

***ITER* Peer Review on Coal Tar Containing Shampoos Meeting Summary Report**

This peer review meeting was conducted by Toxicology Excellence for Risk Assessment (*TERA*); a non-profit organization dedicated to the best use of toxicity data in risk assessment. Expert peer reviewers donated their time and talents to provide an independent review of the assessment. The objective was a comprehensive overall review of the materials as provided by the combined experience of all the reviewers. This meeting summary represents the major discussions and conclusions of the panel as a whole.

After brief introductions, the meeting began with a discussion of conflict of interest. Each reviewer had certified in writing prior to the meeting that he or she did not have a conflict (real or apparent) with the chemicals under review or with the sponsors (identified to the reviewers before the meeting), or identified the potential for such conflicts. *TERA* staff discussed possible conflicts with each reviewer to determine if measures were needed to manage a potential conflict (or appearance of conflict). Options included excluding the reviewer from a particular discussion and consensus, or allowing the reviewer to participate in the discussion, but not be polled for consensus. *TERA* presented a proposed plan for managing conflict of interest to the panel. Panel members each identified themselves and noted any possible conflicts.

Dr. McGee noted that 15 years ago he had worked for ICF for approximately 6 months. He did not feel that this was a conflict. The panel agreed. Dr. Roy and Dr. Findlay each noted that they were currently doing work for one of the Johnson & Johnson (J & J) companies, but not Neutrogena. Dr. Roy is conducting bioavailability research on topical antifungals and Dr. Findlay is working on latex gloves. In addition, Dr. Roy noted that he had a discussion of coal tar issues with Neutrogena scientists during his poster session at the 2000 Society of Toxicology meeting. Some of the reviewers felt that the ongoing work for J&J companies, but not the scientific discussion, could present the appearance of a conflict. None of the panel members felt that this work presented an actual conflict of interest. The panel concluded that the conflict of interest statement should be revised for Drs. Roy and Findlay to include a description of the work being performed for the J & J companies, and that both Drs. Roy and Findlay should participate fully in the discussion and polling for consensus. However, all panel members were asked to look for undue influence on the discussions, and any panel member was given the ability to approach the chair to state an issue of undue influence. The chair would then decide whether additional panel discussion would be needed. The panel unanimously agreed with this approach.

Dr. Brian Adams noted that he had talked to Neutrogena scientists at a scientific poster session, but had not done any work for them. The panel did not see this as a real or perceived conflict. None of the other peer reviewers had a real or perceived conflict of interest and the plan for managing conflict of interest was approved as amended. The approved statement is attached as Appendix A.

This review meeting followed a standard *TERA* process, beginning with a close examination of the supporting documentation and important references by the panel in the several weeks prior to the meeting. At the meeting, the authors of the assessment briefly presented their work. The panel then systematically discussed the assessment, starting with a discussion of the qualitative weight of evidence, followed by a discussion of the quantitative aspects of the assessment.

Full discussion and participation were encouraged and agreement was reached by consensus. Consensus for the purpose of these meetings is defined as "an opinion held by all or most, or general agreement." The meeting was open to the public and individuals from Orrick, Herrington & Sutcliff, LLP; American Home Products; Neutrogena Corporation; Gradient Corporation; Occupational Knowledge International; Preston Gates Ellis LLP; and Murray & Associates observed the proceedings.

Estimation of Lifetime Skin Cancer Risk from the use of Coal Tar Containing Shampoos

Sponsor: Orrick, Herrington & Sutcliff, LLP on behalf of American Home Products and the Neutrogena Corporation

Presenters: Mr. Bruce Allen, ICF Consulting, The K.S. Crump Group, Inc.

- Dr. Jeff Bounds, ICF Consulting, The K.S. Crump Group, Inc.
- Mr. Harvey Clewell, ICF Consulting, The K.S. Crump Group, Inc.
- Dr. Annette Shipp, ICF Consulting, The K.S. Crump Group, Inc.

Chair: Dr. Michael L. Dourson, *TERA*

Review Panel:

- Dr. Brian Adams, University of Cincinnati, College of Medicine
- Dr. Michael L. Dourson, Toxicology Excellence for Risk Assessment
- Dr. Brent Finley, Exponent
- Dr. John Gamble, Exxon Mobil Biomedical Sciences, Inc.
- Dr. Loren Lund, Parsons Engineering Science, Inc.
- Dr. Brian Magee, Ogden Environmental & Energy Services
- Dr. Rebecca Parkin, George Washington University, Department of Environmental and Occupational Health
- Dr. Tim Roy, Petrotec
- *Dr. Ted Simon, U.S. Environmental Protection Agency, Region 4
- Dr. Alan Stern, New Jersey Dept. of Environmental Protection
- Dr. Allan Susten, Agency for Toxic Substances and Disease Registry, Division of Health Assessment and Consultation
- Dr. Glenn Talaska, University of Cincinnati, Department of Environmental Health

- Dr. Rebecca Tominack, Missouri Regional Poison Center

*Could not attend the meeting, but provided written comments which were distributed to panel members, authors and observers.

HAZARD CHARACTERIZATION

Author Presentation and Clarifying Questions from the Panel

Dr. Annette Shipp, ICF Consulting, presented an overview of the hazard characterization conducted as part of the risk assessment for coal tar shampoos. The purpose of the risk assessment was to develop a No Significant Risk Level (NSRL), defined under California's Proposition 65 as the lower bound on dose at a 10^{-5} extra lifetime risk, and an estimate of exposure to coal tar from the use of coal-tar containing shampoos. The carcinogenic weight-of-evidence of pharmaceutical grade coal tar was evaluated by considering two types of evidence: primary evidence consisting of epidemiology and animal studies, and secondary evidence consisting of pharmacokinetic and pharmacodynamic studies. Animal studies consist primarily of mouse skin-painting bioassays. Clear evidence exists that coal tar is carcinogenic to mouse skin and that the PAH (polycyclic aromatic hydrocarbon) fraction is the dominant contributor to carcinogenicity. However, potency is not characterized by BaP [benzo(a)pyrene] content of the coal tar mixture alone, but is dependent on the constituents of the mixture. Therefore, the appropriate dose metric is the amount of coal tar applied. Of the animal studies considered, the Fraunhofer Institute of Toxicology (1994) study was the only study with an adequate characterization of the dose-response.

The epidemiology studies consist of both occupational studies and studies of patients treated with coal tar ointments. Occupational studies suggest a relationship between chronic exposure to high levels of coal tar and other PAH containing materials with skin and scrotal cancers. However, concurrent conditions such as poor personal hygiene, chronic dermal irritation, and co-exposure to UVB light confound assessment of the PAH-containing materials per se. None of the occupational studies were determined to be suitable for quantitative risk assessment. In psoriasis patients treated with coal tar, no significant increase in cancer was observed in several studies of patients receiving the Goeckerman treatment (coal tar ointment plus UVB light) (Stern and Laird, 1994; Stern et al., 1998; Jones et al., 1985; Pittelkow et al., 1981; Grupper and Berretti, 1980; Menter and Cram, 1983; Lindelof and Sigurgeirsson, 1993). Only one study (Maier et al., 1996) reported a significant increase of basal cell carcinoma (BCC) and squamous cell carcinoma (SCC), but not SCC alone, following coal tar treatment in patients who also underwent PUVA treatment for psoriasis. No significant increase in the incidence of skin cancer was noted in patients treated for other chronic skin conditions using the Goeckerman treatment (Maughan et al., 1980; Jemec and Osterlind, 1994). There is no evidence that psoriasis confers protection against developing skin cancer; several studies

indicate that psoriasis patients developed skin cancer at about the same rate as individuals with normal skin (Alderson and Clarke, 1983; Bhate et al., 1993; Stern et al., 1984).

In order to describe the differences between mouse and human skin, a comparison of skin anatomy and physiology, absorption, enzyme activity, and metabolism was conducted. Mouse skin is thinner, has fewer cell layers, and a faster cell turnover than human skin. Absorption of BaP (a reasonable surrogate for other PAHs) is about 60% in mouse skin and 30% in human skin. Basal levels of AHH (aryl hydrocarbon hydroxylase) and induction of AHH by coal tar differed by a factor of 10 between mouse and human skin. Compared to human skin, mouse skin produced more total metabolites of BaP, more total diol metabolites, and about 160 times more of the 7,8-diol metabolite, which is the precursor of the ultimate carcinogenic metabolite. Mouse skin cells readily undergo neoplastic transformation when cultured in the presence of BaP. Human cells did not undergo neoplastic transformation even though they were incubated with higher concentrations and for longer durations than mouse cells. Administration of carcinogenic PAHs at doses that caused tumors in mice failed to produce tumors in human skin grafted to nude mice; all treatments resulted in skin tumors of mouse origin and at the graft margins in 100% of the mice in each group. Similar levels of PAH-DNA adducts have been detected in mouse skin and human skin exposed to coal tar. Human skin has more efficient DNA repair mechanism than mouse skin. Induction of ODC (ornithine decarboxylase) activity, thought to be an obligatory precursor step in mouse skin tumor promotion, appears to be mediated by products of COX2 (cyclooxygenase-2) activity. In contrast, the induction of ODC in human skin was independent of lipoxygenase and COX2 pathways. Therefore, it is questionable if the promotional mechanisms operative in mouse skin are also operative in human skin. Because of the significant differences between mouse and human skin in physiology, absorption, enzyme activity, and metabolism, it was determined that studies in mouse are not relevant to humans and that a quantitative risk assessment should be developed from the human data. The Pittelkow et al. (1981) study was determined to be the most appropriate because it had the best characterization of exposure. This study would be described more fully during the dose-response assessment presentation.

One reviewer asked the authors to describe the therapeutic mechanism of coal tar, because this information may help to answer the question of whether psoriasis patients are the appropriate population to study. The reviewer understood that the Goeckerman treatment sensitizes cells to killing by UVB light; therefore, the treatment could be selectively removing the cell types that could develop into tumors. Alternatively, this reviewer thought that perhaps the treatment may not only be slowing the cell cycle, but also may be killing cells which bind the coal tar constituents to their DNA, thereby keeping those cells from proceeding to tumors. An author was not aware of the therapeutic mechanism, but indicated that using the Goeckerman treatment in other types of patients, such as those with atopic dermatitis, does not affect skin cancer incidence, which might be expected if coal tar was killing cells that could develop into tumors. Another reviewer indicated that there are no data on the therapeutic mechanism of the Goeckerman treatment, however, the treatment does not kill the basal cells and keratinocytes. Rather, it prevents them from being hyperproliferative. The treatment does

not induce mitotic arrest, but rather affects the lymphocytes present in the skin that release interleukins, which in turn stimulate cell growth.

