

BILL LOCKYER
Attorney General

State of California
DEPARTMENT OF JUSTICE



1515 CLAY STREET, 20TH FLOOR
OAKLAND, CA 94612-1413

4555 00 NOV 17 AIO :16

Public: (510) 622-2100
Telephone: (510) 622-2142
Facsimile: (510) 622-2270
E-Mail: @hdcdojnet.state.ca.us

November 16, 2000

By Overnight
Charles J. Ganley, M.D.
Food and Drug Administration
Room 1061
5630 Fishers Lane
Rockville, MD 20852

RE: FDA Docket 00P-1210/CP1: Gottesfeld Petition for formal review of the conditions of sale, use and distribution of FDA-regulated products containing Coal Tar USP

Dear Dr. Ganley:

I am writing concerning the above petition. As you know, the Attorney General of the State of California has filed suit in the California Superior Court under the Safe Drinking Water and Toxic Enforcement Act of 1986, known as Proposition 65. (California Health & Safety Code sections 25249.5 *et seq.*) That action has been stayed at the request of the defendants until January 31, 2001 in order to determine what action the FDA will take on the above petition.

As we previously explained in our letter of August 23, 2000 to Jennifer Butler, (copy attached) we believe that the Gottesfeld Petition has no direct relationship to or bearing on the pending Proposition 65 action. Proposition 65 simply provides that businesses who expose individuals to listed carcinogens or reproductive toxins, must provide a warning, unless the defendant can show that the exposure is below the defined "no significant risk level" ("NSRL"). Coal Tar is listed in 22 California Code of Regulations ("CCR") section 12000 as a carcinogen under "Soots, tars, and mineral oils." Proposition 65 also contains its own very specific set of regulations that set forth the principles for performing a risk assessment to determine if the exposure is below the defined NSRL. (22 CCR § 12703).

During the course of the Proposition 65 litigation, the defendants prepared a series of risk assessments, one of which, the ICF Kaiser Report, has been submitted to the FDA for review. The Attorney General believes that ICF Kaiser risk assessment is seriously flawed scientifically and calculates an NSRL that is significantly above the actual NSRL calculated according to Proposition 65 principles. Our expert, Dr. Lorenz Rhomberg, from Gradient Corporation, has

00P-1210

SUP 1

Dr. Charles J. Ganley
November 16, 2000
Page 2

prepared his own report, calculating the Proposition 65 NSRL for coal tar products to be in the range of 1.5 micrograms per day, rather than the 29 micrograms calculated by the ICF Kaiser Report. Dr. Rhomberg's report relies on the same Fraunhofer study that was used on by the European scientific community to ban outright over-the-counter sales of coal tar shampoos and skin-care products.

Dr. Rhomberg and an additional expert, Dr. Sanford Greenland, who is a biostatistician, have also raised certain significant and important criticisms of the ICF Report. Without going into these in detail, they include the fact that the Report includes a remarkably high degree of uncertainty¹ and has problems in its statistical analysis; the Report relies on usage assumptions that are unrealistically low and based on data that was intended to measure purchase and not use of shampoo; and the Report makes certain critical assumptions about the Pittelkow study that are not supported in the underlying data.

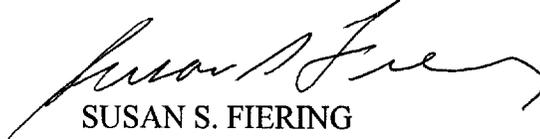
While we understand that the FDA will not be addressing the same issues that are presented in the pending Proposition 65 action, we want to make sure that you have before you some of the additional information that we have developed in the course of our litigation, and have therefore enclosed a copy of the report prepared by Dr. Rhomberg. Because expert discovery has not been completed, we are unable to provide you with all of our critiques of the ICF Kaiser Report. This information will, however, be disclosed in detail once the stay is lifted and expert discovery recommences, and at the time of trial of this matter. We are submitting information to FDA for its review, solely because we believe that the FDA may find this information helpful in its assessment of its own issues and because we believe that the FDA should be aware of the available information that has already been presented and that will be presented in the state court proceeding. Nothing in our submission is intended to suggest that the FDA's decision will have any bearing on the outcome of the Proposition 65 action or that the Proposition 65 action should be stayed pending the outcome of the FDA's decision on the Petition.

¹The ICF Report itself acknowledges that if there were as few as two additional unrecognized cancers in the Pittelkow cohort, the NSRL would be 10 ug per day, rather than the reported 29. The Report does not go on to state what the NSRL level would be if the additional cancers were 3 or 4.

Dr. Charles J. Ganley
November 16, 2000
Page 3

Please call me if you have any questions.

Sincerely,

A handwritten signature in black ink, appearing to read "Susan S. Fiering", written in a cursive style.

SUSAN S. FIERING
Deputy Attorney General

For BILL LOCKYER
Attorney General

BILL LOCKYER
Attorney General

State of California
DEPARTMENT OF JUSTICE



4556 00 NOV 17 A10:16

1515 CLAY STREET, 20TH FLOOR
OAKLAND, CA 94612-1413

Public: (510) 622-2100
Telephone: (510) 622-2142
Facsimile: (510) 622-2270
E-Mail: @hdcdojnet.state.ca.us

August 23, 2000

By Overnight
Jennifer Butler
Docket Management Branch
Food and Drug Administration
Room 1061
5630 Fishers Lane
Rockville, MD 20852

RE: FDA Docket 00P-1210/CP1: Comments concerning Gottesfeld Petition for formal review of the conditions of sale, use and distribution of FDA-regulated products containing Coal Tar USP

Dear Ms. Butler:

I am writing on behalf of the Attorney General for the State of California in response to the August 8, 2000 letter of Carol Brophy, counsel for Bergen Brunswig Drug Company, commenting on the petition of Perry Gottesfeld, Docket No. 00P-1210/CP1.¹

Ms. Brophy's letter discusses in great length a suit filed by the Attorney General in state court on behalf of the People of the State of California under Health & Safety Code section 25249.5 *et seq.* ("Proposition 65") against manufacturers of coal tar shampoos and skin care products. The Attorney General's lawsuit alleges that the manufacturers have exposed individuals to carcinogens listed under Proposition 65, including tars and various polyaromatic hydrocarbons (benzo(a)pyrene, benzo(a)anthracene, etc.) without providing a warning. Mr. Gottesfeld, the petitioner in Docket No. 00P-1210/CP1, filed a separate Proposition 65 lawsuit in state court against the manufacturers of the coal tar products in state court in his capacity as a private citizen; Mr. Gottesfeld's case was consolidated with the case filed by the Attorney General. The consolidated case is currently set to go to trial on October 10, 2000.

¹This letter does not attempt to respond to all legal and factual inaccuracies contained in Ms. Brophy's letter.

Contrary to Ms. Brophy's assertions, the Proposition 65 lawsuit now pending in state court has no direct relationship to or bearing on Gottesfeld petition now pending before the Food and Drug Administration ("FDA"). Ms. Brophy states at the bottom of page 2 and top of page 3 that the state court will decide whether pharmaceuticals containing Coal Tar USP are safe when used in concentrations authorized under the Coal Tar Monograph, or if limitations to the sale, use and distribution of such products are warranted. This is incorrect. Proposition 65 does not address the safety of the product, but simply mandates that businesses that expose individuals to listed carcinogens must provide a clear and reasonable warning. Thus, nothing decided by the state court will determine whether the product is safe and effective, a matter to be determined by the FDA. Furthermore, nothing in the state lawsuit will control the sale, use and distribution of the products. The products can remain on the market, if the manufacturers so choose, subject to the warning requirement.

Ms. Brophy further asks the FDA to adjudicate certain issues that are well beyond the scope of the Gottesfeld petition and beyond the regulatory authority of the FDA. For example, on page 5 of her letter, Ms. Brophy asks the FDA to determine whether coal tar USP is encompassed under the listing of soots, tars and mineral oils contained in Proposition 65. Ms. Brophy also asks the FDA to determine the average consumers' use of shampoos, soaps, and ointments, a determination to be made under Proposition 65 regulations. We respectfully request that the FDA reject Ms. Brophy's request and decline to rule on issues of state law.

In addition, Ms. Brophy asks the FDA to determine if the provision of the Proposition 65 warning on Coal Tar products would constitute "misbranding." In effect, Ms. Brophy requests a ruling on preemption of Proposition 65 by the FDA in the context of the Gottesfeld petition -- notwithstanding the fact that 21 U.S.C. section 379r(d)(2), a provision of the Food and Drug Administration Modernization Act, specifically states that laws such as Proposition 65 are not preempted.² While Ms. Brophy is free to pursue her preemption argument in the appropriate forum, *e.g.*, state court,³ we request that the FDA refrain from taking any position in conflict with Section 379r(d)(2) that would suggest that Proposition 65 warnings are preempted.

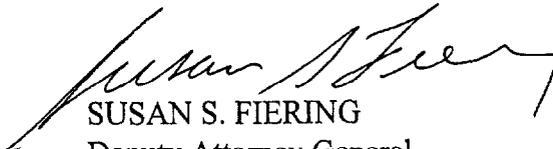
²It is also our understanding that one of the manufacturers of the coal tar products has already approached FDA on this issue and has been told that the FDA does not intend to take a position that the Proposition 65 warning would constitute misbranding.

³A motion for summary adjudication based on preemption is now pending in the state court action.

Jennifer Butler
August 23, 2000
Page 3

If you have any questions concerning the above, or require additional information, please feel free to contact me.

