indoor Exposure to Perchloroethylene (PCE) in Individuals Living with Dry-Cleaning Workers," *The Science of the Total Environment*, **156**: 133-7.
- Allen B.C., K.S. Crump, and A.M. Shipp (1988). "Correlation between Carcinogenic Potency of Chemicals in Animals and Humans," *Risk Analysis*, **8**:531-44.
- Altmann, Lilo et al., (1995). "Neurobehavioral and Neurophysiological Outcomes of Chronic Low-Level Tetrachloroethylene Exposure Measured in Neighborhoods of Dry Cleaning Shops" *Environmental Research*, **69**: 83-89.
- Alexeeff, G.V., R. Broadwin, and W-C. D. Wang (2004) "Establishing Data-Derived Uncertainty Factors for Risk Assessment: Evaluating the LOAEL-To-NOAEL Data for Acute Inhalation Exposures. Office of Environmental Health Hazard Assessment, California EPA (Presentation and 2004 Annual Meeting of the Society for Risk Analysis)
- American Cancer Society (2009). *Lifetime Probability of Developing or Dying from Cancer*. July 13, 2009; accessed at [http://www.cancer.org/docroot/CRI/content/CRI\\_2\\_6x\\_Lifetime\\_Probability\\_of\\_Developing\\_or\\_Dying\\_From\\_Cancer.asp](http://www.cancer.org/docroot/CRI/content/CRI_2_6x_Lifetime_Probability_of_Developing_or_Dying_From_Cancer.asp).
- Benignus, Vernon A. et al, (2009). "Long-Term Perchloroethylene Exposure: A Meta-Analysis of Neurobehavioral Deficits in Occupationally and Residentially Exposed Groups." *Journal of Toxicology and Environmental Health, Part A*, **72**: 824-831.
- CDC, *Morbidity and Mortality Weekly Report (MMWR)*, Dec. 5, 2008, (Vol. 57, No. 48), pp. 1300-2. "Neurologic Illness Associated with Occupational Exposure to the Solvent 1-Bromopropane —New Jersey and Pennsylvania, 2007-2008" 2008
- Colorado Dept of Public Health and Environment, Aug. 17, 2006 (3 pp) "Policy on an Interim Risk Evaluation and Management Approach for PCE." 2006
- Commissioners of the Mayor's Commission on Asian American Affairs (2009). Letter to Air Pollution Control Board (25 signers), 3 pp., March 3, 2009.
- Crombie, Iain K. (1981). "The Limitations of Case-Control Studies in the Detection of Environmental Carcinogens," *Journal of Epidemiology and Community Health*, **35**: 281-7.
- Dugard, Paul H. (2008). Comments of HSIA on EPA's Draft Toxicological Review of Tetrachloroethylene, June 2008, 153 pp.
- Emara, A.M., et al. (2010). "Immunotoxicity and Hematotoxicity Induced by Tetraethylene in Egyptian Dry Cleaning Workers," *Inhalation Toxicology*, e-pub ahead of print, Jan. 4.
- Environmental Protection Agency (1999). "Common-Sense Approaches to Protecting Workers and the Environment: Interagency Cooperation Towards Cooperative Solutions." Conference held in Washington, D.C., June 17-18.

EPA Federal Register Notice, July 11, 2008 (73 FR 39871-75) National Perchloroethylene Air Emission Standards for Dry Cleaning Facilities 2008

EPA Federal Register Notice, July 27, 2006 (71 FR 42724-45) National Perchloroethylene Air Emission Standards for Dry Cleaning Facilities

unreferenced document (but written by funding agency, EPA, to NAS), (5 pp.) Revised Charge to the National Academy of Sciences Panel for the Toxicological Review of Tetrachloroethylene, February 9, 2009.

Fang, Fang, et al. (2009). "Workplace Exposures and the Risk of Amyotrophic Lateral Sclerosis," *Environmental Health Perspectives*, **117(9)**: 1387-92.

Ferroni, C., et al. (1992). "Neurobehavioral and Neuroendocrine Effects of Occupational Exposure to Perchloroethylene." *Neurotoxicology*, **13**: 243-8.

Finkel, A. M. (2007). "Distinguishing Legitimate Risk-Risk Tradeoffs from Straw Men." Presentation at the Annual Meeting of the Society for Risk Analysis, San Antonio, TX, Dec. 11.

Finkel, A.M. and P.B. Ryan (2007). *Risk in the Workplace: Where Analysis Began and Problems Remain Unsolved*. Chapter 9 (pp. 187-237) in **Risk Assessment for Environmental Health**, Association of Schools of Public Health (M.G. Robson and W.A. Toscano, eds.), John Wiley & Sons Inc.

Halogenated Solvents Industry Alliance (HSIA) Update, Sept 2008 (2 pp) "EPA PERC Draft-- A 'Dog's Breakfast'." 2008

Hattis, D., and R. Goble (1991). "Expected Values for Projected Cancer Risks from Putative Genetically-Acting Agents," *Risk Analysis*, **11(3)**: 359-63.