One reviewer asked if the Fraunhofer Institute of Toxicology (1994) study was the only mouse study that the authors found adequate for risk assessment. An author indicated that it was. A reviewer also asked if studies other than the Pittelkow et al. (1981) study were suitable for risk assessment. The author indicated that Pittelkow et al. (1981) was the only one that fully characterized the demographics of the populations and the amount of treatment so that exposure could be characterized.

An author was asked whether the ICF scientists working on the assessment knew the identity of the sponsor of their work. The author indicated that they did.

One reviewer asked if there were data available on the other constituents in coal tar and their action on the human skin response. An author indicated that except for some of the PAHs, not much data on the other constituents was available, and that only the data for BaP were complete enough to allow for a comparison between mouse and human skin.

A reviewer asked which study provided the basis for the statement that the basal levels of AHH are 10-fold higher in mouse than in human. An author indicated that these data are from the Storm et al. (1990) study and are presented in Table 1 of the risk assessment.

One reviewer asked if the charge to the reviewers was to decide if coal tar shampoos required warning labels under California's Proposition 65. The Chair indicated that the panel was to reach consensus on the technical merits of the risk assessment document only and was not to address Proposition 65 issues.

Observer Presentation and Clarifying Questions from the Panel

Written pre-meeting comments were received from one observer, Mr. Perry Gottesfeld of Occupational Knowledge International and a co-litigant with California EPA on the Proposition 65 case. These comments were distributed to the panel the week prior to the meeting. Mr. Gottesfeld briefly presented his technical comments on three additional epidemiology studies (Maier et al., 1996; Stern et al., 1998; and Letzel et al, 1998) which suggested that there is an increased squamous cell carcinoma (SCC) risk from dermal exposure to coal tar. He noted that there are very few pharmacokinetic studies on coal tar mixtures and questioned whether the pharmacokinetics of BaP is appropriate for the analysis of these mixtures. He also said that the skin is the target organ and that absorption is a secondary issue of concern.

A reviewer asked if the Stern et al. (1998) study cited by Mr. Gottesfeld only looked at the effects of PUVA (Psoralen and Ultraviolet Radiation A). Mr. Gottesfeld indicated that Table 6 in the study presents results for coal tar/UVB as well. An author indicated that the odds ratio of 1.4 for this treatment was not statistically significant and that the study authors concluded that the relative risk was small.

Prior to the meeting the Chair had briefly corresponded with Dr. Edward Emmett, a dermatologist from the University of Pennsylvania, and posed the question "How appropriate is psoriatic skin for estimating cancer risk for persons using coal tar-containing shampoos?" Dr. Emmett indicated that it was appropriate under the following conditions:

1. Psoriatic skin is not more or less likely to develop skin cancer (all histologic types) given the same dosage of carcinogen.
2. Absorption across psoriatic epidermis is the same as across non-psoriatic epidermis.
3. Metabolism/binding/destruction of PAH and PAH-DNA adducts is the same in psoriatic and non-psoriatic skin.

A panel member asked that Dr. Emmett be contacted to provide references or confirm that his opinion is based on his experience.

Discussion

Human Data

One of the panel initiated the discussion on human data by indicating that, overall, the conclusions made in the risk assessment document regarding the human studies were accurate and appropriate. This reviewer noted that the occupational studies of chimney sweeps could not be used in this risk assessment because no exposure data were available. Studies in coal tar/pitch workers are confounded by the fact that these workers had sustained exposure to the sun. Two studies, Maizlish et al. (1988) and Hansen et al. (1989) should not have been included in the document because the workers in these studies were not exposed to coal tar. This reviewer indicated that several studies of cutting oil workers should have been included in the risk assessment, including Hendricks et al. (1959), Jarvholm and Lavenius (1987), Jarvholm and Easton (1990), and Jarvholm et al. (1985), along with aluminum reduction plant workers who had exposure to coal tar. The reviewer thought the study in aluminum reduction plant workers by Spinelli (1991) might be a useful study for extrapolating exposure. Studies of coal tar/creosote workers should be added for the qualitative assessment. Likewise, foundry workers have coal tar exposure, but only one study of these workers was conducted, Sherson et al. (1991), and they found no increase in skin cancer risk. Finally, this reviewer noted that there was not a consistent observation of increased relative risk in any of these studies, even for lung cancer. This may be due to the differences in the coal tar mixtures.

Another reviewer observed that none of the occupational studies included women or children and that transfer of coal tar compounds from mother to fetus was an unanswered question. Also, the impact of early childhood exposure to coal tar products on the likelihood of an adult developing skin cancer is not addressed in any of the human studies. Therefore, this reviewer noted that it is important for the risk assessment to account for total life exposures. An author indicated that in the Maughan et al. (1980)

study on atopic dermatitis patients, one patient was an infant. Several of the studies indicated that psoriasis first developed in several patients in the late teens to mid-30s. However, there were no data to assess risk to children. An author acknowledged that this uncertainty was incompletely addressed in the draft report.

Since it is known that dermally absorbed PAHs enter the systemic circulation, one reviewer asked why the discussion was limited to just skin cancer and whether an evaluation was done to see if other organs were affected. An author indicated that there were not adequate data to attribute systemic cancers to exposure by the dermal route.

Another reviewer indicated a concern regarding the controls used in the Pittelkow study. This reviewer suggested that the most appropriate control populations were those that were comparable to the study group at the time of diagnosis, not controls with overall similar residential history. In the 1950s, most of the study population was likely to have been fairly local to the Mayo clinic in Minnesota. Therefore, using controls from Sunbelt states such as Texas and California was not appropriate. Additionally, different cancer registries reported skin cancer differently and skin cancer was difficult to diagnosis. Errors would be minimized by using controls from a single location.

A reviewer highlighted another concern with the Pittelkow study -- that apparently half of the study population had died by the time of the follow up and that most of the data gathered in the study was by either personal or family member recall. Therefore, it was likely that the study did not gather adequate information about the people who died before follow up. A reviewer also noted that there is not information in the Pittelkow study about the extent of the respondents who said "don't know" to any of the questions. From the data presented, it is not possible to determine how much information is missing because people were not available to be asked (e.g., dead, lost to follow up) or because people were asked but provided no answer.

Reviewers next considered the issue of whether psoriatic skin was the appropriate basis for an assessment of skin cancer risk from coal tar shampoos. A reviewer noted that there is no doubt that psoriatic skin is different from normal skin, but there is no biochemical evidence that it is more or less likely than normal skin to develop cancer. When considering the epidemiology data there is no evidence that psoriatics are more (or less) likely to develop skin cancer. In addition, very few patients receive a monotherapy; they are treated for a lifetime with various treatments (including self-therapy in the sun).

A reviewer noted that people who use coal tar shampoos do not have normal skin and asked how different their skin may be from that of psoriasis patients. Another panel member indicated that coal tar is most often prescribed for psoriasis and dermatitis, but that coal tar benefits people with a wide variety of skin conditions. There is not a single standard frequency or regimen of coal tar use. Regimens may range from every day to once every two weeks.

One reviewer discussed reservations with the Pittelkow et al. (1981) study. Principally, this reviewer was concerned that DNA-coal tar-UVB interactions were slowing the cell

cycle so that the cells were prevented from proceeding through the cell cycle to be transformed as tumor cells. This may result in psoriasis patients who receive the Goeckerman treatment being at less risk for skin cancer, compared with other psoriasis patients. This reviewer also noted that size of the cohort was small.

One reviewer indicated that another uncertainty with all the studies is that exposure is to different mixtures. This reviewer noted that coal tar extracts and distillates are different from whole coal tar, which is different from coal tar pitch, etc. Even coal tar shampoos vary in their use of whole coal tar, coal tar extracts, and coal tar distillates.

The panel reached unanimous consensus that the Pittelkow et al. (1981) study was the only existing study from which reasonable estimates of human risk can be made. However, several limitations of this study were apparent. A sensitivity analysis would be helpful and explicit discussion of the impact of these limitations should be presented. The panel discussed a number of these limitations, including:

- Uncertainty as to which individuals got cancer and whether any of them were exposed to coal tar as infants
- The reliability of reporting in Pittelkow (e.g., limitation for 20 patients lost to follow up)
- The potential underreporting of skin cancer for those who were dead at follow up
- How representative the Pittelkow cohort was to estimate risk to a larger population

See also discussion with dose-response for other limitations

Animal Data

The panel discussed the appropriateness of the animal data to qualitatively describe human risk.

One reviewer noted that laboratory animals are selected to be sensitive but that the range of human sensitivity is not known. For example, in one human study, a 100-fold range in normal AHH activity was observed. Therefore, it is difficult to quantitatively describe the difference between mouse and human without an adequate understanding of human variability.

Reviewers discussed the data available from the xenograft studies to address human variability. It was noted that skin from 60 different people was used in the xenograft study by Soballe et al. (1996), but there was not any information that would allow characterization of different individuals as sensitive or not sensitive. One reviewer suggested that the xenograft studies could be used to show the magnitude between human responses in the same way that these studies show the dramatic difference between mouse and human responses.

Another reviewer indicated that the xenograft studies demonstrated very strongly that mouse skin is irrelevant for predicting human risk. However, a reviewer noted that overall mouse skin painting bioassays have been 90% accurate in predicting human tumors (all cancers, not just skin cancer) (Nesnow and Lewtas, 1991). Another reviewer noted that the mouse model has been shown to accurately predict carcinogenic potency, although there is less information on how this ranking is relevant to humans.

One reviewer pointed out that there are significant limitations to using the Fraunhofer study to quantify human risk. Several reviewers could not understand how one could use the mouse for a quantitative estimate for humans because the mouse skin does not have the same progression to cancer; there are different pathways between the species. The study itself is not bad, but it is not appropriate to use the mouse skin as a model for deriving a potency factor when the mouse has both initiation and promotion pathways which are different from humans.

Another reviewer thought that if the qualitative uncertainty to tumor promotion and initiation is so great it precludes use of the mouse model, perhaps one could bound the qualitative differences leading one to determine whether it is appropriate or not to use the mouse model. Another reviewer suggested that the limitations of the model have been discussed in the document, but the door was left open for a quantitative estimate.