Sincerely,



SUSAN S. FIERING
Deputy Attorney General

For BILL LOCKYER
Attorney General

4557 00 NOV 17 09:16

**A Proposed No-Significant-Risk Level
for Dermal Exposure to
Coal Tar from Shampoo**

Prepared by:
Gradient Corporation
238 Main Street
Cambridge, MA 02142

Prepared for:
Office of the Attorney General
State of California Department of Justice
1515 Clay St., 20th Fl.
Oakland, CA 94612-1413

1 Introduction

Coal tar appears on the list of substances identified by the State of California as carcinogens under the Safe Drinking Water and Toxic Enforcement Act of 1986 (widely known as Proposition 65). The State of California has not established a No-Significant-Risk Level (NSRL) for dermal exposure to coal tar. The need to consider such an NSRL has arisen in the context of questions about whether coal tar-containing shampoos, used in treating scalp skin conditions, constitute a cancer hazard under Proposition 65. Gradient has been asked by the State of California Attorney General's Office to conduct an independent analysis of coal tar dermal carcinogenicity and to propose an exposure level that would constitute an NSRL for purposes of Proposition 65. The present report describes Gradient's analysis and proposal for an NSRL based on extrapolation to humans of experimental observation of skin tumors in mice following chronic exposure to coal tar painted on their skins.

1.1 Coal Tar

Coal tars are complex mixtures of hydrocarbons, phenols, and heterocyclic compounds that are produced by the destructive distillation of coal. Coal tars contain hundreds of identified compounds, and thousands of compounds in total. Notably, coal tars contain polycyclic aromatic hydrocarbons (PAHs), several of which are considered to be carcinogens (IARC 1985). Although composition of different coal tars can vary, the profile of compounds in the mixture is largely a function of distillation temperature, and tars created at similar temperatures tend to have similar composition.

Coal tar and coal tar-containing products have been used as treatments for skin conditions such as psoriasis for many years, and several over-the-counter products such as ointments and shampoos contain coal tar as an active ingredient. In particular, certain coal tar shampoos, used as dandruff treatments, have coal tar at levels that range from approximately 0.5% to 5% coal tar. These uses have FDA approval, and the coal tar used is pharmaceutical grade coal tar which is distilled from bituminous coal at the relatively narrow temperature range of 900-1100 degrees.

Coal tar is, however, established as an agent causing skin cancer after high and continued dermal exposure (IARC 1985). This is true both in humans, chiefly in epidemiological studies of formerly high occupational exposures, and in experimental animal studies, where extensive skin painting studies have been conducted in several species over a span of decades. Indeed, coal tar and the similar mixtures of polycyclic compounds produced by incomplete combustion were the first recognized chemical carcinogenic agents. Chimney soot, which shares many components with coal tar, was the first identified occupational carcinogen based on elevated incidence of scrotal cancer in English chimney sweeps reported by Percival Pott in 1775. Coal tar's ability to cause tumors in the skin of animals after dermal application was discovered in 1915 by Yamagiwa and Ichikawa. Since this constituted the first experimental model of carcinogenesis, the effect was extensively studied from the 1920s to the 1950s, leading to the fractionation of coal tar, the identification of the PAHs as the principal sources of carcinogenic potency, and in particular, identifying benzo(a)pyrene (BaP) as a particularly potent component. This system also formed the basis for recognition that many carcinogenic agents need metabolic activation to exert their effects and that interaction of compounds or their metabolites with DNA is an important mode of carcinogenic action.

Coal tars are recognized by IARC (1985) as having sufficient evidence of carcinogenicity in both animals and humans. Coal tar appears on the Proposition 65 list of compounds known to the State of California as carcinogenic.

1.2 The Question of Cancer Risks from Low-Level Dermal Exposure

The cancer risks of coal tar are established at high, repeated exposure levels. Exposures to coal tar through pharmaceutical uses result in lower levels of exposure than levels known to be associated with cancer. The levels of exposure from use of coal tar-containing products is well below those clearly associated with skin carcinogenesis because:

1. lower concentrations are applied to the skin;
2. products are often washed off, taking much (but not all) of the coal tar with it, leaving only a relatively small residue; and
3. the use of products is generally of limited duration.

Nonetheless, the question arises, does dermal exposure to low levels of coal tar *via* use of coal tar shampoo lead to concern for potential cancer risk? In particular, do coal tar shampoos require labeling

under Proposition 65 owing to estimated lifetime cancer risks in typical users exceeding a level of one-in-100,000?

The present report focuses on Gradient's analysis of skin cancer as a result of dermal exposure to coal tar *via* shampoo, based primarily on animal experimental evidence regarding dermal application. For perspective, two points not extensively treated in this report should be kept in mind. First, there are further animal studies that indicate that coal tar can cause internal cancers (forestomach and lung) in mice following oral administration (Gaylor *et al.* 2000). Moreover, dermal exposure can lead to systemic absorption and formation of DNA adducts in internal organs, particularly in lung (Ma *et al.* 2000), that are similar to those in skin thought to be involved in the process of dermal carcinogenesis. Thus, the possibility exists that dermal exposure to coal tar might lead to some risk of internal cancers in addition to any skin cancer risk that may be engendered. The present analysis does not examine the risk of internal cancers that might result from coal tar shampoo.

Second, because of the uncertainties in animal-to-human extrapolation, and because coal tar-containing consumer products and medical treatments designed for use on skin have a long history of use, it is of interest to examine the epidemiological literature for the human evidence regarding the potential carcinogenic effects of chronic low-level dermal coal tar exposure. Although there are several published studies on skin cancers among psoriasis patients treated with coal tar ointments, as with the animal data, there is a paucity of studies reporting sufficient information for a quantitative analysis. A study by Pittelkow *et al.* (1981) is the one most often considered amenable to full analysis, and this study has been adduced as a potential basis for quantitative risk estimation by other investigators (ICF 2000). Even this study has several issues that make interpretation difficult, however. Among them are the likely under-reporting of skin cancers, the lack of full lifetime follow-up on many subjects, questions about appropriate referent populations and the variation in background skin cancer rates among such populations, the relevance of coal tar in ointments to shampoo exposures, difficulties in estimating the amount of tar absorbed from ointments, and sketchy information on amounts and durations of use. In view of these issues, it is not at all clear that use of human data provides a better basis for assessing potential coal tar cancer risks than the analysis of well controlled and carefully reported animal experiments. The human data are not reviewed or analyzed in the present report.

2 Identity of Coal Tar in Products

As noted, coal tars can vary depending on the conditions under which they were created, but coal tars produced by similar distillation temperatures have similar compositions (Wright *et al.* 1985, IARC 1985). There is evidence that different compositions can have toxic effects that show quantitative variation and that some coal tar components can enhance or partially inhibit the effects that individual constituents would have if experienced alone (Warshawsky *et al.* 1993, Goldstein *et al.* 1998). Coal tar in over-the-counter medicines, including coal tar shampoos, is pharmaceutical grade coal tar, which is distilled in the range of 900-1100 degrees. Wright *et al.* (1985) showed that different coal tars, including a pharmaceutical stock solution in alcohol and a coal tar shampoo, had similar profiles of constituents and had similar toxic effects, including *in vitro* mutagenicity and potency in initiation/promotion studies in mouse skin painting *in vivo*.

In view of these facts, it is appropriate to examine coal tar as a mixture, rather than as a complex of individually studied components (some of which are named in their own right on the Proposition 65 list), and the *prima facie* case is that the coal tars of concern in consumer products are similar in composition and toxicological properties to one another and to high-temperature industrial coal tar.

3 Fraunhofer Study

The primary basis for the present analysis is observation of dermally applied coal tar causing skin cancers in mice in a study conducted by the Fraunhofer Institute (Fraunhofer 1997). Although many studies over the years have examined skin tumors in animals as a result of skin painting of coal tars and similar products, the Fraunhofer study is the only available animal study of dermal exposure to coal tar that has multiple exposure levels, appropriate control groups, and well defined doses.

In this study, male CD-1 mice were subject to twice-weekly applications of coal tar solution in toluene painted on their shaved backs for 78 weeks (550 days), at which time survivors were sacrificed. Groups of 62 animals/dose were exposed to coal tar in the amount of 0, 0.1, 0.3, 1.0, 3.0, or 9.0 mg of tar per application. The highest dose group (9 mg) was discontinued after 274 days owing to persistent ulceration of skin and focal loss of epidermis, and the doses of 0.1 and 0.3 (and a second set of concurrent

controls) were started at this point. (One mouse in the high dose group was mis-sexed, and this animal was eliminated from analysis.) No effect of treatment on body weight was seen, but survival was decreased in the two highest dose groups.

A second, parallel, study examined a second coal tar (CTP-1)—one chosen to be very low in benzo(a)pyrene content. For equivalent doses in mg of tar, responses were lower from this second coal tar type. Although a few squamous cell tumors were generated (and none in controls), results did not achieve statistical significance. The data on CTP-1 were not analyzed.

The material designated "CTP-2", tested in the Fraunhofer study, was a mixture of tars from several sources and was chosen to have a "medium" benzo(a)pyrene content. Compared to most industrial coal tars and to pharmaceutical grade coal tar, the CTP-2 material has a somewhat lower than typical concentration of some of the larger multi-ring PAHs, although the differences are ones of proportion and not of kind of constituents. Since these larger multi-ring PAHs are the components of most concern for carcinogenicity based on fractionation and testing of individual compounds (Nisbet and LaGoy 1992), the carcinogenic potency of CTP-2 as a material may be somewhat less than for the coal tars of concern in coal tar shampoo. To the extent that this factor operates, use of the Fraunhofer study to estimate coal tar carcinogenic potency would tend to underestimate the risks from coal tars with higher concentrations of BaP and larger PAHs.