Koh, J.M., et al. (1997). "Primary Ovarian Failure Caused by a Solvent Containing 2-Bromopropane." *European Journal of Endocrinology*, **138(5)**: 554-6.

Lee, Jong-Song (2008). *Tetrachloroethylene (PCE)*. Report by the Texas Commission on Environmental Quality, April 15, 20 pp.

Li, X., J. Sundquist, and K. Sundquist (2008). "Socioeconomic and Occupational Risk Factors for Epilepsy: A Nationwide Epidemiological Study in Sweden," *Seizure*, **17(3)**: 254-60.

Lynge, Elsebeth et al. (2006). "Cancer in Persons Working in Dry Cleaning in Nordic Countries," *Environmental Health Perspectives*, **114(2)**: 213-9.

Majersik, Jennifer J. et al. (2007). "Severe Neurotoxicity Associated with Exposure to the Solvent 1-Bromopropane (n-propyl bromide)," *Clinical Toxicology*, **45**: 270-6.

McDermott, Michael J. et al (2005). "Tetrachloroethylene (PCE, Perc) Levels in Residential Dry Cleaner Buildings in Diverse Communities in New York City," *Environmental Health Perspectives* **113(10)**: 1336-43.

National Academy of Sciences (1994). **Science and Judgment in Risk Assessment.**

National Academy of Sciences (2009). **Science and Decisions: Advancing Risk Assessment.**

National Toxicology Program (2009). *NTP Technical Report on the Toxicology and Carcinogenesis of Studies of 1-Bromopropane in F344/N Rats and B6C3F1 Mice.* NTP TR-564, November 19, 223 pp.

New York State Department of Health, May 2003 (5 pp) "Fact Sheet: Tetrachloroethylene (Perc) In Indoor and Outdoor Air"

Perrin, M.C., et al. (2007). "Tetrachloroethylene Exposure and Risk of Schizophrenia: Offspring of Dry Cleaners in a Population Birth Cohort, Preliminary Findings," *Schizophrenia Research*, **90(1-3)**: 251-4.

Pralong, P., et al. (2009). "Diffuse Systemic Sclerosis after Occupational Exposure to Trichloroethylene and Perchloroethylene," *Annals of Dermatology and Venereology*, **136(10)**: 713-7.

Rascoff, Samuel and Richard Revesz (2002). "The Biases of Risk Tradeoff Analysis: Towards Parity in Environmental and Health-and-Safety Regulation." *Univ. of Chicago Law Review*, **69**: 1763-1836.

Qiu, J., Y.C. Chen, J.V. Bruckner, and J.W. Fisher (2010). "Bayesian Analysis of a Physiologically-Based Pharmacokinetic Model for Perchloroethylene in Humans." *Journal of Toxicology and Environmental Health*, **73(1)**: 74-91.

Raymond, Lawrence W. and Marsha D. Ford (2007). "Severe Illness in Furniture Makers Using a New Glue: 1-Bromopropane Toxicity Confounded by Arsenic," *Journal of Occupational and Environmental Medicine*, **49(9)**: 1009-1019.

Schreiber, Judith S. et al (1993). "An Investigation of Indoor Air Contamination in Residences above Dry Cleaners" *Risk Analysis*, **13(3)**: 335-44.

Seidler, A., et al. (2007). "Solvent Exposure and Malignant Lymphoma" A Population-Based Case-Control Study in Germany," *Journal of Occupational Medicine and Toxicology*, **2(2)**, 11 pp.

Sinsheimer, Peter, et al. (2007). "The Viability of Professional Wet Cleaning as a Pollution Prevention Alternative to Perchloroethylene Dry Cleaning," *Journal of the Air and Waste Management Association*, **57**: 172-8.

Storm, Jan E. and Kimberly A. Mazor (2004). "Update of Residential Tetrachloroethylene Exposure and Decreases in Visual Contrast Sensitivity"), *Environmental Health Perspectives* **112(15)**: A863-5 (NY Dept of Health).

Sunstein, Cass (1996). "Health-Health Tradeoffs," *Univ. of Chicago Law Review*, **63**: 1533-71.

Tanios, M.A., H. El Gamal, B.J. Rosenberg, and P.M. Hassoun (2004). "Can We Still Miss Tetrachloroethylene-Induced Lung Disease? The Emperor Returns in New Clothes." *Respiration*, **71(6)**: 642-5.

Walsh, Bryan (2009). "Guilt-Free Laundry," *TIME*, Nov. 23, 2009, page 101.

Wolf, Katy (2008). Institute for Research and Technical Assistance (IRTA), May 2008, 57 pp. "AB998 Non-Toxic Dry Cleaning Incentive Program: Demonstrations of Safer Technologies for the Textile Cleaning Industry"2008