Additional Preliminary Comments from Dr. Daniel Zelterman

Charge Question 2 - Breast cancer incidence

I read Steenland et al 2004 for some follow-up on the background, data, and methods that are missing in the present report. Table 1 of Steenland 2004 examines 16 causes of mortality by three populations (overall, male, and female) for a total of 48 hypothesis tests. Each of these was expressed as a 95% confidence interval of the standard mortality ratio (SMR). When there is no health effect of EtO, each of these 48 ratios excludes the value "1" with probability 5%. Then we would expect 2 or 3 to exclude this value. As expected, there is only one and this is for bone cancer, for all populations combined.

(Why doesn't this EPA report contain a section on bone cancer?)

Overall, Steenland 2003 and Steenland 2004 estimate the standard incidence ratio (SIR) of breast cancer to be .87 (95% CI: .77 - .97) relative to SEER rates. Even those with the highest quintile of exposure have an estimated SIR with 95% confidence interval (.94 -1.69) that includes "1."

So: Exposure to EtO exhibits no, or possibly a beneficial, effect in preventing breast cancer.

This is probably because those exposed to EtO have well-paying jobs in the healthcare industry.

With these comments made, let us address the EPA report.

There is no description of the data in Section 4.1.2. The description on page 1-2 is superficial. The footnote praising the study at the bottom of the page only adds to the vague sense that the authors of this report either don't understand the data or its shortcomings.

The data analysis is Section 4.1.2 is poorly explained and not well motivated. Is the purpose to try several different models in order to find a significant p-value? The first full paragraph on page 4-24 fails in many ways. What makes one method as the "best way"? The concept of "low-exposure extrapolation" is wrong on so many levels.

Why is a 15 year lag used in some places and 20 year lag in others? Where did these numbers come from? Are these estimates? How were they estimated? By Steenland? Were other values examined? What were those findings?

Models are described as "best fitting," last paragraph on page 4-30, and other places. What does this mean? Was there a formal goodness-of-fit test performed? Which one? Or was this the model that provided the most statistically significant result?

How were these models chosen out of the many approaches available to us? Were other models tried? Why or why not?

Who wrote this?

Are the methods clearly and transparently described? No.

Are these scientifically appropriate? No.

Charge Question 4 - Uncertainty in the cancer risk estimates

Extrapolation is a bad word in any statistical report. This word should never be used and any methods that rely on extrapolation is useless. The fact is, we have no idea what the low-exposure risk should be. Models such as splines may fit the data well (what does that mean?) but models with noticeable "corners" are biologically indefensible. More straightforward is a quadratic risk in exposure, or log-exposure. The study of 2-AAF (a known carcinogen) in the 1970's proved that low-exposure extrapolation cannot be done with any precision.

But we DO know a lot about unexposed individuals. The NCHS NCI, etc. data provides national rates for all kinds of cancers, with further breakdown in age- and sex-specific risks. We can safely assume that these are the rates for unexposed populations. Is this data used in the model fitting? Probably not, is my guess. A Bayesian analysis that includes these population values would greatly benefit the modeling process.

Bayesian analyses are controversial and shunned in many circles. Yet, rodent data suggests elevated cancer risk in mammary and lymphoid cancers but the evidence is sparse in human populations. So why are the authors of this report Bayesian in some of their actions but not others?

So, to meet half-way, why not perform a meta-analysis with the studies listed in Table 4-25 in Section 4.6, page 4-83?

Right now, Section 4.1.4 sounds so indecisive. Some models were suggested and tried but it doesn't sound as though here were good reasons for doing so. If there were strong motivation for the methods used then this report would read as a much stronger document.

The data is based on the findings of Steenland et al 2004 is criticized, last paragraph, beginning on page 4-56, and addresses the bias due to the "healthy-worker effect." This was the only discussion of the data that I read.

The Cox model assumes that relative risk (RR) is independent of age. Is this assumption tested? If it is not a valid assumption then why use Cox regression? The test for this is an interaction term of exposure and age. But the power of the Cox model is so low that it is no surprise that this was found not statistically significant.

Section 4.5: First paragraph may have been written on the basis of careful thought but all of these numbers are meaningless to the typical reader. In fact, every estimated risk statistic in this section is on the basis of parts per billion of EtO exposure. So what is the typical exposure in Steenlan's cohort?

When the authors talk about increased lifetime risk, is this a lot or a little? Is it negligible? When I travel to the EPA from Connecticut, I am moving to a much warmer climate, thereby greatly increasing my risk of shark attack. Should I be concerned? No. The increase in risk is great but the probability is small.

Section 4.7: Again we are looking at estimated exposure levels but not given an idea about what a typical exposure is. How do I interpret Table 4-26, page 4-89? There are lots of scary numbers but how do these compare with other causes of mortality, or shark attack, for example?

What's with the all the AIC's, anyway? This is not useful as goodness-of-fit unless the models are nested (one is a special case of the other).