3.1 Responses and Doses

Squamous cell papillomas and carcinomas were the type of skin tumor showing an increase in response with dose. The incidences of these were combined. Time of death or sacrifice was noted, and this time was used in time-to-tumor analysis, as described further below. (One hemangioma is included in the data of the highest dose group; it could not be differentiated from squamous cell tumors in the available records.)

Responses were as follows: zero tumors in any of the 62 mice in either of the two solvent control groups, one tumor at 0.1 mg, three at 0.3 mg, 9 at 1.0 mg, 23 at 3.0 mg, and 20 at 9.0 mg. The decrease at 9.0 mg reflects the early termination of this dose group.

Doses were first described as average daily mg applied (*i.e.*, they were multiplied by 2/7 to account for the 2 days per week exposed). These amounts of tar were then expressed as an amount per g of the mouse's whole skin by dividing the amount by 6.612 g, the estimated weight of a 40 g mouse's skin based on the Brown *et al.* (1997) estimate that a mouse's skin constitutes 16.53% of its body mass.

It was assumed that 60% of the applied amount of PAH is absorbed into mouse skin, based on the *in vitro* studies of absorption of benzo(a)pyrene by Storm *et al.* (1990). That is, the applied amounts were multiplied by 0.6 to estimate amounts absorbed into mouse skin. BaP is in the middle of the range of absorption fraction among PAHs that have been examined (Dankovic *et al.* 1989), and these authors proposed BaP as a reasonable choice as a representative marker for absorption of 4- and 5-ring PAHs. Strictly speaking, the absorption fraction applies to the PAH component of coal tar and not to the whole mass of tar, but we apply the fraction to the mass on the understanding that we are using the mass of tar as a surrogate for the PAH fraction that we presume is largely responsible for the carcinogenic effect seen in mouse skin. The dermal absorption experiments of Storm *et al.* (1990) examined BaP dissolved in acetone, reasonably similar to the toluene vehicle in the Fraunhofer bioassay. The absorption was determined over 24 hours, while the Fraunhofer mice, being exposed twice-weekly, had several days between applications. It is likely that little further absorption into skin happens in this longer interval, with unabsorbed material either volatilizing, being rubbed off, or being shed with superficial stratum corneum.

It should be noted that the normalization is by the mass of the whole skin, not just exposed skin (which is about 10% of the whole skin). This method is used as a normalization aimed at the question of cross-species equivalence, in which organ-to-organ projection of risk potential will be assumed, as discussed further below. Human exposures will also be estimated as average daily amount absorbed per g of whole skin, even though humans exposed on only part of their skin as well—also about 10%. This obviates questions about the exact area of skin exposed, since if the given amount of tar is spread around a larger area, the concentration per unit skin is decreased, but the area is increased in a way that exactly offsets it.

Doses calculated in this way were 0, 0.00259, 0.00778, 0.0259, 0.0778, and 0.233 daily mg tar absorbed into skin per g of skin for the controls and the 0.1, 0.3, 1.0, 3.0, and 9.0 mg dose groups, respectively.

3.2 Dose-Response Analysis

Dose-response characterization of these data was carried out using the Multistage-Weibull time-to-tumor model (MULTI-WEIB; Howe and Crump 1991). This model combines a multistage model in dose (as prescribed by the Proposition 65 defaults) with an empirically estimated exponential increase in tumor incidence over time at a given dose. By using information on the time of appearance of tumors and of deaths without tumors, it is possible to correct for intercurrent mortality (causes of death other than from the tumor of interest), and to estimate the level of risk that would be expected to occur at various durations of exposure.

Such analysis is appropriate in this case because:

1. the reduced survivorship in higher dose groups in the Fraunhofer study could cause animals to die before tumors they would have developed have yet appeared. Time-to-tumor analysis is the best way to correct for intercurrent mortality;
2. the Fraunhofer study ran only 78 weeks, somewhat less than a conventional mouse lifetime of 2 years. One wants to know how much risk might be engendered by full lifetime exposure, that is, if the mice were exposed at the same rate for a full 2 years, and this can be estimated with a time-to-tumor model.

Tumor-bearing animals dying before scheduled terminal sacrifice were counted as having fatal tumors; animals killed at scheduled terminal sacrifice were counted as having incidental tumors. (The distinction between fatal and incidental is made because, for the former but not the latter, the time of initiation of the tumor dictates the time at which it is recorded.) We considered a multistage component of degree 3, with no time-lag parameter.

3.3 Fraunhofer Dose-Response Results

We have used the Multistage-Weibull model with no estimated lag-time between initiation and detection of tumors, and with a multistage degree of three. (Sensitivity analysis examining models with lag times produced estimates of zero lag, *i.e.*, the same results as without lag, and different degrees of the multistage produced results that did not differ markedly from degree 3.) Our specific use of the Multistage-Weibull model produces an equation of the form

$$P(d,t) = 1 - \exp[(-q_0 - q_1d - q_2d^2 - q_3d^3)(t^C)]$$

where $P(d,t)$ is the probability of response (*i.e.*, bearing a tumor of interest) at dose d after time t , the q 's are nonnegative estimated parameters of the multistage component, and C is the estimated exponent of time (measured in days).

A good fit was achieved. The fitted parameters are tabulated below.

Table I. Fitted Parameter Values for the Multistage-Weibull Model

Parameter	Value
q_0	0.00000
q_1	2.96889×10^{-17}
q_2	0.00000
q_3	1.03857×10^{-14}
C	6.35082

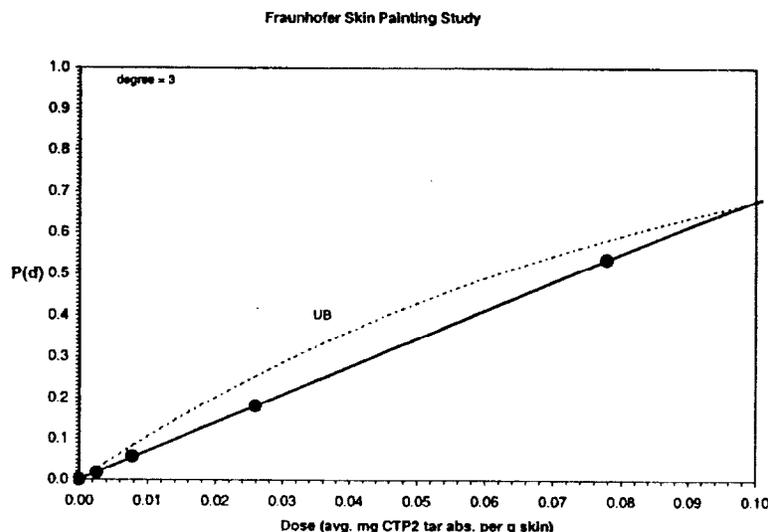
A linear component (*i.e.*, a positive value of q_1) is necessary in the Maximum Likelihood Estimate (MLE) of the parameters. (This was true for all fits for different degrees.) As a result, there is no pronounced nonlinearity in dose-response curve shape. It is not S-shaped or steeply rising with dose. That is, there is no indication that the effects leading to skin tumors in mice come into play only as high-dose effects. Instead, the effect among mice in the Fraunhofer study seems to diminish gradually with decreasing dose over coal tar amounts differing by almost two orders of magnitude. The shape is consistent with a mode of action such as mutagenic initiation, in which risk is expected to be proportional to the exposure to mutagenic components.

The report on the Fraunhofer study (Fraunhofer 1997) indicates that higher doses had higher frequencies of skin ulceration. (Dosing was temporarily stopped for individual mice having such ulceration until the lesions healed.) Thus, high-dose mice may have had more cytotoxic responses in skin than those at lower exposure levels, but there is no evidence that this acts synergistically with the presumed mutagenic mode operational at all dose levels, based on previous understanding of PAH carcinogenicity. The dose-response curve shape is also consistent with results for skin painting of BaP at several doses conducted by Schmähl *et al.* (1977).

On the other hand, the increase in tumor incidence with time (at a given dose) is quite steep; the estimated value of the exponent of time (parameter C) is about 6.3, which is high compared to values more often seen in carcinogenicity studies, which are often in the range of 3-4. The large power of time means that, if the fitted equation is used to project risk estimates after 104 weeks (2 years), the estimated risks a good deal higher than those seen after the 78 weeks of the experiment's conclusion. (The effect on estimated low-dose potency is about a factor of 5 lowering of the lower bound dose to produce a 10^{-5} risk.) This represents an estimate that, if the Fraunhofer mice were kept on study (including continued dosing) for 104 weeks, the rise in tumor rates with time (based on rises observed during the study) would lead to a potency estimate 5-fold higher.

The 2-year rodent lifetime as a basis for risk determination is partly a matter of convention. The Fraunhofer study's duration of 78 weeks (1½ years) is nearly a full lifetime. Such a duration was considered a full mouse lifetime in earlier lifetime bioassays as conducted by the US National Cancer Institute (before the testing program was moved to the National Toxicology Program). Thus, this five-fold adjustment seems a large one to make based on model projections, and the relevance of this difference to projection of human risks (where exposure continuing uninterrupted for full lifetime is not expected and is handled by linear prorating, not with exponential effects) could be questioned. Accordingly, both the 78-week result and the 104-week projected result will be carried through the analysis.

Figure 1. Fit of the Multistage model of degree 3 to mouse skin tumors in the Fraunhofer study.