Overall, the panel agreed that the document's conclusions regarding the animal studies were sound and appropriate. The panel discussed the nature of the material tested in the various mouse studies. Reviewers noted that none of the animal studies tested coal tar at multiple dose levels with a multiple dosing design without promotion with a separate promoting agent. Therefore, none of the animal studies were suitable for a risk assessment of coal tar *per se*. The animal studies used mixtures, such as creosote, which are dissimilar from coal tar because they contain components which may cause skin irritation (at specific doses) or act as promoters. These differences need to be acknowledged and discussed in the risk assessment. The panel agreed that the results of the quantitative risk assessment using the animal studies are most applicable to the exact mixture used in the referenced study.

Although most of the panel agreed that the mouse data are not an appropriate basis for a model to estimate human cancer risk from exposure to coal tar shampoos, it was recognized that additional discussion on this issue would have been beneficial. The panel unanimously agreed that the human data were more appropriate for estimating risk. The panel reached unanimous consensus that the animal analysis should be included in an Appendix for comparison purposes only and that the limitations should be made more explicit in the discussion.

Pharmacokinetics

Reviewers first addressed the question of whether BaP content can be used to estimate the potential carcinogenicity of coal tar mixtures. As part of this issue, reviewers also considered whether BaP is a reasonable surrogate for comparison of pharmacokinetic

differences between mouse and human skin responses to coal tar mixtures. Reviewers noted that the components of PAH and coal tar mixtures, including the vehicle, influence the carcinogenicity of the mixtures. While BaP equivalents are difficult to use in risk assessment, many agencies are using them as a measure of exposure to PAH containing mixtures. However, BaP equivalents will not accurately predict the carcinogenicity of the mixture. One reviewer suggested that a discussion about the mode of action for skin cancer, the tumor type of concern, would help address the question. Possibilities include whether AHH enzyme activity is the appropriate measure of carcinogenicity or whether there is a particular DNA adduct or metabolite that accurately predicts carcinogenicity.

An author indicated that BaP content alone is not suitable for a quantitative estimate of potency of coal tar mixtures; although the data clearly indicate that the PAH fraction is largely responsible for the carcinogenicity of coal tars. BaP and DMBA (7, 12-Dimethylbenz(a)anthracene) are the most potent of the PAHs. Metabolism of PAHs follows two general metabolic pathways and BaP is an appropriate surrogate for estimating metabolism by one of the pathways. BaP is also an appropriate surrogate for estimating absorption of PAHs. Relative AHH induction is a good measure for species differences. Both BaP and DMBA are good surrogates for evaluating the multistage process of carcinogenesis. Therefore, the author concluded that BaP is a good surrogate for making species comparisons, if not for making absolute statements regarding the potency of a given coal tar mixture.

A reviewer suggested that it might be possible to develop a predictive model for carcinogenicity based on PAH content of mixtures that is similar to an approach used by the petroleum industry to characterize different petroleum streams. This approach defines a boiling point range, and then plots PAH content versus carcinogenicity with good correlation.

The panel reached unanimous consensus that it is inappropriate to use BaP content of a coal tar mixture to predict the absolute potency of that mixture. Reasons mentioned by the reviewers include:

- Mixture components could bind receptors without having a toxicological action; therefore the components could have some inhibitory effects.
- Concentration of BaP-DNA adducts is not predictive for cancer risk, so BaP cannot be used to predict the overall carcinogenicity of a mixture.
- Predicting carcinogenic potency based on BaP content does not account for the promoting effect of other mixture components.

However, in general, the panel agreed that BaP is a reasonable surrogate for understanding the different responses between mouse and human skin in order to conduct a species extrapolation. Reasons mentioned by the reviewers include:

- The volume of data on BaP and the lack of data on other PAHs means that BaP is the only useful surrogate.
- There is good correlation between PAH content and carcinogenicity for coal tar mixtures

The panel next considered the issue of whether the differences between mouse and human skin can be quantitatively described and whether a kinetic model can be developed. One reviewer opened the discussion by indicating that he agreed that the differences between mouse and human skin could be described and modeled. However, several reviewers indicated that any further quantitative analysis beyond what was done in the document would be too much. The revised draft should more clearly state that the estimate from the mouse data is a "ball park" figure. In addition, the mouse analysis should be de-emphasized and the assessment should include some uncertainty bounds around the mouse estimate to give readers an understanding of how imprecise the mouse estimate would be.

Several reviewers mentioned that the summary paragraphs in the risk assessment report describing the differences between mouse and human are accurate. These paragraphs, particularly the information on repair and thickness of skin, contribute to the conclusion that the mouse is at higher risk of skin tumors than are humans.

One reviewer indicated that although there are enough data on passive processes to make quantitative Human Equivalent Dose (HED) adjustments to extrapolate from mouse to human, there are not enough data on active processes such as metabolism to make quantitative adjustments.

Several reviewers indicated that it is reasonable to say that mouse skin is more sensitive than human skin to BaP or PAH carcinogenesis. In a more general sense, it is not unreasonable to expect that mouse skin is more sensitive than human skin to coal tars, but data do not exist to support this statement.

A reviewer indicated that while it is reasonable to quantitatively compare the differences between mouse and human for the separate components of carcinogenic response, such as AHH induction, there are not enough data to "connect the dots" to quantitatively compare the difference between mouse and human in overall carcinogenic response. This reviewer also recommended that the variability seen in the studies should be better characterized. Another reviewer suggested looking at the variability among different mouse strains in AHH induction and other components of the carcinogenic response.

An author stated that a key question in pharmacodynamics is that the promotional pathway is relevant to mouse but not to humans. He does not believe that a pharmacokinetic HED based on likelihood of generating DNA adducts is relevant. What is really relevant is how long the promotional status in the animals is maintained. He questioned whether it is a function of irritation from lower molecular weight chemicals and activation of a pathway from higher molecular weight chemicals. Both of these swamp the question of interspecies extrapolation. The real question is whether there is a

pathway activated in the animal that is not activated in the human. A reviewer asked that this information be more strongly conveyed in the document.

The panel reached general areas of agreement regarding the pharmacokinetic discussion:

- BaP is not an appropriate surrogate for estimating carcinogenic potency of coal tar mixtures.
- BaP is a reasonable surrogate for understanding the differences between mouse and human skin.
- Mouse skin is more sensitive than human skin for specific PAH compounds (e.g., DMBA, BaP).
- In general, the document's conclusions regarding species sensitivity are appropriate, but the document needs to clearly specify the exact compound or mixture to which the mouse is more sensitive.
- The differences between mouse and human skin can be quantitatively described and modeled. However, generalizations to coal tar mixtures, while perhaps not unreasonable, should be done with caution and uncertainties need to be stated. Furthermore, since the use of mouse data is problematic, an analysis of the variability of responses in the various mouse strains is needed to put bounds of uncertainty on the mouse estimate.

DOSE RESPONSE ASSESSMENT

Author Presentation and Clarifying Questions from the Panel

Mr. Bruce Allen presented the dose-response approach used in the risk assessment document to estimate relative risk from the Pittelkow et al., (1981) study. The assessment was conducted based on a 10^{-5} risk level as required by California's Proposition 65. Relative risk is a function of dose and is estimated using a linear equation. Estimates of potency are based on $P(d)$ and $P(0)$ which are obtained from the Pittelkow study, along with dose. Because of data reporting limitations in the Pittelkow study, inputs were treated as distributions to represent the uncertainties. Members of the cohort were classified according to mode of treatment. There were 19 members with cancer and 241 without cancer. Note that the 241 was revised from that presented in the document (261) based on the earlier panel discussion that 20 members of the cohort could not be accounted for. Three treatment categories were clinic treatment only, clinic plus at-home treatment for <50 days/year, and clinic plus home treatment for >50 days/year. Treatment categories could not be determined for 3 patients with cancer or for 116 without reported cancer (116 was adjusted from the 136 used in the draft document). For those with unknown treatment, it was assumed that the probability for assignment to a treatment category could be based on the frequencies observed among those with known treatment.

Dose was estimated using a Monte Carlo analysis based on estimated clinic and home treatment exposures. Inputs to these equations included the percent coal tar in the

ointment, grams of coal tar applied during treatment, and days of treatment. For home treatment, additional inputs included fraction of body surface covered and years of treatment. The Monte Carlo simulation was run 100,000 times. For each iteration 260 (20 less than the draft document value of 280) individual doses were estimated and averaged to give a value for d. Estimates of P(0) ranged from 15.5 to 49.2, but the best estimate was 26.6. For P(d), the best estimate was 19/260, but values were simulated based on a Poisson distribution. That is, assuming the observed number of skin cancers was the result of an underlying Poisson process, it was possible to determine the relative likelihood of any particular value for p(d). Sampling according to this relative likelihood gave a distribution for P(d). A distribution of values was generated for b, and the authors chose the 95th percentile (to be conservative). This 95% value was used in a life table analysis to estimate the dose at 10⁻⁵ risk. Without uncertainty in P(0) this is 49 ug/day and with uncertainty in P(0) it is 45 ug/day (these numbers may change slightly when the upper bound and observed incidence is based on 260 individuals).

One reviewer asked how the triangular distribution for estimating P(0) was constructed. The author indicated that a median of 26.6 was used with a lower bound of 15 and an upper bound of 49. Another reviewer asked how "grams of ointment used" was estimated. The author indicated that it was based on a distribution of body surface area with a mean of 1.84 m². Another reviewer asked how the duration of clinic visit treatment was estimated. The author replied that the clinic treatments varied from 8 to 38 days with a median value of 16-17 days. This was what they had determined to be usual for the Goeckerman treatment. Each iteration was sampled from each distribution.

One reviewer asked about the process for randomly assigning probabilities for patients with unknown treatment and whether any of the iterations assigned all unknowns to the no home treatment group or all unknowns to one of the home treatment groups. The author indicated that they did not do this, but that this might be a reasonable way to bound the assessment.

Another reviewer asked the author to describe the procedure used to calculate the confidence intervals on the expected number of skin cancers. The author indicated that the frequency data followed a Poisson distribution (a standard assumption for frequency type data). The variance was equal to the mean and by definition is independent of sample size. The reviewer noted that with a small cohort such as this, the accuracy of estimating the true response was a function of sample size. The author noted that with decreased sample size the curve is spread wider, the maximum likelihood estimate would not be as high and the confidence intervals would be wider.

