February 1, 2005

Ms. Silvia M. Garrigo
ChevronTexaco Global Downstream
2333 Ponce de Leon Boulevard, 4th Floor
Coral Gables, Florida 33134

Dear Ms. Garrigo:

This letter is a summary epidemiologic assessment of three reports relating to cancer occurrence and other health problems in the Amazon basin of Ecuador. The three reports that we assess here are the following:

1. Hurtig A-K, San Sebastián M: Geographical differences in cancer incidence in the Amazon basin of Ecuador in relation to residence near oil fields. Int J Epidemiol 2002;31:1021–1027

2. The Center for Economic and Social Rights: Rights Violations in the Ecuadorian Amazon. The Human Consequences of Oil Development. Manuscript dated March 1994.

3. San Sebastián M, Córdoba JA: Yana Curi Report: Impact of oil activity on the health of the Ecuadorian Amazon Basin Population.

Of these three documents, the second and third appear to be unpublished reports, and only the first is a scientific paper that is published in a peer-reviewed scientific journal. Below we give a description and critique of each of the studies, along with some general comments.

1. Hurtig A-K, San Sebastián M: Geographical differences in cancer incidence in the Amazon basin of Ecuador in relation to residence near oil fields. Int J Epidemiol 2002;31:1021–1027

This article is a study of cancer occurrence in four Amazon basin provinces. The authors do not describe the age range of the population studied, but it appears to be a study of adult males and females. They aimed to learn the extent to which proximity of residence to oil fields is related to cancer occurrence. The study design was an ecologic design, which means that the study compared the rates of disease for geographic regions, rather than measuring the proximity of residence to oil fields for individual people. The authors classified the 15 counties in the four provinces of Sucumbios, Orellana, Napo and Pastaza into two categories, those where oil exploitation had been ongoing for a minimum of 20 years (four counties with 118,264 people), and those without oil development activities (11 counties with 115,710 people). Apparently

there were no counties in the region in which there was oil development for less than 20 years. The study period encompassed the years 1985–2000. The incidence rate of all cancers and specific types of cancer of was calculated for the exposed counties, separately for men and women, standardized for age, and reported as the ratio of standardized incidence rates, along with 95% confidence intervals, separately by sex.

The authors report a moderately higher cancer rate in the counties that they designated as exposed than in the counties that they designated as unexposed. They note that the exposed counties have "severe contamination of water sources" and imply that there might be a causal connection, although with a caveat: "The results suggest a relationship between cancer incidence and living in proximity to oil fields, although this ecologic study cannot lead to causal inference." They then proceed to offer arguments supporting a causal interpretation. They offer four arguments:

1. The association between exposed counties and cancer occurrence is strong.
2. Benzene and PAH, which are components of crude oil, have been associated with cancer.
3. The findings are consistent with other investigations.
4. The time sequence is plausible.

In a well balanced scientific assessment of various theories to explain an observation, one would expect to find careful weighing of evidence that supports and detracts from various theories. These authors mention broadly the weaknesses of ecologic studies for causal inference, but there is a notable absence of balanced criticism that one would expect to find in a reasoned scientific assessment. Scientists are usually hesitant to infer causal connections, especially from ecologic data for which a broad array of alternative explanations could be readily invoked to explain the findings. These authors have not even mentioned any alternative theories to explain the differences that they report, making it seem that their role is closer to that of advocate than that of a skeptical scientist.

Their arguments in favor of a causal interpretation are weak to begin with:

1. They declare the association between "exposure" and cancer to be strong, but they do not state which of the many associations presented that they refer to, nor do they define "strong." If we take the data for all cancers, there is an approximately 50% greater rate among both men and women in the exposed counties (RR $\approx$ 1.5), compared with unexposed counties. Few epidemiologists would describe a RR of 1.5 as strong, however, and this magnitude of association is one that is readily compatible with uncontrolled confounding.

2. Merely noting that people in the exposed counties have more benzene and PAH exposure does not contribute to any causal inference. They have not established individual exposure, nor have they analyzed the pattern of specific cancer findings to determine whether the pattern supports the theory that these exposures are causing the increase. In fact, the data suggest otherwise (see below).

3. Consistency with other investigations is claimed, but not shown. Consistency implies more than the superficial idea that more than one study shows an increase. The similarities of exposure and types of disease, control of confounders, and many other factors need to be considered before consistency can be claimed. Furthermore, if several studies are all subject to the same bias, such as a confounder that goes uncontrolled in all of them, the consistency can result from a consistent error.