The results of this modeling, expressed as doses to mice (in average daily mg tar absorbed into skin per gram of entire skin) expected to produce a 10^{-5} lifetime risk, are tabulated in Table II. For comparison, the results from using the more commonly applied Multistage model (GLOBAL86, Howe and Crump 1986) are also shown. In running this model, which has no time component, mortality adjustment is by eliminating from the count of animals on test all those individuals dying before the first appearance of the tumor of interest in the study (day 361), a method that approximates the more exact allowance for mortality that the Multistage-Weibull model can perform. (The 9 mg dose group had to be dropped from this Multistage analysis owing to its early termination.) Also, the Multistage model cannot by itself project past the 78 weeks of the study's duration. Its results are very similar to those from the Multistage-Weibull for 78 weeks. The results from the Multistage model are presented graphically in Figure 1. The linearity of the dose-response relationship in the range of bioassay doses is evident, even for the maximum likelihood fit (dark line). The upper 95% bound curve with respect to low doses (dotted line) is only slightly higher. That is, in this case, the difference between a best-fitting and an upper-bound low-dose extrapolation is minimal.

Unlike the time-to-tumor analysis, the Multistage model provides no direct way to extrapolate the effects to be expected if the experiment had been run longer. When an experiment lasts less than a lifetime (and a time-to-tumor analysis has not or cannot be used), a standard practice is to calculate a correction factor to the potency as

$$\left(\frac{L_L}{L_E} \right)^3$$

where L_L and L_E are the length of lifetime and the length of the experiment, respectively. In the present case, if a standard mouse lifetime is 2 years (104 weeks) and the experiment lasted 78 weeks, the factor is $(104/78)^3 = 2.37$. That is, it is estimated that if the experiment were run out for a full 104 weeks, the potency would be 2.37-fold higher (and the risk-specific dose for 10^{-5} risk 2.37-fold lower) than the value observed at 78 weeks. This correction is based on invoking a "typical" power of time for the increase in tumor incidence with time. In the case of the Fraunhofer study, it is evident that the actual power of time (6.3) is greater than is typically seen (a default of 3), and so the approximate generalized correction of 2.37-fold is smaller than the study-specific correction (embodied in the time-to-tumor analysis) of about 5-fold. The approximate correction is shown here for completeness, since it is what would typically be done for a default analysis.

Table II. Doses to mice (in average daily mg tar absorbed into skin per gram of entire skin) expected to produce a 10^{-5} lifetime risk.

	MLE	Lower 95% Bound
Time-to-Tumor model - 78 weeks	1.33×10^{-6}	8.42×10^{-7}
Time-to-Tumor model - 104 weeks	2.20×10^{-7}	1.67×10^{-7}
(Multistage model)	1.37×10^{-6}	8.90×10^{-7}
(Multistage – corrected to 104 weeks)	5.77×10^{-7}	3.75×10^{-7}

The results in Table II could be re-expressed as average daily mg tar applied per mouse by multiplying by the assumed mass of mouse's skin, namely by 6.612 g.

4 Cross-species equivalence of doses, and extrapolation to humans

By basing projections of human risk on the results of experimental exposures to mice, we are accepting the proposition (at least provisionally) that mice constitute an appropriate model, and that analogous biological processes occur in humans. To apply this assumption quantitatively, one must define the doses that are expected to lead to equivalent degrees of response in humans and in mice.

How to represent doses to achieve cross-species equivalence in toxic effects is a difficult area and one that has enduring controversy. In practice, default assumptions are used. These defaults reflect both (1) empirical experience regarding the dose-scaling methods that appear to reconcile toxicological responses in different species (including different experimental animal species and these compared to human risk based on epidemiology), and (2) an attempt to account logically for the general differences in scale (size, duration, and the pace of physiological processes) between small, short-lived, rodents with fewer cells but faster physiological processes, and large, long-lived humans with more cells but relatively slower physiological processes (US EPA 1992). That is, the default attempts to reflect our general experience with doses that are observed to produce equivalent toxicity in differently sized species, and it attempts to do so by accounting for the influence of the principal differences in size-scale and time-scale of organisms that are presumed to be substantially similar qualitatively.

In considering the question of appropriate scaling of doses to skin, we first consider the default method set out in the Proposition 65 law. Then we consider possible modifications of this approach that, while

logically consistent with it, make it more appropriate for assessing doses to skin. Finally, we briefly consider whether further modifications can be based on knowledge of PAH metabolism in human and mouse skin, and information regarding relative sensitivity of human and mouse skin.

The Proposition 65 law specifies that scaling daily administered doses in proportion to the $^{2/3}$ power of each species' body mass is to be used as a default, unless agent- and species-specific data "are of such quality that physiologic, pharmacokinetic and metabolic considerations can be taken into account with confidence."

The use of the $^{2/3}$ power of body mass reflects the relatively slower pace of physiological processes in larger species. (Current thinking favors use of the $^{3/4}$ power of body mass as an index for scaling these rates [USEPA 1992, 1996].) The slower pace includes the processes of uptake, distribution through the body, metabolism (either activating or detoxifying) and clearance from the body. It is not just pharmacokinetic processes that are affected; human cells have longer lifetimes and hence slower division rates. Maturation, growth, aging processes, and so on go more slowly as well. Because the slower tempo affects physiological rates generally, the component processes stay in proportion to one another, and the fates of compounds in the bodies of rodents and humans are expected to operate in parallel. Thus, under the assumptions of the default cross-species dose scaling method, the fraction of a dose that is metabolized is expected to be about the same, but the time over which this process proceeds is drawn out in humans *vis-à-vis* mice. Measures of target organ exposure that are integrated over time, such as areas under concentration-time curves, are larger in larger species because the processes that change concentrations operate more slowly.

Understanding the use of the $^{2/3}$ power scaling hinges not on thinking about the fate of a single day's dose, then, but rather on considering that the increments of exposure come every 24 hours, a repeat-interval that represents a much shorter span of physiological time for a human than it does for a mouse. The scaling is really not about dose as such, but about time. The daily dose is adjusted so that the dose *rate* matches the slower pace of physiological time in humans. Use of the $^{2/3}$ power scaling of doses (or, in more modern usage, the $^{3/4}$ power scaling) represents the embodiment of the principal that differently sized species experience similar toxic reactions when the dose rate is matched to their tempo of physiological time.

The situation for dermal uptake of coal tar residues is somewhat different than the administered doses directly addressed by the default scaling method specified in the Proposition 65 law. Uptake is directly

into the target organ, metabolic activation is local, and the rest of the body mostly serves as a sink, providing the means by which absorbed PAHs and their metabolites are cleared away from the skin.

There are three issues to consider: amount of agent per unit of skin, scaling to account for dose rate, and the volume of tissue.

4.1 Amount of agent per unit of skin

PAHs require metabolic activation for their carcinogenic effects, and the effects of the metabolites are believed to occur through their covalent binding to DNA (*i.e.*, the formation of DNA-adducts), presumably leading to somatic mutations in the cells that bear them when cell division is attempted in the presence of unrepaired adducts.

All the PAHs in coal tar deposited on the backs of bioassay mice, or those forming a residue on the skin surface of humans after shampooing with coal tar shampoo, must be absorbed into the skin before they can have any effect, and all coal tar constituents that enter the body must do so through such absorption. Experimental data indicate that PAHs are fairly readily absorbed, but the process is not quick. The residue left on the skin of mice by the process of skin painting (and on human skin as a residue after use of coal tar shampoo) will undergo substantial (but not necessarily complete) absorption in the interval between episodes of new application. Once absorbed, there appears to be little lateral transport through the skin. Loss of PAHs from the skin to the blood (and thence to the rest of the body) is a relatively slow process, so residence time in the skin of absorbed material appears to be substantial. The skin has metabolic activity toward PAHs, and the formation of macromolecular adducts in skin tissue appears to be due to such local metabolism.

Thus, the amount of PAHs metabolized in the skin tissue should be a good measure of dose, and the presence of DNA-adducts should constitute a relevant biological marker.

A measure of the daily amount of coal tar constituents absorbed per unit of skin volume is analogous to the kind of absorbed dose envisaged by the default, since for internal organ effects, the exposure of the organ is assumed to be proportional to the systemically absorbed dose.

In humans, Van Schooten *et al.* (1994) showed that pyrene metabolites appeared in urine of subjects using a coal tar shampoo only after a considerable time lag. These results suggest that absorption is a fairly slow process acting on the residue left behind from use of shampoo rather than from systemic uptake during the shampooing event itself. It is useful to distinguish the issue of amount of residue left by use of coal tar shampoo and the amount of absorption of that residue. These issues are discussed further under the topic of human dosimetry.

In principle, the daily amount of absorbed PAHs that are metabolized in skin would be a better dose measure than the amount absorbed, since metabolic activation is necessary for their carcinogenic effects. In the present case, however, there is insufficient basis to make a sound estimate of such a metabolized dose. Under the default dosimetry expectations, this metabolized amount would be expected to be proportional to the absorbed amount in both mice and humans, and so the amount absorbed into skin serves as a surrogate for an estimated amount metabolized in skin. The default's presumption of proportionality of absorbed and metabolized doses arises because the (assumed) slower metabolic activity in humans is balanced by the slower removal of unmetabolized PAHs from the skin by its blood supply. (It must be remembered that the realized metabolic rate is a product of the enzyme activity and the substrate concentration; humans have lower activity, but the substrate concentration is kept higher for longer owing to the slower clearance. In the end, metabolism and nonmetabolic clearance are competing as causes of removal of compounds from the skin, and when they are both slower in the same proportion, the fraction metabolized is equal.)