One reviewer noted that for duration of hospital treatment, the simulation estimated a higher median value than that actually reported in the study. An author indicated that they would adjust the analysis to change the peak of the triangular distribution. Another reviewer asked why exposure for two clinic visits or fewer was estimated when the Muller and Perry (1984) study showed that 11% of patients had three or more visits. An author indicated that Pittelkow et al. (1981) only reported that patients visited the Mayo

Clinic for the first time during 1951-55 and 36 patients revisited the Mayo Clinic during 1956-60.

One reviewer asked the basis for the estimates of duration of home treatment. An author indicated that it was based on patient responses to phone interviews or mailed questionnaires, as specified in Pittelkow et al. (1981).

Discussion

Human Dose Response

The first issue discussed was the influence of early childhood exposure on estimating exposure. One reviewer noted that if childhood exposure was considered in estimating the exposure for the dose-response assessment, then several assumptions about the distributions, particularly body surface area and absorption of residue, would change. The authors indicated that Proposition 65 directs that the NSRL look at the average person, not sensitive populations. For this document, exposure of the cohort in Pittelkow was estimated so that the dose response could be characterized, and this cohort did not include children. Under discussion of the exposure assessment information about potential exposure to children is presented. However, reviewers noted that it was likely that the Pittelkow cohort had some prior exposure before being included in the study, although there was no information on this available.

Because the incidence of skin cancer is based on self reporting or reporting by surviving relatives rather than medical records, a reviewer questioned the estimates of skin cancer reported by Pittelkow. However, another reviewer noted that in 64% of the cohort, historical specimens were reviewed to confirm skin cancer.

A reviewer asked if there were data on where a person lived (e.g., north or south) at time of diagnosis in the Pittelkow study and whether this information could be used in the analysis to calculate expected cancer rates. The authors indicated that this information could be used to weight a distribution. However, to recalculate the expected numbers one needs person years of follow up, broken down by where they lived. These data are not available. There was some confusion in interpreting the study write up as to whether the Pittelkow data represented where people lived at time of diagnosis or at time of follow up.

The panel discussed the accuracy of skin cancer reporting in the cohort compared with control populations. Reviewers noted that while misreporting would also be expected for control populations, the degree and or type of misreporting would vary from locale to locale. Therefore, it would be necessary to know more about where the people in the cohort lived at the time of diagnosis in order to select the most appropriate control population. In addition, skin cancer would not have been a reportable disease for the cancer registry in the 1950s.

The panel discussed how the authors might consider potential underreporting in expected and observed cancer incidence from Pittelkow to put bounds on the uncertainty analysis. The panel agreed that the authors should consider the potential underreporting in observed and expected incidence, given there is uncertainty in both Pittelkow and the background rates. They should use skin cancer reporting and accuracy rates to anchor the projections to bound the estimates on the expected. One reviewer closely examined the Scotto et al. (1974) paper, which reports that underreporting by a factor of 2-3 was the case from the mid-1950s to the early 1970s. With this information, the panel agreed that the Pittelkow data on expected cases are as good as are available and there is more uncertainty among the observed data than the expected.

The panel extensively discussed the methods for estimating exposure to the cohort in the Pittelkow study. The first issue addressed was the treatment of the three of 19 patients with skin cancer for whom no data on treatment was provided. The risk assessment document had assumed that these patients had similar probabilities of falling into any one of the three classes of treatment as the rest of the cohort of patients with skin cancer. One reviewer had submitted written comments that the treatment classification of a patient would depend on the severity of that patient's condition and not on the frequencies observed in the other patients. Therefore, this reviewer suggested that using a probability of 0.3 of being in any of the three treatment classes was a more appropriate approach. Reviewers also suggested using a larger cohort to define the probability that could be assigned to each treatment class. An author indicated that using a probability of 0.3 was a less conservative approach than the one they used and that no larger cohort was available for defining the appropriate probability. The reviewers suggested bounding the estimate by alternately assigning all of the unknown patients to each of the different treatment classes.

The reviewers discussed the assumptions and distributions that were used to estimate Pittelkow cohort exposure during clinic and at-home treatments with coal tar ointment. The reviewers agreed that assuming 2% coal tar concentration in the ointment was reasonable and that normalizing grams of coal tar per 1% ointment was appropriate.

A reviewer asked whether the distribution of men and women was taken into account in estimating body surface area for calculation of grams of coal tar used in the clinic. If there was an even distribution of genders, then the distribution used was appropriate, but if there were more of one gender, then the distribution was inappropriate. Reviewers suggested that the study authors be contacted to learn the genders of the cohort members. Alternatively, the National Psoriasis Foundation Internet site could be examined to find out if the disease is more prevalent in men or women. Finally, the reviewers suggested assuming that the entire cohort was either all men or all women in order to provide bounding estimates.

The reviewers discussed the exclusion of the scalp in estimating body surface area for calculation of grams coal tar used in the clinic. Reviewers noted that the standard Goeckerman treatment [as reported in Muller and Perry (1984)] includes shampooing with 10 to 20% solutions of liquid carbonis detergens in mineral oil, applied two or three

times a day with daily shampoo. Although some reviewers felt that an estimate of exposure through the scalp should be included, others felt that this was not appropriate since there was no indication in the Pittelkow study that the patients in that cohort had exposure by this route; therefore, scalp should not be included.

In discussing the number of days of treatment at the clinic, one reviewer suggested that a Johnson SB distribution would be more appropriate than the triangular distribution used in the risk assessment report. An author indicated that the Johnson SB distribution was complex, needed four input parameters, and that the shape of the distribution varied widely. They chose to use the simple triangular distribution because the data were simple. None of the other reviewers were familiar with appropriate uses for the Johnson SB distribution, and it was recommended that the triangular distribution be used.

In the Pittelkow cohort, it was known that six of the patients with cancer and 30 without cancer had one or more additional visits in the second five year period reported (some may have had one or more visits in the original five year period reported), but it was not known which of the patients these were. Therefore, in the analysis, a nonzero value for E_2 (second visit) was randomly assigned to six and 30 patients, respectively, and a value of zero was assigned to the remaining members. Reviewers agreed that this approach was reasonable.

Reviewers discussed the assumptions used to estimate amount of coal tar exposure at home. For the percentage of coal tar in ointments, the document cites literature indicating that ointments can range from 1% to 25% coal tar; however, the authors used a uniform distribution with a range of 1% to 5% to estimate this parameter. Reviewers suggested that a skewed distribution would be more appropriate. One reviewer noted that the data on coal tar concentration in ointments is from the 1990s and asked if this was relevant to coal tars used in the 1950s. An author indicated that 5% was commonly used in the 1960s and that it did not seem reasonable to use a higher concentration in the distribution. Another reviewer indicated that the Physicians Desk Reference (PDR) shows a range of concentrations for coal tar ointments of 2% to 6%, and that using higher concentrations in the distribution will not be conservative. Reviewers agreed that the approach used in the document was appropriate but that the document text should be revised to support the choice of the distribution.

In estimating the fraction of body surface area treated at home, the risk assessment used a uniform distribution with a minimum of 5% and a maximum of 25%. Reviewers questioned the reasonableness of this approach. One reviewer indicated that his analysis of the literature showed that a reasonable maximum body surface area treated at home was 30%. Other reviewers, however, questioned whether patients would actually cover this much of the body at home because it is time consuming and uncomfortable. Another reviewer indicated that 30% body area is the absolute maximum for psoriatic lesions on a person (the condition is found mostly on scalp, hands, elbows, and knees), but that the Goeckerman treatment covers all body surfaces. Follow-up use at home would be just for visible lesions and therefore, 5-25% body area is a reasonable range for this reason. The reviewer also mentioned that steroids have replaced coal tar for psoriatic treatment today

and that while coal tar is mentioned in textbooks as a possible treatment for cradle cap, it is not used for cradle cap today. An author suggested using a triangular distribution with 30% as the median and a range of 5-55%. However, a reviewer suggested that a skewed distribution might be more appropriate. The author then suggested that they look at several different distributions and ranges to see the impact of the different choices. The panel agreed with presenting alternatives and also suggested that the justification of the choices in the text be revised as discussed above.

Reviewers discussed the appropriateness of the assumptions for the number of days/year that patients treated themselves at home. An author indicated that the possible range was from zero days to every day for 27 years. One reviewer suggested a range of 25 to 150 days, while another reviewer pointed out that the actual data reported by Pittelkow was "used tar at home for 50 days/year or more" and the percentage who answered "yes" to this. It has been inferred that if they did not answer yes, then they used the ointment for 50 days/year or less. However, their response was "other" and there is no way of knowing what "other" means. Another reviewer suggested that it is unlikely that people would use coal tar products for less than one course of treatment, which is approximately 10 days. Therefore, the panel recommended that a lower bound of 10 days and an upper bound of 50 days be used for the "<50 days/year" treatment group.

The panel also questioned the maximum 365 days for at home treatment. This was based on the reported usage of a single individual in the Maughan paper on patients treated for atopic dermatitis, one of whom used "coal tar products...everyday use for 26 years" (Maughan et al., 1980). The panel asked *TERA* to check with another member of the panel with dermatology expertise to confirm that 365 day maximum was reasonable for use during the years of treatment for the Pittelkow cohort. The panel recommended that the bounds for the ">50 days/year" group be based on the known frequency of reoccurrence of psoriasis, and that this information be obtained from the dermatologist on the panel and provided to the authors.

The panel discussed the assumption for number of years of at-home treatment. One reviewer questioned the approach used in the risk assessment because approximately half of the cohort dropped out of the study due to death and the average years of follow-up for them might have been only 12 years. Therefore, the distributions need to be structured to take this into account. One reviewer suggested that the data needed for constructing this distribution might be found in the Maughan et al. (1980) study of patients with atopic dermatitis.

A reviewer asked how a study without enough power to detect an increase in cancer incidence above background could be used for this analysis. An author explained that first, it is a mistake to think of power as a fixed and constant entity, and that "enough power" is ill defined. More importantly, whether or not a null hypothesis of no effect is rejected by some hypothesis test procedure is irrelevant when the issue is one of estimation, as it is here. The panel suggested that this issue be better explained in the document.

A reviewer commented that the authors had assumed a linear dose-response, and asked why if there is evidence of non-linear mode of action, they did not present a non-linear analysis. An author responded that they generated a relative risk from the Pittelkow data which, because it had only one dose group, could only be used to estimate a single parameter. The simplest and most commonly applied one-parameter model is a linear model.