4. The time sequence may be plausible, but is of very little value. Suppose that the growth in the number of pharmacies in the exposed counties increased over the same time period. That would hardly be evidence in favor of the theory that the increase in pharmacies was responsible for an increase in cancer occurrence.

The most striking problem with this paper is that the authors have completely neglected consideration of competing explanations for their data. Most epidemiologists are preoccupied with the possibility that biases might account for observed differences, and for good reason, especially in ecologic studies, in which biases abound. To suggest just a few issues toward which the authors should have directed their attention: 1) Is case ascertainment equivalent for exposed and unexposed counties? 2) Can exposure to carcinogens at levels known to be carcinogenic be documented? 3) The authors have controlled for age and sex, but not for other confounders. For example, could there be smoking differences between exposed and unexposed counties?

Among the data presented, the most interesting finding is the RR of 4 for cervical cancer. There were 96 cases of cervical cancer, accounting for one-third of all female cancer observed in the exposed counties, and most of the excess cases among females in exposed counties. Cervical cancer is basically a venereal disease, and it is far-fetched to think that environmental pollution could explain this increase. More plausibly, this increase, if real, results from different social conditions that accompanied economic development in a frontier area. Similarly, by far the largest number of cases in males are cases of stomach cancer, which is known to be strongly associated with diet. The different diet in an economic frontier area is a plausible explanation for an increase in stomach cancer. We do not propose these alternative theories as better or worse than the causal theory, which is the only one that the authors considered. Apparently, however, the authors have not presented a thorough review of all possible explanations for their findings.

2. The Center for Economic and Social Rights: Rights Violations in the Ecuadorian Amazon. The Human Consequences of Oil Development. Manuscript dated March 1994.

This document, unlike the published paper by Hurtig and San Sebastián, is not a peer-reviewed scientific publication. Indeed, it is not proposed as such, as is evinced by the document's summary, which states that "The indigenous peoples of Ecuador, after more than 500 years of being silenced, are here to denounce publicly the continued destruction of our lands and cultures, and what you in the modern world call the environment." Regardless of the merits

or demerits of the document as a denunciation of cultural and environmental destruction, it is not a scientific study of environmental effects. A scientific study would be an attempt to apply skepticism and reason to a specific problem, whereas an advocacy document such as this is presumably intended to marshal arguments in favor of a particular point of view. As scientists asked to view the science related to a specific question, we have no opinion to offer on the advocacy argument itself. Indeed, we have always been opponents of deforestation. That caveat notwithstanding, the document does propose that one of its objectives was "to collect data on contamination levels and associated health effects" in Ecuadorian region of the Amazon, and we have been asked to comment solely on the scientific merit of that assessment.

The document describes environmental oil contamination that it attributes to inadequate standards of business practices. The health effects of the contamination, which is only described by indirect reference, are reported in section IV of the document, which had two stated objectives: "(1) to measure exposure levels of PAHs and VOCs faced by residents of the Oriente living in proximity to oil producing areas; and (2) to determine whether residents are experiencing observable health problems associated with oil exposure." The document, however, does not offer any study design, so one must conclude that there was no formal study that this document reports. Rather, there were discussions with people in the area about exposure to PAHs and VOCs, and then health examinations by a physician. In an epidemiologic study, at a minimum, the health of exposed people would be compared with that of unexposed people. In this instance, however, "An occupation and environmental medicine physician interviewed and examined residents in areas reportedly contaminated by oil in order to identify adverse health effects attributable to oil exposures." This procedure bears little resemblance to an epidemiologic study. As mentioned, there is no comparison, without which there can be no basis for inference. In addition, the physician is attempting to infer the cause of each exposed person's health status from his or her examination, a procedure that is generally invalid and has been questioned by scientists and skeptics alike at least as far back as David Hume (see *A Treatise of Human Nature*, 1739). The problems with attempting to learn about causal relations from inferences on individuals was discussed critically in the following publication: Rothman KJ, Ray W: Should epidemiologists exclude cases with a "known" cause of their disease? Pharmacoepidemiol Drug Safety 2002; 11:11-14; Rothman and Ray showed that making an inference on each person is not valid unless the disease is rare in the absence of exposure and occurs immediately after exposure occurs, a situation that would not apply in this case.