It is noteworthy that experimental determination of DNA-adduct levels in human and mouse skin within 24 hours of similar amounts of coal tar application showed that adduct levels were very similar in the two species (Phillips *et al.* 1990, Schoket *et al.*, 1988a,b). This implies a similar degree of absorption and metabolic activation of the absorbed amount. (The interpretation must be tempered because the human skin was examined *in vitro* in a setting that may have diminished nonmetabolic clearance; the impact of this fact depends on the unknown importance of nonmetabolic removal of PAHs from the skin.)

ICF (2000) cites data from several sources to estimate the relative metabolic activity toward PAHs in mouse and human skin. Their estimates of uninduced activity show mice to have 7.5-fold higher activity (on a per g of protein basis) than humans. If this translates to *in vivo* activity, it would be almost exactly in accord with the presumption of scaling of physiological rates by the $\frac{3}{4}$ power of body mass (which implicitly assumes a metabolic activity about 7-fold lower). The fully induced rates estimated by ICF

differ by a factor of 16.5—somewhat more difference between species than is presumed by the default. ICF goes on to cite the default's presumption about nonmetabolic clearance (a 7-fold lower rate in humans) as exacerbating the metabolic difference in its impact on amount metabolized in the skin. In fact, the slower rate of clearance in humans should enhance rather than reduce the amount metabolized by slowing the removal of substrate for the enzymes in the skin. How these opposing tendencies combine in their effect on metabolic activation in skin should really be evaluated in the context of a model of skin pharmacokinetics, since the effect will depend on absolute amounts and not just on ratios. Nonetheless, one can note that the ratio of relative AHH enzyme activity to relative nonmetabolic clearance is about unity for uninduced skin, and 2.3 for fully induced skin, which would seem consistent with the notion that the degree of metabolic activation of the absorbed amount of PAHs is at most only slightly greater in mouse than in human skin.

4.2 Scaling to account for dose rate

As discussed above, the scaling of doses by the $^{2/3}$ power is not to correct the daily dose measure as such, but to adjust the dose *rate* so that the incremental daily amount is in proportion to the species' rate of physiological time. Under this interpretation, it is only when the dose measure already includes the consideration of the different pace of physiological processes (as does an area-under-the-curve measure) that the scaling by a power of body weight is obviated by information on daily target-organ-level dose. Accordingly, the daily dose in terms of amount of PAHs absorbed per unit of skin (or in terms of daily amount metabolized per unit of skin) that is assumed to lead to equal lifetime cancer risk in mice and humans is an amount proportional to the $^{2/3}$ power of body weight. Because current thinking has it that scaling daily amounts in proportion to the $^{3/4}$ power of body mass is in closer accord with the pace of physiological processes that this method is intended to address, scaling by the $^{3/4}$ power of mass is also considered.

4.3 Volume of Tissue

Transformation of a cell from a normal to a carcinogenic state is a phenomenon of individual cells; a tumor begins when one of the many cells in a tissue undergoes the rare events that lead it to lose control of cell division and cause it to divide without limit, growing in mass and invading surrounding tissue. The risk of developing a tumor for the whole organism is a function of the collective risk for each of its potentially affected cells accumulated over all the cells at risk.

Larger organisms such as humans have many more cells in a given organ than do mice or rats. The number of cells is approximately proportional to organ mass, since cell size is relatively constant across species. Despite this, larger organisms do not have massively higher risk of cancer. In the default scaling procedure, an assumption of organ-to-organ equivalency is made in the sense that per-organ risks are assumed to be the same in rodents and humans when all the cells in those organs have the equivalent exposure, despite the fact that humans have so many more cells at risk. For this to be so, the per-cell risk must be very much less in human tissues than in rodent tissues.

Similarly, the toxicologically equivalent doses are figured as amount that can be experienced daily for a lifetime to produce an equivalent lifetime cancer risk. Humans live some 35 times longer than experimental rodents, so the cells of their target organs experience the daily dose deemed toxicologically equivalent for a much longer time and their per-cell cumulative lifetime exposure is much higher. Again, in order for the assumed lifetime-to-lifetime equivalence to hold, the per-cell per-unit-time risk must be much less in long-lived humans than in short-lived rodents. (This factor is partially, but not wholly, addressed by scaling dose rates in proportion to the species-specific pace of physiological time.)

In making an organ-to-organ and lifetime-to-lifetime extrapolation of carcinogenic risk when (appropriately scaled) target organ concentrations are equal, the sensitivity of human tissue on a per-cell, per-day basis is assumed to be radically less than that of rodents. The reasons for this insensitivity are not well understood, and no method exists to make an exact accounting of the roles of the various possible factors, which are thought to include better DNA repair, slower rates of cell turnover, and a generally slower pace of physiological processes. It is clear, however, that demonstrations that mouse skin or mouse skin cells are "more sensitive" than human skin to carcinogenic transformation on a per-cell per-day basis are not inherently in conflict with the existing presumptions implicit in the default approach to cross-species extrapolation. (For instance, xenograft studies, in which small amounts of human or mouse skin are grafted onto the backs of mice and then treated with carcinogens, inherently compare mouse and human skin on a per-cell per-day basis, and not on a per-organ per-lifetime basis. The amounts of skin and durations of treatment are of a scale appropriate to mice, but to the extent that grafted human skin retains its human properties (as is, of course, the intention of the experiment), one must allow that one is examining a much smaller fraction of the total human skin that is exposed for a much smaller fraction of a total human lifetime than is true for the mice.)

The usual default cross-species scaling methodology makes the implicit assumption that the number of cells at risk is proportional to body mass. In the case of skin, humans certainly have many more cells at risk than do mice, but the proportionality is less than the difference in body mass would indicate. Brown *et al.* (1997) cite data indicating that human skin constitutes about 3.71% of body mass, while mouse skin constitutes an estimated 16.54%. The difference reflects the lower surface:volume ratio of humans compared to mice, partly tempered by the fact that human skin is thicker.

An implicit organ-to-organ scaling based on mass would overestimate human risk by about a factor of 4.5 based on the difference between a body mass ratio of 1,750 for a 70 kg human and a 40 g mouse, and a skin mass ratio of 392 (*i.e.*, $1750/392 = 4.5$). It would seem appropriate to reduce the estimate of human risk (*i.e.*, to raise the NSRL) by this amount to address the fact that humans have fewer cells at risk than is implicitly assumed by the usual methods. One could further consider that the skin cells at risk belong not to the whole skin, but to a fraction of it (the epidermis), and that the relative sizes of this fraction might make a better adjustment. To do this, however, one would want to know the relative doses experienced by different fractions of the skin, since PAHs absorbed into skin and metabolized there could have distribution more widely than in any one layer. We have a rather crude measure of absorbed dose to the skin as a whole, and so too fine an adjustment does not seem warranted.

5 Calculation of NSRL

Doses are expressed as amount of tar absorbed per unit of the whole skin since the arguments regarding the relative efficacy of absorbed tar to cause skin cancer in mice and humans were made on the basis of an organ-to-organ comparison referring to comparative skin mass. The issue of fraction of the skin that is exposed is handled implicitly. For example, if one assumes that the fraction of mouse skin painted with coal tar in the Fraunhofer study is 10%, then the mass per unit of exposed skin is 10-fold higher than a mass per whole skin, but the number of cells in that skin is 10-fold lower. A whole-skin basis obviates the concern about estimating a difficult-to-know parameter (the fraction of skin exposed in mice and humans) that drops out of the equation anyway.

Two bases for an NSRL estimate are considered from the Fraunhofer study: the lower bounds on dose associated with a risk of 1-in-100,000 based on a 78-week duration or based on a projected 104-week duration. The values calculated from fitting the Multistage-Weibull time-to-tumor model are 8.42×10^{-7} and 1.67×10^{-7} mg tar absorbed to the skin per g of entire skin, respectively. For completeness (since it is

the traditional method), calculations are also made using the Multistage model and the approximate correction to 104 weeks (*i.e.*, $[L_I/L_E]^3$), as described at the end of Section 3.3. The lower bound dose for 1-in-100,000 risk from this approach is 3.75×10^{-7} .

It is considered that no appropriate basis exists to estimate relative degrees of metabolic activation of the absorbed amount in mice and humans. To the degree that data exist (the DNA-adduct data described by Phillips *et al.* 1990), they suggest that a similar proportion is activated. Application of proportionality corrections regarding relative enzyme activity and relative clearance rate are deemed inappropriate in principle for addressing species differences in the degree of metabolic activation, and in any case the implications of such an approach are belied by the DNA-adduct data. Similarly, it is considered that no appropriate basis exists for making quantitative corrections for the relative sensitivity of mouse and human skin to carcinogenic transformation by coal tar. Observations of per-cell, per-unit-time differences in sensitivity are in accord with the already existing presumption that on such a basis human skin should be much less sensitive. No data (except for the skin volume relative to body mass) suggest that the usual presumptions about organ-to-organ comparability of cancer risks need be altered.

It is considered appropriate to modify human risk estimates by a factor of 4.5, however, to account for the fact that the whole human skin has fewer cells than is implicitly assumed by organ-to-organ methods that are based on assumptions of proportionality of organ mass to whole body mass. Using the usual organ-to-organ assumption without correcting for nonproportionality of skin mass to body mass is also considered for comparison.

Finally, it is considered appropriate to apply the correction of body weight to the $^{2/3}$ power to account for differences in the pace of physiological time in mice and humans, since this factor is included in the default methodology specified by the Proposition 65 law and has not been obviated. Scaling by the $^{3/4}$ power of body weight is also considered.

For convenience, the proposed NSRLs are expressed in units of lifetime average daily μg of tar absorbed into skin. (This is done by multiplying the human equivalent doses in mg per g of whole skin by 1000 $\mu\text{g/g}$ and by the weight of human skin, assumed to comprise 3.71% of a 70 kg body mass [Brown *et al.* 1997].) Human exposures should be rendered in these units for comparison. The alternatives are laid out in Table III.