Several panel members questioned whether net grams of coal tar in the ointment applied is the appropriate dose metric for estimating exposure in the psoriasis patients. They suggested that it would be more appropriate to use the dose of coal tar that penetrates and stays in the skin (and is bioavailable) rather than the amount of coal tar in the ointment applied. A major concern was that the current approach for estimating dose includes a significant portion of the dose that is in the ointment that does not have contact with the skin. Several reviewers thought that whether a thin layer or larger amount is applied, there would be no real difference in bioavailable dose because only that portion in contact with the skin will be absorbed and available. For ointment, its viscosity factor and ability to retain from the partition coefficient standpoint will determine penetration into the skin. One reviewer noted that the loading of the ointment in Pittelkow was 8 mg/cm^2 , which is not a monolayer and that the ointment was wiped off at the end of the day. Ointment used frequently would be more akin to a long-term steady state or infinite dose.

The panel discussed how ointments and shampoos might compare. A reviewer noted that the shampoo scenario is different than the ointment in that shampoo is applied to a wet scalp, the PAHs are quickly diluted from the shampoo, and the partition coefficient will probably drive the PAHs toward the stratum corneum. In a bathing situation the skin acts like a solid phase sorbent, the PAHs get picked up rather readily, are partitioned to the stratum corneum, the shampoo is washed off, and what was absorbed will be a quick and finite dose, perhaps a small percentage of what was actually applied. Reviewers noted that the coal tar ointments and shampoos are not applied at comparable rates and differences in flux and time between the shampoo and ointment may be important.

An author noted that Van Rooij et al. (1993a) measured absorption of coal tar ointment in human volunteers by fluorescence disappearance and found 22% absorption over 6 hours.

A panel member indicated that application rates are highly variable even with patient training. The reviewer estimated that when psoriasis patients apply the coal ointments at home they may apply 2-4 times more ointment than recommended. Some excess is likely in the hospital setting, but not to the same degree. A reviewer asked whether there are any studies of applying and wiping off and then measuring what remains behind. The authors indicated that there are such studies and explained further what they have done in their assessment. In their approach they have used the metabolic data to back calculate the residue of coal tar from shampooing. Their estimates are reasonably bounded by exposure data, but they make an implicit assumption that the applied dose to the patient and the person shampooing will result in the same absorption. They redid their analysis with a distribution of absorption from 25 to 100%. They assumed the same distribution for the psoriasis patient, as there are no data to suggest that psoriatic patient skin has

more penetration. With the revised analysis, the dose to the patient and the person shampooing is reduced, but the ratio does not change very much.

The panel also discussed the difficulty in understanding due to different permeabilities in various PAHs.

The panel agreed that the most important factor is the concentration of coal tar in the skin. Therefore, whether patients may have used an excess of ointment in home treatment will not make a difference in dose. Looking qualitatively at the underestimate in hospital versus overestimate for in home use might get at some concerns.

The panel unanimously agreed that it would be more appropriate to use the dose of coal tar that penetrates and stays in the skin rather than the coal tar applied. They recommended that to strengthen the risk assessment the authors do an analysis on the basis of absorbed dose or flux over time for shampoos and ointments. The risk assessment should characterize the existing Monte Carlo assessment. This characterizes a balance of the potential overestimation of applied dose in the hospital with the potential underestimation of applied dose in home use. One reviewer suggested looking at the Van Schooten et al. (1994) data and correlating with data from Van Rooij et al. (1993a and b) on fluorescence disappearance related to absorption.

A reviewer asked how the model handles uncertainty in the 19 observed cases from Pittelkow and if one started with something other than 19, how one would capture the uncertainty. The authors noted that potential underreporting for those who died is a qualitative uncertainty that may be addressed in the uncertainty section. Another reviewer noted that years of follow up were reduced for those who died, suggesting that this reduced the probability of detecting cancer which otherwise would occur, and asked how this is accounted for. The authors responded that the expected numbers are based on the years of follow up and that, therefore, the observed number is appropriate for the expected numbers reported.

Another reviewer suggested that in the uncertainty analysis the authors could make some assumptions of perhaps 10, 20, or 30% underreporting by relatives for those who were dead and may have had cancer, but their relatives did not report. Another reviewer thought there might be data in the literature to help determine what might be typical in these cases that could help bound this uncertainty estimate. However, several reviewers agreed that it would not be appropriate to adjust observed incidence without adjusting expected incidence.

A reviewer questioned if it is appropriate to start the life table analysis at age one when the risk is based on an adult cohort. An author indicated that if one assumes that relative risk is the same across all ages, then one could start the life table analysis at any age point. The panel recommended that the life table analysis start with age one, rather than age 15, because coal tar shampoos may still be recommended for children.

The panel discussed whether the Pittelkow data should be stratified by squamous cell carcinomas (SCC) and basal cell carcinomas (BCC), because SCC are more susceptible than BCC. The SCC response is diluted when combining the types. The authors explained that the expected values that they used were for the combination based on age specific rates for the Pittelkow cohort. Several reviewers indicated that to stratify by type, one would need cohort person years and that there is not enough detail in the Pittelkow study to quantify this. The panel recommended that the authors address this qualitatively in the uncertainty discussion.

Another reviewer asked if there were ways to analyze the twenty individuals lost to follow up.

The panel agreed that the approach used in the risk assessment to calculate the one in 100,000 risk was appropriate, and suggested the authors use the best guess with Pittelkow, but do a sensitivity analysis for expected and observed. In addition, they should review the literature for information on skin cancer underreporting. Reviewers noted areas of uncertainties, including that the sample size potentially affects the 95% upper confidence limit, the uncertainty of the effect of the Goeckerman treatment compared to coal tar alone, and identification of an appropriate comparison population.

The panel also agreed that an overall sensitivity analysis should be done with input parameters ranked as high, medium, or low for their influence on the outcome. A qualitative ranking (high, medium and low) of uncertainty will give one a sense of what needs more care in bounding. Bounding could be done separately for each parameter, and then for a reasonable combination in one bounding analysis.

The panel acknowledged that the analysis might change if the authors base it on absorbed dose (that which penetrates the skin and is bioavailable) rather than grams of coal tar applied in the ointment. Some of these suggested items can be incorporated into the quantitative estimates while others are ways to characterize uncertainty and would be discussed in the risk characterization.

Animal Dose Response

Confronted with a shortage of time, the panel decided that a full animal dose response discussion would be deferred until the exposure assessment and risk characterization discussions were completed. The panel therefore discussed only briefly some aspects of the animal dose response modeling.

One reviewer asked why the Weibull model was chosen, and whether and to what extent the coal tar in the Fraunhofer study was relevant to what is in shampoo. An author responded that the multistage Weibull model was used to model time to tumor data to adjust the less than lifetime Fraunhofer results of 78 weeks to 104 weeks for both the light and heavy coal tar data. Including a time to tumor analysis is a standard practice used in many risk assessments and results in greater potency and a larger estimate of risk. The authors used the heavier coal tar data with the upper dose group, which also resulted

in greater potency. An author noted that the Fraunhofer study saw ulceration, which was progressively more severe and more frequent with increasing dose, but that they made an assumption that the ulceration had no impact on tumor formation. They assumed the dose response to be linear.

One reviewer asked why fraction of body surface area was included in the HED equation in addition to the volume of skin. An author indicated that this factor is recommended by U.S. EPA guidance (U.S. EPA 1992) and that it calculates the concentration of active species in the skin. This adjustment accounts for the fact that the whole organ (skin) is not being exposed. The reviewer questioned whether skin is considered to be an organ in that sense and that the only relevant location is the point of application. The author responded that a default value of 1 is used for the factor, so omitting it does not quantitatively change the result.

Several reviewers thought it might be helpful to include a measure of the uncertainty of the animal NSRL, rather than just presenting a point estimate. An author noted that one cannot bound the uncertainty in the pharmacodynamics and that a distribution would provide a perception of accuracy greater than is appropriate given the data. The panel briefly discussed whether the animal analysis should even be included in the document. The panel reached consensus on a recommendation to the authors that the animal analysis should be kept in the document for comparison purposes. It should be enhanced with parameter distributions when possible to characterize the highly uncertain nature of the analysis. They recommended that the analysis be put into an appendix.

EXPOSURE ASSESSMENT

Author Presentation and Clarifying Questions from the Panel

Mr. Bruce Allen of ICF Consulting made a brief presentation on the exposure assessment. He presented the Lifetime Average Daily Dose (LADD) equation and noted that they calculated the LADD for coal tar concentrations ranging from 0.5% to 2.5%. Ounces of shampoo used each year, number of years usage, and fraction of coal tar in applied shampoo that is left as a residue were treated as variable inputs. These inputs were considered to be distributions representing variability across shampoo users. The distribution for ounces of shampoo (U) was estimated from market surveys (Nielsen, 1999) and phone interviews (Toppmeyer 1994). The years of usage distribution (Y) was estimated from phone interviews (Toppmeyer 1994) and literature on duration of psoriasis treatment (Menter and Cram, 1983). The fraction of coal tar in applied shampoo that remains as a residue (F) is that portion of applied shampoo that remains in contact with the skin. Shampoo residue studies indicate this parameter should be 1-2%. Absorption and metabolism studies measuring pyrene in urine were used as a surrogate measure (Van Schooten et al., 1994; Van Rooij et al., 1993a and b; Jongeneelen and Bos, 1990; and Singh et al., 1995). With this information, a Monte Carlo simulation was run

100,000 times, sampling from all the distributions to yield a distribution for the LADD for hypothetical shampoo users. The median from this was then compared to the NSRL.

Mr. Allen presented some unpublished data from J & J that ICF recently obtained. In this study coal tar shampoo was applied at a rate of 5 uL/cm² to a 20 cm² patch of volar forearm skin of human volunteers. The shampoo was left in place for one minute, rinsed with a minute amount of body temperature water (100 mL) and the residue was measured using fluorescence. This study found 15.2% of applied coal tar shampoo remained as a residue. This value is much higher than the value from Von Rooij et al. (1993) cited in the document (0.9%), and other, less directly obtained, values.

To account for this new information, ICF considered this F value of 15% to be possible and included it in the range of what one might experience. It is a relatively unlikely value falling in the 99th percentile of the range from a log normal distribution with a mean of 1-2%. Given a 2.5% coal tar shampoo, adding this value to the distribution increases the average LADD from 26.6 to 33.3 ug/day. Use of these data in this manner provides an overestimation of importance of the one data point. This reanalysis was based on the parameter F representing the fraction of coal tar in applied shampoo left as a residue. If ICF were to adjust its analysis to base it on absorption, they would have to see how things changed.