The actual findings relate only to dermatoses in a few exposed people. The authors do acknowledge that there are some limitations to what they present, by calling their work a "limited study" and suggesting that there is a need for additional studies. This work, however, is not a formal epidemiologic exercise, and cannot be evaluated in that context. Furthermore, this "study" lacks the central characteristic of any scientific inquiry, which is skepticism regarding the explanation for observations.

3. San Sebastián M, Córdoba JA: Yana Curi Report: Impact of oil activity on the health of the Ecuadorian Amazon Basin Population.

This report describes oil pollution in the Ecuadorian Amazon region, and the investigation of a cancer cluster in the village of San Carlos. We will comment only on the investigation of the cancer cluster.

Cluster investigations are notorious for being subject to biases and providing little information of value to aid in inferring the causes of reported clusters (Rothman KJ: Clustering of disease (Editorial). Am J Public Health 1987;77:13-15; Rothman KJ: A sobering start for the cluster busters' conference. Am J Epidemiol 1990;132(S):6-13.) The first issue to consider in any cluster investigation is whether the reported cluster really represents an excess occurrence. As Schulte et al. showed in the occupational setting, most reported clusters do not represent a real excess (Schulte PA, Ehrenberg RL, Singal M: Investigation of occupational cancer clusters: theory and practice. Am J Public Health 1987;77:52-55).

The authors report a three-step investigation: case-finding; calculation of standardized incidence ratios; and investigation of water pollution. This investigation amounts to a comparison of the cancer incidence in San Carlos with the cancer incidence for Quito, controlling for age and sex by using the population of Quito and its cancer registry as the source of comparison data. The author has standardized the findings to the age distribution of the estimated population of San Carlos, separately for men and women, thus controlling for age and sex. The author found a total of 10 cancer cases (men and women combined) among San Carlos residents over the 10 year period 1989-1998. According to the author's calculation, if the residents of San Carlos had experienced the cancer rates of the population of Quito, they would have experienced an estimated 7.5 cases of cancer (the "expected" number).

As discussed above, a comparison such as this one that contrasts cancer rates across two different geographic areas is susceptible to a variety of potential confounding factors that were not controlled in the present situation; these would include factors such as smoking, diet, socioeconomic level, and physical activity. Although uncontrolled confounding is an important concern, there appears to be a methodologic error in this study that is even more important. Few epidemiologists would describe the difference between 10 observed and 7.5 expected as a striking excess. Furthermore, the value of 7.5 for the expected number appears to be a serious underestimate. The author has excluded from the expected number calculation all types of cancer for which there were zero cases in the San Carlos population. Thus, he includes no expected number for lung cancer, colon cancer, head and neck cancer, and many other cancers that did not occur in San Carlos but do occur in Quito. Because he has counted all the cancer sites in which there were cases observed in San Carlos, but omitted all the cancer sites and the expected numbers for which there were no cases observed in San Carlos, he has exaggerated the difference between the observed and expected numbers. In fact, the sites for which cancers were observed in San Carlos account for fewer than half of the cancers that occurred among all sites in the study by Hurtig and San Sebastián. Therefore a rough estimate suggests that the expected number in the Yana Curi study is about half of what it should be. In that case, the correct expected number should be closer to 15, and the 10 cases of cancer that were observed would amount to a 33% reduction from the estimated 15 cases that would be expected if the residents of San Carlos had the cancer rates of those in Quito.

In summary, only one of these papers, the one by Hurtig and San Sebastián, represents a peer-reviewed epidemiologic study. The Hurtig and San Sebastián paper is an ecological study, with no exposure information at the individual level, and it fails to consider the full range of potential explanations for the modest differences that were reported. Its unbalanced assessment of the epidemiology makes it appear more of an advocacy exercise than a scientific paper, but even its authors would appear to agree that it alone does not provide a good basis for any causal inference relating to environmental effects on cancer occurrence. Neither of the other documents sheds any further light on the topic. The CESR report is not an epidemiologic study, and presents little data that could have been used in an epidemiologic study. The Yani Curi report, which is an unpublished manuscript that includes an investigation of a reported cancer cluster, appears to have a methodologic error that has led to a bias that reverses the direction of the association among residents of San Carlos. As a body of data, these reports collectively contain little material information about the relation between oil development in the Ecuadorian Amazon region and health effects among residents of that region.

Yours,

Kenneth J. Rothman
Senior Scientist

Félix M Arellano
Senior Scientist