The alternative corresponding to the Proposition 65 default is to use: (1) the Multistage model corrected by the factor $[L_A/L_E]^3$ (*i.e.*, "MS- corr to 104," meaning the Multistage model corrected to 104 week duration), (2) the $^{2/3}$ power of body mass scale factor, and (3) no correction for the relative numbers of cells in skin. The lower bound on dose associated with a 1-in-100,000 estimated human risk according to this Proposition 65 default approach is 0.08 $\mu\text{g/day}$ of coal tar absorbed into skin.

Several alternative methods for calculating an NSRL are also presented, embodying various modifications to the default procedure as discussed earlier in the present report. These alternative calculations differ from the Proposition 65 default in that: (1) the Multistage-Weibull time-to-tumor model is used to provide a more accurate accounting of intercurrent mortality; (2) the time-to-tumor model is considered as a more accurate, compound-specific, empirically-based method for extrapolating from 78 weeks to the risk expected at 104 weeks in the bioassay mice; (3) in addition to the cross-species dose scaling assumption of proportionality of daily doses to the $^{2/3}$ power, we consider scaling in proportion to the $^{3/4}$ power, as better accomplishing the goals of such scaling in view of data on the rates of physiological processes; (4) a cells-in-skin factor of 4.5 is applied to account for the fact that the default method implicitly assumes that target organs are in proportion to body mass, and skin is a smaller fraction of body mass in humans than in mice; and (5) for all alternatives, the 1-in-100,000 risk-specific dose based on the maximum likelihood estimate (MLE) is also shown (in parentheses) to illustrate that, owing to the linearity of the best-fitting curves, this central estimate is not markedly different from the 95% lower bound (LB) estimate. Table III lays out all of these alternatives, considering the various departures from the Proposition 65 default in all combinations.

Of these alternatives, the one that is considered most appropriate is represented by the second line of Table III. It is the analysis that: (1) uses the time-to-tumor analysis of the Fraunhofer study; (2) uses the 78-week results without extrapolation to 104 weeks, on the grounds that human partial-lifetime exposure is also expected, and that such exposures are estimated in humans based on lifetime averaging; (3) employs the $^{3/4}$ power of body mass scaling method as better embodying the goals of physiological time scaling (US EPA 1992); (4) employs the 4.5-fold allowance for differences in cells-in-skin relative to body mass; and (5) uses the lower bound dose estimate. This line of analysis yields a No-Significant-Risk Level of 1.51 $\mu\text{g/day}$ of coal tar absorbed into skin, averaged over lifetime. It is this value that the present report recommends as an NSRL for dermal exposure to coal tar.

Table III. Calculated No-Significant Risk Levels for Dermally Applied Coal Tar, Using Various Bases for Extrapolation

duration in mice	scale factor	cells-in-skin factor	NSRL	
			$\mu\text{g/d}$ absorbed into skin (based on MLE)	based on LB
78 weeks	2/3 power	4.5	(1.28)	0.81
78 weeks	3/4 power	4.5	(2.38)	1.51
78 weeks	2/3 power	none	(0.29)	0.18
78 weeks	3/4 power	none	(0.53)	0.34
104 weeks	2/3 power	4.5	(0.21)	0.16
104 weeks	3/4 power	4.5	(0.39)	0.30
104 weeks	2/3 power	none	(0.047)	0.036
104 weeks	3/4 power	none	(0.088)	0.067
MS – corr to 104	2/3 power	4.5	(0.55)	0.36
MS – corr to 104	3/4 power	4.5	(1.03)	0.67
MS – corr to 104	2/3 power	none	(0.12)	0.08
MS – corr to 104	3/4 power	none	(0.23)	0.15

6 Human Exposure to Coal Tar *via* Shampoo

For consumer products, Proposition 65 considers the risk from an average exposure to an average user. The use of coal-tar shampoo by consumers would typically involve one use per day, but the number of days of use may vary considerably. Some users will be occasional employers of medicated shampoo, and the daily use may extend for only a few days, with such bouts repeated at widely separated intervals. Other users (those with chronic scalp conditions) may be expected to use coal-tar shampoo more or less regularly for extended periods.

The duration of contact with the shampoo for each daily use is quite brief, and most of the material is washed off, leaving a small fraction of the amount used as a residue on the skin. Since the medicinal effectiveness of coal tar depends on this residue, it is clear that some remains after shampooing. What we count as residue includes some amount adhering to the skin surface, but it is likely that much of it is superficially absorbed into the outermost stratum corneum, the topmost, nonliving layer of the skin.

Estimates of the amount of residue that may be left by shampooing vary. Barker and Winrow (1979) found that about 1-2% of radiolabeled zinc pyridinethione in a non-coal-tar shampoo product remained on skin in human volunteers using the product in sink wash or shower conditions. It is not clear whether residues for this marker compound would be predictive of coal tar residues, since superficial absorption will depend on the affinity of the material for the stratum corneum. In unpublished in-house work at a shampoo manufacturer, Orth and Widjaja (1993) used fluorescence spectroscopy to measure skin residues after applying coal-tar shampoo or coal-tar conditioner to the forearms of volunteer subjects, leaving the material for one minute, and then rinsing it off. They estimated that 15.2% and 16.1% of the tar in two shampoo products remained on the skin after this procedure, and that 39.7% of the tar in a coal-tar conditioner remained. The volume of rinsing-off water (100 ml used over 1 minute of rinsing and rubbing of about 20 cm² of exposed skin) may be somewhat small compared to the rinsing occurring during showering, but it is unlikely that further rinsing would remove superficially absorbed material. Residues on the scalp might be higher owing to residue on the hair as well as the skin of the scalp, although it is not clear how much of the material adhering to hair becomes available for eventual absorption into the skin. Also, users of coal-tar shampoo may leave the material on the scalp for longer than one minute (indeed, they are usually instructed to do so by the product labels), perhaps leading to a larger superficial absorption than found by Orth and Widjaja. On the other hand, during shampooing, not all the material is in as close contact with skin as it was in the experiments of Orth and Widjaja, reducing the opportunity to adsorb onto the skin surface.

The experiments of van Schooten *et al.* (1994) show appearance of metabolites of pyrene (a PAH in coal tar) in the urine of volunteers using coal-tar shampoo, indicating that PAHs are absorbed into the body as a result of coal-tar shampooing. The time pattern of appearance of metabolites of pyrene in the urine suggests that systemic absorption is delayed. Presumably, the observed delay is due to the gradual absorption of post-shampooing residue into and through the skin, and thence into the bloodstream. These data have been used by ICF (2000), together with information on the amount of pyrene in shampoo, the fraction of residue absorbed systemically, the fraction of pyrene metabolized, and the fraction of the metabolite appearing in urine, to estimate residues from use of coal-tar shampoo. Their estimates (which depend on acceptance of a variety of estimates for contributing factors, some of which could be debated), range from 0.003% to 5.4%, with a mean of 0.86%. The uncertainty in these results, and the difference between them and the results of direct observations of residues, illustrates that the amount of residue left

after shampooing with coal tar shampoo constitutes an important uncertainty in any analysis of coal tar shampoo risks.

The NSRL developed in Section 5 of the present report is described in terms of amount of coal tar absorbed into skin (and not just adsorbed onto the skin surface). Although the contact of skin with the shampoo is brief, the time available for absorption into the skin of the residue that remains is substantial. This process may be slow, because it involves diffusion of the material through the stratum corneum and the underlying viable part of the epidermis (where metabolism occurs). Removal from the epidermis is through metabolism and through diffusion into the underlying dermis, which is supplied with blood vessels that can carry away absorbed material, leading to systemic exposure. Because of the slow, gradual nature of these processes, the skin constitutes a reservoir of absorbed material; residence time in the skin appears to be substantial, with gradual redistribution *via* the bloodstream.

Estimates of the degree of absorption of residue into human skin also vary. Storm *et al.* (1990) found about 30% of applied BaP was absorbed within 24 hours by viable human skin in a flowthrough diffusion cell. (Most of the absorbed material remained in the skin, and little appeared in the receptor fluid, consistent with the skin's reservoir role, described above.) ICF (2000) estimates the degree of absorption of residue in human skin to be between 25 and 100%, with the latter figure deemed most likely.

Taken together, the estimates of residues and of absorption of those residues suggest that as much as 16% or as little as 0.00075% of the amount of coal tar in the shampoo used during a shampooing event may be absorbed into skin. The higher figure is based on 16% residue from Orth and Widjaja (1993) and 100% absorption of residue from ICF (2000) and the lower is based on 0.003% residue from ICF (2000) and 25% absorption of residue. The chief reason for the wide range is the uncertainty regarding residues. The estimates based on the van Schooten *et al.* (1994) study seem the most relevant, being based on actual shampooing use of coal-tar shampoo, but they are also rather indirect and contingent on a number of assumptions about imperfectly known extents of metabolism and appearance of metabolites in the urine. Since the amount of residue is estimated by "backing out" from a calculation of absorbed dose, the estimate of absorbed dose should be somewhat less uncertain than the residue estimate.

The present report does not attempt to conduct a full human exposure assessment. Doing so entails not just estimating residues and absorption per shampooing event, but also estimating the average lifetime frequency of shampooing events using coal-tar shampoo. Instead, we refer to the assessment conducted

by ICF (2000). Elements of this assessment can be debated. For example, its estimates of residues are a good deal smaller than those of Orth and Widjaja (1993), and the interpretation of data on amounts and frequencies of shampoo use can be questioned.