A reviewer asked why the J & J data were considered an outlier. The authors indicated that the shampoo was applied to hairless skin on the volar forearm and left on for one minute. One hundred milliliters of water was used to rinse off the skin, there was no wiping or blotting. The surface area in this experiment was much smaller than the scalp (20 cm² vs. 700 cm²) and orders of magnitude more water is used in shampooing in the shower (100 mL vs. a showering rate of 60-80 L/minute).

Another reviewer asked whether there is any information on actual conditions of use. The PDR recommends lathering and letting the shampoo sit for a certain number of minutes. The authors indicated that the Van Schooten study provides a good simulation of showering scenario and they therefore assumed the Van Schooten procedure of shampoo, leave on for 1 minute, and rinse.

A reviewer asked whether the pyridinethione used in the Barker and Winrow (1979) study was a PAH. The authors indicated it was not and that the data from this study were used to represent a shampooing scenario.

Observer Presentation and Clarifying Questions from the Panel

Mr. Perry Gottesfeld from Occupational Knowledge International and a co-litigant with the State of California in the Proposition 65 case asked the panel to consider a number of issues, including the following:

- US Pharmacopia allows for a wide range of distillation temperatures from 900 to 1100 degrees. As shown in the literature there is a wide range of distribution of

- PAHs in these shampoos, something like a 50-fold difference in BaP has been reported in the literature for coal tar shampoos.
- Studies cited concerning consumer use of shampoo shows a bimodal distribution among those who buy these products. There seem to be very many people who buy one bottle and throw it away and a huge number that use it every day and buy large quantities of it. He did not see where the authors have accounted for that.
 - Data on years of usage do not make sense. The maximum number of years usage can only be 20 years since the product has only been marketed for 20 years. Proposition 65 considers a 70 year exposure.
 - The Toppmeyer (1994) and Nielsen (1999) data do not show how much extra strength versus regular shampoo is used. For the American Home Products product the difference is 0.5% versus 2.5% coal tar.
 - Toppmeyer (1994) also shows that females use a lot more than males; former users switch brands and they then get exposure from other brands, and 6% of current users use over 200 ounces per year. The data do seem to suggest that those who shampoo the longest use the most shampoo and those who buy the most, use the most.
 - The ICF study did not account for exposure to those who shampoo in a bathtub.
 - Differences in absorption for things like temperature have not been accounted for (people take hot showers). Nor has the presence of alcohol in these products is key in terms of exposure as is the fact that soaps are surfactants and all this would increase absorption.
 - These products are known to be used to treat cradle cap in babies, these are average users.
 - The amount deposited on hair, the residue on hair has not been addressed.
 - The skin acts as a reservoir for coal tar and build up over time is not addressed.
 - The ICF average of 6.1 ounces per year is not in agreement with a European Union document, which says average users use 102.9 ounces/year. Toppmeyer study indicates that the average user uses 60 ounces per year.

Mr. Gottesfeld also asked the panel whether coal tar could cause cancer in humans, and whether there is a presumption by panel members that coal tar was a probable human carcinogen. The Chair of the panel indicated that the risk assessment made the presumption that coal tar could cause cancer in humans and the panel has presumed this for its deliberations. The panel asked Mr. Gottesfeld to provide his list of issues on an overhead so that they could consider these in their discussion. Mr. Gottesfeld agreed. The panel had no clarifying questions.

Discussion

A reviewer asked for clarification of the purpose of the review and who the NSRL is intended to represent. An author indicated that under California's Proposition 65, the NSRL is calculated for the average user and average usage patterns. The Chair clarified that the purpose of the peer review is to provide a scientific critique of the document in order to answer the question of whether the authors estimated the average user risk appropriately and provide suggestions on the document.

The panel addressed the question of whether the exposure assumptions were reasonable.

Several reviewers questioned the upper bound of 81% for the range of values of the conversion factor representing the amount of the absorbed pyrene that is excreted in the urine. One reviewer noted that the percentage of pyrene absorbed into the systemic circulation, which ends up in the urine and can then be measured as 1-OHP, is a key element to back calculate either residue or absorbed dose. This reviewer mentioned a study by Withey et al. (1980), wherein rats were dosed by intravenous and oral routes and total pyrene in the urine and feces was measured at multiple time points. That study found 45% of pyrene ends up in the urine regardless of the form it is in. An author indicated that data on humans indicate that up to 100% of the pyrene is excreted in the urine, while for rats it is approximately 50% in urine and 50% in feces (biliary excretion is important in rodents, but not in humans). The authors explained that the 13% conversion factor is based upon 1-OHP in the urine of coke oven workers (Van Rooij et al., 1993b). The amount of absorption was not measured, but Van Rooij et al. created a model from absorption of ointment and applied the model to the absorption for coke oven workers. Van Rooij estimated that 13% of the amount of pyrene absorbed was excreted in the urine as 1-OHP. The reviewer also mentioned an oral study by Viau that indicates that very little (1.4%) is excreted in the urine. Several reviewers suggested that the Singh et al. (1995) study may have been miscited in the Withey et al. (1980) report, which is total amount in urine, not total pyrene, and would result in an overestimate. They thought that more appropriate bounds on what is absorbed through the skin and excreted in the urine would be 4 to 20%. The authors indicated they need to revisit this and the panel agreed.

The panel discussed the distribution for ounces of coal tar shampoo per year. Several reviewers thought the mean of 9.88 ounces of coal tar shampoo used was too low. The authors explained how they estimated this distribution from the Nielsen (1999) and Toppmeyer (1994) data. They explained that the Nielsen data were from approximately 44,000 households (of which 8148 purchased coal tar shampoos) who recorded all purchases (barcode swiping) for a period of two years. The Nielsen product specific data were reviewed by the sales and marketing staff of Whitehall Robbins and the estimates of purchases closely matched sales data. The Toppmeyer study reported the results of a random phone survey of 300 persons and information for individual users of one product was estimated. The estimates of purchases from Toppmeyer were 7 times greater than actual sales. The authors explained that they took the data on users from the Toppmeyer study and added these to the appropriate usage category to increase the N for each category. For the Nielsen data they assumed that just one person used the entire purchased product, thereby potentially overestimating usage by not considering the sharing of shampoo within the household or unused shampoo.

Another reviewer questioned whether the Toppmeyer data should be used at all, given the superior quality of the Nielsen data. The authors indicated that including the Toppmeyer data increased the mean from 7.4 to 10 ounces per year. They clarified that the estimates they used were brand specific, as the purpose of their assessment was to assess risk from a single brand. The authors noted that the Nielsen survey included 8148 households who bought coal tar shampoo during the 2 years. Denorex was purchased by 3263 households

(mean annual average use 7.4 ounces) and 4578 households bought Neutrogena's product (mean annual average use of 6.8 ounces). The average total coal tar shampoo purchased was 7.55 ounces/year, which is just slightly higher than the draft document average of 7.4 for Denorex. The panel questioned the appropriateness of only using one product and recommended the authors present total coal tar shampoo usage as well so that the reader has a clear idea of the risk from using coal tar shampoo, rather than just one product. A reviewer raised an uncertainty with the Nielsen study, suggesting it was possible that respondents might not have remembered to bar code swipe all purchases. The panel agreed that the Nielsen data on total coal tar shampoo usage was more reliable than Toppmeyer.

An observer stated that the ICF average was not in agreement with the value of 102.9 ounces per year used by the European Union (EU) in a document entitled "Guidance for Testing of Cosmetic Ingredients for Their Safety Evaluation" by the Scientific Committee on Products and Nonfood Products Intended for Consumers. The chair asked the observer to clarify what specifically the EU value represented, whether it was all coal tar, some coal tar, or something else. The observer indicated that it was a generic number that the EU uses in their risk assessment for all shampoos, but he did not know what it specifically represented. The panel asked the observer to provide the authors with the full EU citation so that they could investigate what this value represents.

The panel recognized that the Nielsen and Toppmeyer data each have their own notable limitations that affect data quality. Reviewers asked the authors to try to confirm the Nielsen data from other data sources, including the EU data. The panel asked the authors to add an analysis of total coal tar shampoo usage, in addition to the individual product estimates, and to reconfirm and document the Nielsen data with product sales data. The panel asked the authors to discuss qualitatively the assumption that one person in the household uses the entire product and the possibility that not all products might have been scanned.

To calculate years of shampoo usage (Y), the authors supplemented the data from Toppmeyer with information from Menter and Cram (1983) and Pittlekow et al. (1981). They used a piece wise uniform distribution to calculate Y. Their assumption was that years of shampoo usage would correspond to years of treatment for a skin condition like psoriasis. Reviewers questioned whether years of suffering with psoriasis is relevant and noted that the Pittlekow and Toppmeyer data had limited follow up. With longer follow up it would be expected that people would be reclassified into greater duration of use categories. ICF agreed that removing Pittlekow data from the calculations for years of shampoo usage is necessary to eliminate short follow up bias. One reviewer noted that this is a good distribution, but questioned whether it needs to be shifted to account for years of product availability. Several panel members recommended that the psoriasis data should not be included in the distribution. To account for potential usage greater than 20 years, the panel suggested the authors look at Toppmeyer to determine age of respondents, which might help determine usage beyond 20 years. The panel also suggested the authors consider 60 years as the upper bound on years of usage.

The panel agreed the LADD equation was correctly used and recommended that absorbed dose is more appropriate than fraction of coal tar in applied shampoo residue (as discussed above).

The Chair raised a question posed by an observer of how the authors handled the bimodal nature of amount of shampoo purchased in the Toppmeyer and Nielsen data. Several reviewers had difficulty understanding the basis of this question. A reviewer thought that some of the factors affecting the shape of the combined distribution for years of usage (e.g., recall bias) also may have contributed to the apparent bimodal nature of amount of shampoo purchased in the Toppmeyer and Nielsen data. Some reviewers noted that the Nielsen coal tar purchasers were assumed to be coal tar users, while Toppmeyer respondents were asked whether and how much they used.