The ICF assessment of exposures employs a Monte Carlo analysis to characterize the impact of variation in the parameters of the exposure calculations. In this analysis, most variable elements represent interindividual variability (e.g., years of shampoo use, ounces used per year) while others (the fraction of residue absorbed, excretion of pyrene metabolites) represent uncertainty about a quantity that presumably applies nearly equally to all individuals. The range of the variable describing fraction of residue absorbed in the Monte Carlo analysis (25% to 100%, represented as a uniform distribution) refers to uncertainty regarding the human-specific absorption fraction, not to variation among individuals. By taking the mean of the Monte Carlo analysis output distribution, one is effectively assuming that the mean value of the uncertainty distribution of absorption fraction (62.5%) applies.

Despite some questions and reservations, the results of the ICF (2000) exposure analysis can be accepted provisionally for comparison with the present report's NSRL. This includes ICF's assumptions about residues and absorption as well as its means for including uncertainty about these parameters in its calculations.

ICF (2000) estimates that "average" human exposures expressed as average daily μg of tar appearing as residue on skin (and not yet absorbed into skin) are about 5 to 26 μg . The approximate 5-fold range reflects the range of coal tar content of shampoo products. That is, a separate Monte Carlo analysis of exposure was conducted for each of several different coal-tar contents from 0.5% to 2.5%, and each assumed content has its own output distribution, the mean of which constitutes an estimate of mean lifetime average daily dose (LADD) in $\mu\text{g}/\text{day}$. The 5 μg figure arises from analysis assuming 0.5% coal-tar shampoo and the 26 μg figure arises from assuming 2.5% coal-tar shampoo. (It should be noted that some shampoo labels claim up to 5% coal tar content; 5% coal-tar shampoo use would presumably lead to an estimated average LADD of 52 $\mu\text{g}/\text{day}$).

This is a residue amount, and for comparison with the NSRL in the present report, the figures must be corrected for the fraction of residue that is absorbed into skin. As noted above, by including the uniform distribution of absorption fraction in the Monte Carlo analysis and then taking the mean of the resulting distribution, the ICF assessment is in effect assuming 62.5% absorption. It would seem preferable to

remove this uncertainty from the calculation of mean exposures, since it represents uncertainty, not variation among shampoo users. If the true value is higher or lower than 62.5%, then population mean absorbed doses will be higher or lower in direct proportion. (Since the absorption factor is uncorrelated with other variable factors and is applied as a simple multiplicative term, one can calculate the mean population exposure for a given absorption amount by multiplying the ICF result by the ratio of the assumed absorption fraction to 62.5%.) Using ICF's "most likely amount of residue absorbed" of 100%, the estimates of mean absorbed daily dose should be given by multiplying the ICF means by $(100/62.5) = 1.6$, giving estimates of population mean daily absorbed dose of 8 μg and 41.6 μg for 0.5% and 2.5% coal-tar shampoo, respectively. (Shampoos with other coal-tar contents would have means that differ in proportion to coal-tar content. For example, a 1.0% coal-tar shampoo would have an estimated mean of 16 $\mu\text{g}/\text{day}$.)

If the lower end of the range of absorption fraction (25%) applies, the mean values reported by ICF (2000) should be multiplied by a factor of $(25/62.5) = 0.40$ to give the estimated mean daily absorbed amount. For 0.5% and 2.5% coal-tar shampoos, these estimates are 2.0 μg and 10.4 μg , respectively.

7 Comparison of Exposure Levels to the NSRL

The recommended NSRL, developed in Section 5, is 1.5 μg of coal tar absorbed into skin per day, averaged over lifetime. The dermal exposures to coal tar from use of coal tar shampoo estimated by ICF (2000) and discussed in Section 6 of the present report depend, *inter alia*, on the concentration of coal tar in the shampoo (with values from 0.5% to 2.5% considered) and on the assumption about the fraction of residue that is absorbed (with values from 25% to 100% considered). If we provisionally accept the ICF exposure analysis, the population mean lifetime average exposures in amount absorbed into skin range, from as low as 2.0 $\mu\text{g}/\text{day}$ (for 0.5% coal-tar shampoo and assuming 25% absorption of the residue remaining after shampooing) up to 41.6 $\mu\text{g}/\text{day}$ (for 2.5% coal-tar shampoo and assuming 100% absorption of residue). Values intermediate between these are found for shampoos of intermediate strength and for assumptions about absorption of residue intermediate between the extremes.

Based on this analysis, the exposures exceed the NSRL of 1.5 $\mu\text{g}/\text{day}$ absorbed into skin. They do not do so markedly, however, especially for lower strength coal-tar shampoos and for lower estimates of absorption. If the NSRL were calculated strictly according to Proposition 65 defaults without

modification (0.08 µg/day), the exceedance would be considered to be more marked. If the NSRL were calculated using projected 104-week risks for the bioassay mice using the time-to-tumor model (but otherwise as proposed herein), *i.e.*, a value of 0.30 µg/day, the exceedance would also be notable.

Given the uncertainties in both the estimation of coal tar carcinogenic potency in humans and in the estimation of human exposures from use of coal-tar shampoo, conclusions about the existence and magnitude of human risk must be tempered. Nonetheless, the *prima facie* case is made that typical use of coal-tar shampoo can result in estimates of human skin cancer risks that may moderately exceed 1 in 100,000.

8 Conclusions

This report develops a recommendation for a No-Significant-Risk Level for ongoing dermal exposure to coal tar of 1.51 µg per day of tar absorbed into skin as a lifetime average. This recommendation is based on a time-to-tumor analysis of the skin tumors produced in male CD-1 mice after 78 weeks of twice-weekly skin painting with a toluene solution of coal tar at various doses conducted by the Fraunhofer Institute (Fraunhofer 1997).

The Fraunhofer mouse study is notable for an especially wide range of dose levels having been tested, spanning almost two orders of magnitude at 5 distinct exposure levels. The dose-response curve shape is notably linear across this range, suggesting that the processes operating to cause skin tumors in mice do so similarly at moderate as well as at high doses. That is, these data do not suggest a high-dose-only syndrome of tumor formation only secondary to a toxic effect operating with an exposure threshold.

Compared to pharmaceutical coal tar, the material tested in the Fraunhofer bioassay has somewhat lower concentrations of some of the larger, multi-ring PAHs—components that are presumed to account for much of the carcinogenic effects of coal tars on skin. Hence, use of the Fraunhofer data when doses are expressed as amounts of coal tar may somewhat underestimate human risks from pharmaceutical coal tar, all else being equal.

In extrapolating the dermal carcinogenicity results from mice to humans, allowance is made for the different pace of physiological processes in mice and humans using dose scaling in proportion to the $\frac{3}{4}$

power of body mass and for the fact that, compared to internal organs for which default scaling rules are primarily intended, humans have a relatively smaller mass of skin (and hence skin cells at risk) than mice, compared to their body mass. These factors represent modifications of the default cross-species scaling methodology specified in the Proposition 65 law that, while consistent with the rationale and intent of the defaults, make the methods more appropriate to extrapolation of risks of skin tumors following dermal exposure.

There is evidence that humans may have lower metabolic activity than do mice toward the polynuclear aromatic hydrocarbons in coal tar, but the differences are not very well characterized in terms of impact on the degree of metabolic activation of these compounds. To the extent that quantitative data can be brought to bear, they suggest that metabolic differences are largely in accord with the assumptions already embodied in the default dose scaling methods, and that absorbed dose provides a reasonable surrogate for the dose metabolized in skin. Data on DNA-adducts in human and mouse skin following PAH application support this notion.

Although evidence exists that human skin is less sensitive to carcinogenic effects than is mouse skin on a per-cell per-day of exposure basis, there is no evidence to suggest that this sensitivity difference deviates from the assumptions about relative sensitivity already embodied in default scaling, which compare animals and humans on a per-organ per-lifetime basis (and hence implying that human tissues are much less sensitive on a per cell per-day basis). Although anatomical and physiological differences between mouse and human skin can be documented, there is no clear basis to conclude that such differences are critical to the potential carcinogenic process in a way that obviates the conclusion that responses in mouse skin provide *prima facie* evidence that similar responses might appear in human skin.

Nonetheless, extrapolations from animal bioassays to make projections of potential carcinogenic potency in exposed humans are inherently uncertain. There is no reason to feel that the extrapolations in the present case are any more uncertain than in other cases of use of animal data to estimate potential human cancer risks. Indeed, in the case of dermal exposure to coal tar we have evidence of histologically similar tumors being caused in skins of a number of animal species in experiments ranging over decades, and similar tumors have been observed among workers in former occupational settings, albeit at very high exposures and conditions of poor hygiene that no longer prevail. We are also presented with a particularly well characterized dose-response curve in animals, with more doses over a wider range of exposure levels than is typical, and the curve that emerges is well behaved and linear in the observed

range, increasing confidence in extrapolation. Human and mouse skin both show qualitatively similar metabolic capabilities toward absorbed PAHs, and data on quantitative differences, although limited, present no unusual species differences.

On the other hand, much remains to be known about the pharmacokinetics of PAHs in mouse and human skin, including levels of absorption, metabolic activation, enzyme induction, and clearance of compounds from skin. The possible carcinogenic role of PAHs absorbed into the bloodstream and then transported to organs in the body has not been clearly addressed. Many details of the mode of carcinogenic action in skin remain to be elucidated, and so the relevance of processes in mouse skin at higher doses to those in human skin at lower exposure levels remains a matter of presumption, with elements that can be questioned and debated. Coal tar-containing consumer products and medical treatments designed for use on skin have a long history of use in human patients. There is sparse epidemiological information on such uses, and there are important issues to bear in mind in interpreting them. Nonetheless, such use has not led to obvious evidence of increased skin cancer risks. Whether analysis of such data might provide evidence that human potency of dermal coal tar exposure cannot be as high as estimated from the animal data is a question beyond the scope of the present report.

Estimation of the degree of human exposure to coal tar through use of coal tar shampoo is challenging. Key parameters that appear to be not well established are some elements of the frequency and duration of use and the amount of residue remaining on the skin after use. The present report provisionally examines the exposure analysis conducted by ICF (2000). Depending on the strength of the shampoo and on assumptions about the amount of residue absorbed, mean daily exposures of tar absorbed into skin are estimated to range from as low as 2 µg/day to as high as 41.6 µg/day. These results suggest that the NSRL developed herein (1.51 µg/day) is exceeded by a moderate amount in average users of coal tar shampoo.

9 References

Barker E.A. and M.J. Winrow. 1979. Percutaneous (Scalp) Absorption of Zinc Pyridinethione in Humans. Report to Procter & Gamble. [As cited in Food and Drug Administration OTC Volume 160369]

Brown, R.P., M.D. Delp, S.L. Lindstedt, L.R. Rhomberg, and R.P. Beliles. 1997. Physiological Parameter Values for Physiologically Based Pharmacokinetic Models. *Toxicology and Industrial Health* 13(4): 407-484

Dankovic, D.A., C.W. Wright, R.C. Zangar, and D.L. Springer. 1989. Complete Mixture Effects on the Dermal Absorption of Benzo[a]pyrene and other Polycyclic Aromatic Hydrocarbons from Mouse Skin. *Journal of Applied Toxicology* 9(4): 239-244.

Fraunhofer. "Dermal Carcinogenicity Study of Two Coal Tar Products (CTP) by Chronic Epicutaneous Application in Male CD-1 Mice, Final Report." October 1997.

Gaylor, D.W., S.J. Culp, L.S. Goldstein, and F.A. Beland. 2000. Cancer Risk Estimation for Mixtures of Coal Tars and Benzo[a]pyrene. *Risk Analysis* 20(1): 81-85.

Goldstein, L.S., E.H. Weyand, S. Safe, M. Steinberg, S.J., Culp, D.W. Gaylor, F.A., Beland, and L.V. Rodriguez. 1998. Tumors and DNA Adducts in Mice Exposed to Benzo[a]pyrene and Coal Tars: Implications for Risk Assessment. *Environmental Health Perspectives* 106(6): 1325-1330.

Howe, R.B., and K.S. Crump. 1986. "Global 86: A Computer Program to Extrapolate Quantal Animal Toxicity Data to Low Doses." Report to Environmental Protection Agency. K.S. Crump and Co., Ruston, LA.

Howe, R.B., and K.S. Crump. 1991. "Multi-Weib: A Computer Program to Extrapolate Time to Tumor Animal Toxicity Data to Low Doses." K.S. Crump and Co., Ruston, LA.

ICF. 2000. "TERA Review Draft: Estimation of Lifetime Skin Cancer Risk from the Use of Coal Tar-Containing Shampoos." ICF Consulting, K.S. Crump Group, Inc. May.

International Agency for Research on Cancer (IARC). 1985. "IARC Monographs on the Evaluation of the Carcinogen Risk of Chemicals to Humans: Polynuclear Aromatic Compounds, Part 4, Bitumens, Coal-tars and Derived Products, Shale-oils and Soots." World Health Organization. Volume 35, January.

Ma, B.L., A. Yarborough, B. Frimpong, and K. Rozett. 2000. Comparative Analysis of Chemical DNA Adduct Formation in Mouse Lung Following Dermal Application of Coal Tar Shampoo and MGP Tar. *The Toxicologist*.

Nisbet, I.C. and P.K. LaGoy. 1992. Toxic Equivalency Factors (TEFs) for Polycyclic Aromatic Hydrocarbons (PAHs). *Toxicology and Applied Pharmacology* 16:290-300.

Orth D. and J. Widjaja. 1993. PNAH Concentration in Neutar-Containing Products. Memorandum to Warren Shapiro, dated 8-10-93.

Phillips, D.H., B. Schoket, A. Hewer, and P.L. Grover. 1990. DNA adduct formation in human and mouse skin by mixtures of polycyclic aromatic hydrocarbons. In *International Agency For Research On Cancer: Complex Mixtures and Cancer Risk* (Eds. H. Vainio, M. Sorsa, and A.J. Michael), World Health Organization.

Pittelkow M.R., H.O. Perry, S.A. Muller, W.Z. Maughan, and P.C. O'Brien. 1981. Skin Cancer Patients With Psoriasis Treated With Coal Tar: A 25-Year Follow-up Study. *Archives of Dermatology* 117:465-468.

Pott, P. 1775. *The Chirurgical Works of Percivall Pott, F.R.S. volume V, Part III*. Hawes, W. Clarke and R. Collins, London.

Schmähl, D., K.G. Schmidt, and M. Habs. 1977. Syncarcinogenic Action of Polycyclic Hydrocarbons in Automobile Exhaust Gas Condensates. In *International Agency for Research on Cancer: Air Pollution and Cancer in Man* (Eds: U. Mohr, D. Schmähl, and L. Tomatis), World Health Organization.

Schoket, B., A. Hewer, P.L. Grover, and D.H. Phillips. 1988. Covalent Binding of Components of Coal-Tar, Creosote and Bitumen to the DNA of the Skin and Lungs of Mice Following Topical Application. *Carcinogenesis* 9(7): 1253-1258.

Schoket, B., A. Hewer, P.L. Grover, and D.H. Phillips. 1988b. Formation of DNA Adducts in Human Skin Maintained in Short-Term Organ Culture and Treated with Coal-Tar, Creosote or Bitumen. *International Union Against Cancer* 42: 622-626.

Storm, J.E., S.W. Collier, R.F. Stewart, and R.L. Bronaugh. 1990. Metabolism of Xenobiotics during Percutaneous Penetration: Role of Absorption Rate and Cutaneous Enzyme Activity. *Fundamental and Applied Toxicology* 15: 132-141.

USEPA. 1992. Draft Report: "A Cross-Species Scaling Factor for Carcinogen Risk Assessment Based on Equivalence of mg/kg^{3/4}/Day; Notice." *Federal Register*, June 5.

USEPA. 1996. Proposed Guidelines for Carcinogen Risk Assessment. EPA/600/P-92/003C. Office of Research and Development, US Environmental Protection Agency. Washington, DC.

Van Schooten, F.-J., E.J.C. Moonen, E. Rhijnsburger, B. van Aagen, H.H.W. Thijssen, and J.C.S. Kleinjans. 1994. Dermal Uptake of Polycyclic Aromatic Hydrocarbons After Hairwash and Coal-Tar Shampoo. *The Lancet* 344(8935): 1505-1506.

Warshawsky, D., W. Barkley, and E. Bingham. 1993. Factors Affecting Carcinogenic Potential of Mixtures. *Fundamental and Applied Toxicology* 20: 376-382.

Wright, C.W., D.W. Later, R.A. Pelroy, D.D. Mahlum, and B.W. Wilson. 1985. Comparative Chemical and Biological Analysis of Coal Tar-Based Therapeutic Agents to Other Coal-Derived Materials. *Journal of Applied Toxicology* 5(2): 80-88.

Yamagiwa, K. and K. Ichikawa. 1915. Experimental study of the pathogenesis of carcinoma. *J. Cancer Res.* 3(1): 187-189.

TER

Call 1-800-PICK-UPS (1-800-742-5877) or visit our Web site at www.ups.com

Labels and Space

Place the Tracking Label and your address label

to the left of the address label.

Apply your UPS Shipper Number or place parcel register tape above the address label.

Air Shipping Document

Waybill

ups
UPS Next Day Air
UPS Worldwide Express
 Shipping Document

SHIPMENT FROM
 UPS ACCOUNT NO. **6124EX**

REFERENCE NUMBER
 Docket #49008430SF1999CV0269
 TELEPHONE
 510-622-2100

Susan S. Fiering
 STATE ATTORNEY GENERAL
 1515 CLAY ST FLR 20

OAKLAND CA 94612-1499

DELIVERY TO TELEPHONE

Charles J. Canley
 Food and Drug Administration
 5630 Fishers Lane, Room 1061
 Rockville, MD 20852

0101911202609 2/99 W

United Parcel Service, Louisville, KY

WEIGHT	WEIGHT	DIMENSIONAL WEIGHT

The shipper authorizes UPS to act as forwarding agent for export control and customs purposes.
 The shipper certifies that these commodities, technology or software were shipped from the United States in accordance with the Export Administration Regulations. Diversion contrary to U.S. law is prohibited.

- NEXT DAY AIR
- WORLDWIDE EXPRESS (INTERNATIONAL)
- DOCUMENTS ONLY
- SATURDAY PICKUP
- SATURDAY DELIVERY



1Z 612 4EX 22 1000 332 8

1Z 612 4EX 22 1000 332 8



1Z 612 4EX 22 1000 332 8

BILL RECEIVER

ups **UPS Next Day Air**
 EXTREMELY URGENT

1Z 612 4EX 22 1000 332 8



1Z 612 4EX 22 1000 332 8

TRACKING NUMBER

SHIPMENT ID NUMBER **6124 EX79 XTF**

DATE OF SHIPMENT
1/16/00

EXPORT 1 EXPORT 2 DELIVERY 1 DELIVERY 2
UPS USE FOR INTERNATIONAL SHIPMENTS

DELIVERY 2