RISK CHARACTERIZATION

The panel discussed children's potential usage of these products, for example use on infants with cradle cap. One reviewer noted that coal tar shampoos are not commonly prescribed for cradle cap today. Several reviewers noted areas in the document that contained information that would allow the authors to address potential children's risk:

- on use of Goeckerman treatment as young as one year from Nagawa et al. (1970),
- data on human neonatal foreskin and adult homogenates,
- data on xenografts and potential age differences,
- data on basal AHH levels for fetal to adult and DNA repair rates from fetus to adult, and
- treatment of psoriasis in children.

Panel members also noted that if one were to quantitatively address a child as a specific target population, parameters in the exposure assessment that could be considered include:

- absorption differences between adults and children. A reviewer agreed to share some data on this (Guzelian et al., 1992 and Yaffe and Aranda, 1992).
- differences in surface area of scalp to adjust ounces of shampoo used per shampooing event.

An author asked whether the distribution of 25-100% for absorption of residue captures all of these concerns. Several reviewers suggested that in the Monte Carlo assessment some of the parameters (e.g., surface area of scalp and body weight or absorption) would need to be correlated, or stratified, so that the child values would be matched with other child values. A third approach would be to calculate these deterministically. The appropriate approach would depend upon what specifically the authors were trying to estimate. Just looking at the distributions alone may not be sufficient to account for these potential differences, although one reviewer noted that given other uncertainties this is a

minor issue. The panel recommended that the authors address these issues in the risk characterization discussion of uncertainties.

The panel agreed that the general approach for calculating the NSRL was correct, but that the estimations need to be redone with absorbed dose rather than remaining residue.

The Chair thought it was a fair statement to make on the panel's behalf that the authors have done a very nice job in putting together an assessment document that addressed many different pieces of information and certainly a lot more than have been seen for many different kinds of assessments. However, there were a number of the authors' uncertainties that the panel enhanced. The panel also added a few additional ways to characterize the uncertainty. The chair asked whether panel members would offer individual opinions on whether the existing document is a) great and should be distributed, or b) good, but needs additional work, or c) not even close to a point where it could be distributed.

The first reviewer noted that the assessment is a nice piece of work, if one had to do the work. Given the caveats that the panel has raised and these caveats notwithstanding, the reviewer thought that they even did an admirable piece of work. The reviewer had some serious questions as to whether or not the assessment should be done, given the data and that this is getting close to the margins of what risk assessment can accomplish. If the reviewer had to assign a confidence level it would certainly be low, but further clarified that what is being asked in this assessment is to determine whether the point estimate falls above or below a bright line and one needs to consider where the confidence intervals fall. This makes it particularly difficult.

Another reviewer noted that while the panel has not been asked to compare this work to other assessments, one could not help but compare and note the large amount of data here. Not only are there epidemiology studies, given certain limitations, but also there are epidemiological data upon which one can base a risk assessment. Although the panel has not yet fully discussed the animal study data, when one incorporates pharmacokinetic differences and includes the initiation and the mechanistic information into a mix of uncertainty, it qualitatively supports what the epidemiology data are demonstrating. Therefore, this reviewer would state that this document goes a lot further than many other documents that are making statements about potential for risk.

The next reviewer wanted to add to the previous statement that although there are uncertainties about the quantification, and although one might have misgivings about a lot of the issues the panel has discussed in the meeting, one should not overlook the total, overall database. This includes many negative epidemiology studies that are not of sufficient quality to use for quantification, but that nonetheless add to our overall understanding. Therefore, this reviewer was not uncomfortable going forward with doing quantification.

The next reviewer also agreed and noted that although this reviewer personally does not do quantitative risk assessment, the reviewer liked the weight of evidence approach taken

in the document, using all the available data. The panel has had some questions and has at least addressed one of the exposure portions of the document. Given that the authors had to do the assessment, it was a good piece of work, they have made a really good effort at it, and it is really very good.

Another reviewer agreed that it is an admirable piece of work, it is voluminous and thorough. While the tendency has been to discount the animal data, this reviewer thinks that it needs to be retained and put in perspective. This is because one can arrive almost at the same point using both the epidemiological data and the animal data. It is supportive despite the fact that there are some concerns about the pharmacodynamic considerations. Nonetheless, it is useful to include given the enormous amount of mouse skin painting data available. It would be unwise to discount it completely. That said, the reviewer agreed with the colleague who thought it seemed to be on the edge of what one can do with risk assessment, but that the authors are working with the data available. Without a large amount of additional data from actual experimentation, one could add supportive data, perhaps exposure data to help trim up some of the numbers used in the LADD equations.

Another reviewer agreed with all of the previous positive comments and asked whether this was a required assessment. An author indicated that it is a voluntary analysis that is required.

Another reviewer was somewhere in the middle. The reviewer noted that while not a risk assessor familiar with all the permutations of the analysis, attaching a level of confidence to an estimate of risk based upon the one epidemiological study seemed tenuous. This reviewer somewhat agreed that the animal data suggest that humans appear to be less sensitive than the mouse and that there are various physiological reasons for that, and would have to agree that probably the risk is relatively low. While this is not in a league with asbestos, given the uncertainties of both the animal and the human data, the reviewer was not sure that there is no excess risk at the level presented.

Another reviewer noted that while the document represented a lot of work, more justification and documentation were needed. While this reviewer was also uncomfortable having to go forward with such limited information, this is not outside the range of general practice. Going forward with quantification in these circumstances should be done with careful clarification in the document itself and with a better description of the uncertainties in the risk characterization. There were a number of issues discussed, but there are a number of others to be added which would strengthen that particular piece.

Another reviewer indicated that this type of risk assessment was a new experience. The reviewer was glad to see that at least there were some human data. Both the animal and human data were imperfect but the imperfections were different and it needs to be recognized that coming out with the same answer or similar answers even though the imperfections were in different places, gives one comfort that maybe you are close to being right. The reviewer thought that this was a good document that will be made much

better from this process, with a better understanding of what that number really means and how much uncertainty surrounds it.

Another reviewer indicated that there is enough rigor and thoroughness in this analysis to provide personal confidence in recommending these shampoos to people who need to use them; that they could use them without fear of adverse consequence. The reviewer thought this document is at the limits of risk assessment. If there is a risk, it is so low that one has got to do layers and layers of Monte Carlo to try to quantitate it. But the direction that the qualitative weight of evidence is going is very favorable and as a practitioner this reviewer would be confident in using this.

Another reviewer was very impressed with the thoroughness of the review of the animal data. The reviewer's personal experience has been that, in general, epidemiology data can be inconclusive, but in this case came to the conclusion that the human studies are sufficient. However, the authors need to do some additional work and additional uncertainty analysis with the Monte Carlo analyses.

The Chair noted again that the panel *presumed* that coal tar is carcinogenic, which is consistent with the underlying assumption that the authors have made. There may be data that indicate otherwise, but this panel is not operating outside of the presumption that coal tar is carcinogenic, but perhaps with a very low potency. A reviewer expressed concern regarding the statement that the panel presumed that coal tar is carcinogenic. While this reviewer is certain that PAHs are carcinogenic, the reviewer is not sure about the presumption that coal tar is carcinogenic just because it contains PAHs. Another reviewer noted that coal tar is on the California list of carcinogens, and therefore the reviewer is agreeable with assuming it is carcinogenic for this exercise. But, this reviewer does not believe coal tar is carcinogenic and does not want to go on record as having it implied such. Therefore, presumption is an important key word.

Another reviewer brought the panel's attention to the last sentence of the report that said "use of these products should be considered safe and these products should not require a warning label under Proposition 65." This reviewer thought that this comes close to making a risk management decision and would want to see more discussion about Proposition 65 before agreeing with this statement. Other reviewers asked whether the last statement was needed in the document and expressed discomfort with making this judgement regarding safety. An author explained that information on Proposition 65 is necessarily included to provide context for the approach taken, but agreed that they could stop at the point of comparison of the no significant risk levels to the intake, characterize the uncertainties as best they can, and for those uncertainties that they can quantify, show the impact on that ratio. The panel recommended that conclusions regarding safety under Proposition 65 be removed from the document.

REFERENCES

Alderson MR, Clarke JA. 1983. Cancer incidence in patients with psoriasis. *Br J Cancer* 47:857-859.

Barker EA, Winrow MJ. 1979. Percutaneous (Scalp) Absorption of Zinc Pyridinethione in Humans. Report to Procter & Gamble (As cited in Food & Drug Administration OTC Volume 160369).

Bhate SM, Sharpe GR, Marks JM, Shuster S, Ross WM. 1993. Prevalence of skin and other cancers in patients with psoriasis. *Clin Exp Dermatol* 18:401-404.

Fraunhofer Institute of Toxicology. 1997. Dermal Carcinogenicity Study of Two Coal Tar Products (CTP) by Chronic Epicutaneous Application in Male CD-1 Mice (78 Weeks). Final Report. Fraunhofer Institute of Toxicology, Federal Republic of Germany. October 1997.

Grupper C, Berretti B. 1980. Tar, ultraviolet light, PUVA, and cancer. *J Am Acad Dermatol* 3:643-646.

Guzelian PS, Henry CJ, Olin SS, eds. 1992. Similarities and Differences Between Children and Adults: Implications for Risk Assessment. ILSI Press, Washington, DC.

Hansen ES. 1989. Cancer incidence in an occupational cohort exposed to bitumen fumes. *Scand J Work Environ Health* 15:101-105.

Hendricks NV, Berry CM, Lione JG, Thorpe JJ. 1959. Cancer of the scrotum in wax pressmen. *AMA Arch Ind Health* 19:524-529.

Jarvholm B, Easton D. 1990. Models for skin tumour risks in workers exposed to mineral oils. *Br J Cancer* 62:1039-1041.

Jarvholm B, Fast K, Lavenius B, Tomsic P. 1985. Exposure to cutting oils and its relation to skin tumors and premalignant skin lesions on the hands and forearms. *Scand J Work Environ Health* 11:365-369.

Jemec GBE, Østerlind A. 1994. Cancer in patients treated with coal tar: a long-term follow up study. *J Eur Acad Dermatol Venereol* 3:153-156.

Jones SK, Mackie RM, Hole DJ, Gillis CR. 1985. Further evidence of the safety of tar in the management of psoriasis. *Br J Dermatol* 113:97-101.

Jongeneelen F, Bos R. 1990. Excretion of pyrene and hydroxypyrene in urine. Letters to the editor. *Cancer Lett* 51:175-179.

Letzel, S and H Drexler. 1998. Occupationally related tumors in tar refinery workers. *J of the Am Acad Derm* 39(5):712-720.

- Lindelof B, Sigurgeirsson B. 1993. PUVA and cancer: a case-control study. *Br J Dermatol* 129:39-41.
- Maier, H., M. Schenmper, B. Ortel, M. Binder, A. Tanew, and H. Honigsmann. 1996. Akin tumors in photochemotherapy for psoriasis: a single-center follow-up of 496 patients. *Dermatology* 193:185-191.
- Maizlish N, Beaumont J, Singleton J. 1988. Mortality among California highway workers. *Am J Ind Med* 13:363-379.
- Maughan W, Muller S, Perry H, Pittelkow M, O'Brien P. 1980. Incidence of skin cancers in patients with atopic dermatitis treated with coal tar. *J Am Acad Dermatol* 3:612-615.
- Menter A, Cram DL. 1983. The Goeckerman regimen in two psoriasis day care centers. *J Am Acad Dermatol* 9:59-65.
- Muller SA, Perry HO. 1984. The Goeckerman treatment in psoriasis: six decades of experience at the Mayo Clinic. *Cutis* 34:265-269.
- Nagawa et al. 1970.
- Nesnow, S. and Lewtas, J. (1991) Interspecies sensitivity to chemical carcinogens: Relationship between mouse skin tumors and human lung cancer, *Environ. Carcinoge. Ecotox Revs.*, c9(2), 229-259.
- Nielsen. 1999. Nielsen Household Panel Distribution Report. Consumer stated usage of Denorex shampoo.
- Pittelkow M, Perry H, Muller S, Maughan W, O'Brien P. 1981. Skin cancer in patients with psoriasis treated with coal tar. *Arch Dermatol* 117:465-468.
- Scotto J, Kopf AW, Urbach F. 1974. Non-melanoma skin cancer among Caucasians in four areas of the United States. *Cancer* 34:1333-1338.
- Sherson D, Svane O, Lynge E. 1991. Cancer incidence among foundry workers in Denmark. *Arch Environ Health* 46:75-81.
- Singh R, Tucek M, Maxa K, Tenglerova J, Weyand EH. 1995. A rapid and simple method for the analysis of 1-hydroxypyrene glucuronide: a potential biomarker for polycyclic aromatic hydrocarbon exposure. *Carcinogenesis* 16:2909-2915.
- Soballe PW, Montone KT, Satyamoorthy K, Nesbit M, Harlyn M. 1996. Carcinogenesis in human skin grafted to SCID mice. *Cancer Res* 56:757-764.
- Spinelli JJ, Band PR, Svirchev LM, Gallagher RP. 1991. Mortality and cancer incidence in aluminum reduction plant workers. *J Occup Med* 33:1150-1155.

Stern RS, EJ Liebman, L Vakeva. 1998. Oral psoralen and ultraviolet-A light (PUVA) treatment of psoriasis and persistent risk of nonmelanoma skin cancer. *J of the Nat Cancer Inst* 90(17):1278-1283.

Stern RS, Laird N. 1994. The carcinogenic risk of treatments for severe psoriasis. *Cancer* 73:2759-2764.

Stern RRS, Laird N, Melski J, Parrish JA, Fitzpatrick TB, Bleich HL. 1984. Cutaneous squamous-cell carcinoma in patients treated with PUVA. *N Engl J Med* 310:1156-1161.

Storm J, Collier S, Stewart R, Bronaugh R. 1990. Metabolism of xenobiotics during percutaneous penetration: Role of absorption rate and cutaneous enzyme activity. *Fundam Appl Toxicol* 15:132-141.

Toppmeyer Research, Inc. 1994. Denorex Usage Study. Job #02-01-94. Toppmeyer Research, Inc., Great Neck, NY.

U.S, EPA. 1992. A cross-species scaling factor for carcinogen risk assessment based on equivalence of $\text{mg/kg}^{3/4}/\text{day}$; notice. Draft report. *Fed Regist* 57(109):24152-24173.

VanRooij JGM, DeRoos JHC, Bodelier-Bade MM, Jongeneelen FJ. 1993a. Absorption of polycyclic aromatic hydrocarbons through human skin: differences between anatomical sites and individuals. *J Toxicol Environ Health* 38:355-368.

Van Rooij J, Bodelier-Bade M, Jongeneelen F. 1993b. Estimation of individual dermal and respiratory uptake of polycyclic aromatic hydrocarbons in 12 coke oven workers. *Br J Ind Med* 50:632-632.

Van Schooten F-J, Moonen EJC, Rheinsburger E, van Agen B, Thijssen HHW, Kleinjans JCS. 1994. Dermal uptake of polycyclic aromatic hydrocarbons after hairwash with coal-tar. *Lancet* 344:1505-1506.

Withey et al. 1980.

Yaffe SJ, Aranda JV. 1992. *Pediatric Pharmacology. Therapeutic Principles in Practice*. 2nd ed. W.B. Saunders, New York.

Managing Potential Conflicts of Interest

***ITER* Peer Review Meeting**

June 5 and 6, 2000

(approved by panel on June 5, 2000)

ITER peer reviewers donate their time and talents to this effort. They are selected based upon their expertise and qualifications and are employed by many types of organizations. *TERA* strives to create a balance of expertise and affiliations for each meeting. However, individual peer reviewers are representing their own expertise and views, not those of their employer. The *TERA* Board of Trustees approves *ITER* peer reviewers for inclusion in this program. A complete list of potential reviewers and more information on the *ITER* peer review program are available at <http://www/tera/org/peer>. Additional, *ad hoc* reviewers may be selected to participate for their special expertise that may be needed for a particular chemical or discussion.

TERA requested that each peer reviewer identify potential conflicts of interest related to the review of the document "Estimation of Lifetime Skin Cancer Risk from the Use of Coal Tar Containing Shampoos" and/or American Home Products or Orrick, Herrington & Sutcliffe LLP, or Neutrogena, the sponsor of these discussions. Each reviewer has signed a statement disclosing any potential conflicts of interest concerning this assessment review. The following statements were considered by the panel and have been revised and agreed upon by the panel.

Brian B. Adams. Dr. Adams is an Assistant Professor of Dermatology at the University of Cincinnati and Director of Dermatology at the Veterans Administration Hospital, Cincinnati. He has been asked to participate as an *ad hoc* reviewer for his dermatology expertise. He does not have any conflicts and will participate fully in all discussions and polling for consensus.

Michael L. Dourson. Dr. Dourson is the Director of Toxicology Excellence for Risk Assessment and will serve as the Chair for this panel. He does not have any conflicts and will participate fully in all discussions and polling for consensus.

Brent L. Finley. Dr. Finley is a toxicologist with Exponent. Dr. Findlay has performed worked for a Johnson & Johnson company (not Neutrogena) on issues related to latex gloves. The panel agreed that Dr. Finley should participate fully in the discussion and polling for consensus. However, all panel members were asked to look for undue influence on the discussions, and any panel member was given the ability to approach the chair to state an issue of undue influence. The chair would then decide whether additional panel discussion would be needed.

John Gamble. Dr. Gamble is a Senior Staff Epidemiologist with Exxon Mobil Biomedical Sciences. He has been asked to participate as an *ad hoc* reviewer for his epidemiology expertise He does not have any conflicts and will participate fully in all discussions and polling for consensus.

Loren Lund. Dr. Lund is a Senior Toxicologist with Parsons Engineering Science Inc. He has been asked to participate as an *ad hoc* reviewer for his dermal risk assessment expertise. He does not have any conflicts and will participate fully in all discussions and polling for consensus.

Brian Magee. Dr. Magee is Vice President and Principal Toxicologist with Ogden Environmental and Energy Services. He has been asked to participate as an *ad hoc* reviewer for his exposure assessment and coal tar expertise. Dr. Magee noted that he worked for ICF for six months approximately 15 years ago. The panel did not think this a real or perceived conflict and agreed that he will participate fully in all discussions and polling for consensus.

Rebecca T. Parkin. Dr. Parkin is an epidemiologist and Associate Professor in the Department of Environmental and Occupational Health of the George Washington University Medical Center. She does not have any conflicts and will participate fully in all discussions and polling for consensus.

Timothy A. Roy. Dr. Roy is a toxicologist and chemist, and President of Petrotec Inc. Until earlier this year he was with Mobil Oil Corporation. He has been asked to participate as an *ad hoc* reviewer for his dermal expertise with complex mixtures including coal derived products. Dr. Roy has carried out several studies for one of the Johnson and Johnson companies. He is conducting bioavailability research on topical antifungals. Dr. Roy has informally discussed the coal tar issue with Neutrogena personnel, but has not consulted or carried out studies for Neutrogena. The panel decided that Dr. Roy should participate fully in the discussion and polling for consensus. However, all panel members were asked to look for undue influence on the discussions, and any panel member was given the ability to approach the chair to state an issue of undue influence. The chair would then decide whether additional panel discussion would be needed.

Ted W. Simon. Dr. Simon is a toxicologist with the U.S. Environmental Protection Agency, Region IV. He was not able to attend the meeting, but has provided written comments for the panel's consideration. Dr. Simon has no conflicts and his comments will be fully considered.

Alan H. Stern—Dr. Stern is a risk assessor and toxicologist, and is the Chief of the Bureau of Risk Analysis for the New Jersey Department of Environmental Protection (NJ DEP). He does not have any conflicts and will participate fully in all discussions and polling for consensus.

Allan S. Susten. Dr. Susten is Assistant Director for Science, Division of Health Assessment and Consultation, Agency for Toxic Substances and Disease Registry. Dr. Susten's wife inherited shares of American Home Products in late 1980s. They presently own 28 shares valued at \$1575 (4/30/2000), which comprises about 1.1% of their stock assets. Estimated yearly income is about \$25 and they have never traded this stock. The panel did not consider this a conflict and he will participate fully in all discussions and polling for consensus.

Glenn Talaska. Dr. Talaska is a toxicologist, industrial hygienist, and an Associate Professor at the University of Cincinnati, Department of Environmental Health. He has been asked to participate as an *ad hoc* reviewer for his expertise in PAH toxicology and

biomarkers. He does not have any conflicts and will participate fully in all discussions and polling for consensus.

Rebecca Tominack. Dr. Tominack is the Residency Director, Occupational and Environmental Medicine, St. Louis University Health Sciences Center, and Assistant Medical Director of the Missouri Regional Poison Center. She is board certified in internal medicine and medical toxicology. She does not have any conflicts and will participate fully in all discussions and polling for consensus.