Fle VDC/CMA

CHEMICAL MANUFACTURERS ASSOCIATION

September 25, 1990

TO: Vinylidene Chloride Panel

SUBJECTS: (1) EPA's Calculation of Cancer Risk from Exposure to Indoor Air Containing Vinylidene Chloride
(2) ATSDR and Data Needs Report
(3) European Study on Mechanisms of Toxicity

Bill Hayes (Dow) asked that I distribute the enclosed study to the Vinylidene Chloride Panel. There are actually two papers enclosed; however, they are slightly different papers of the same study. The study author is Lance Wallace of EPA and the study concerns the cancer risks associated with exposure to various indoor air pollutants, one of which is vinylidene chloride. Wallace has calculated an excess cancer risk of 320×10^{-6} from exposure to vinylidene chloride. Please review the Wallace study. The Vinylidene Chloride Panel may want to respond to the study and to Wallace in some fashion. Wallace himself notes several study limitations.

I have enclosed a document recently published by the Office of Management and Budget (OMB) which outlines general weaknesses of the risk assessment process. This document may be helpful in identifying risk assessment criticisms which can be used to fault the Wallace study. CMA is beginning to use this OMB document in its advocacy efforts as a "guide" to respond to a variety of environmental and toxicological issues.

I will be talking with Dr. Cebulus of ATSDR in the next day or so about what the status of vinylidene chloride is in terms of ATSDR developing a Data Needs report. If ATSDR has identified vinylidene chloride as a "high priority" chemical for which additional data will have to be developed, then the Panel may want to respond to ATSDR. Bill Hayes also indicated that he would be contacting the Europeans planning to conduct a study on the mechanisms of vinylidene chloride toxicity. He will find out more about the study for the Panel. One issue is whether or not the Panel wants to participate financially in this European study.

If you have any questions, please call me at 202/887-1189.

Sincerely,

Jonathon T. Busch

Jonathon T. Busch
Manager, Vinylidene Chloride Panel

cc: Enclosures (2)

SL 064711

90-159.9

CANCER RISKS FROM 35 ORGANIC CHEMICALS IN AIR AND DRINKING WATER: RESULTS FROM EPA'S TEAM STUDIES OF VOCS AND PESTICIDES

Lance A. Wallace

U.S. Environmental Protection Agency
Warrentown, Virginia

AIR & WASTE MANAGEMENT
ASSOCIATION

SINCE 1907

For Presentation at the
83rd Annual Meeting & Exhibition
Pittsburgh, Pennsylvania
June 24-29, 1990

INTRODUCTION

EPA's TEAM Studies have provided data on the exposures to organic chemicals of about 1000 persons representing about 1,000,000 residents of ten U. S. cities. About 35 of these chemicals cause cancer in animals and may cause cancer in man. In this paper, we calculate the upper-bound lifetime risk associated with airborne and waterborne exposures to each chemical. Although the absolute magnitudes of these upper-bound risks are very uncertain, the relative rankings of these chemicals may be useful in highlighting the more important on s for exposure reduction priorities.

METHODS

The calculation of cancer risk requires two factors: carcinogenic potencies of chemicals and mean exposures of people. Chemical potencies are taken from EPA sources[1,2]. Exposure data for 12 volatile organic compounds (VOCs) in personal air and drinking water are taken from TEAM Studies carried out in eight U.S. cities between 1980 and 1987[3,4]. Exposure data for 23 carcinogenic pesticides in air and drinking water were collected in two cities between 1986 and 1988[5]. Outdoor air measurements were made for all chemicals in the back yards of the residents; therefore an estimate can be made of the relative contribution of outdoor and indoor air to total airborne exposure to all the target VOCs and pesticides.

In a previous study of cancer risks of six prevalent VOCs[6], the TEAM cities were divided into "metropolitan" and "non-metropolitan" categories. The mean exposures calculated for ach city were averaged to provide a risk associated with each of th two categories. Assuming that the TEAM cities represented typical values, U.S. Census figures were employed to calculate a risk for the U.S. population. The results from that study have been reproduced here, with two additions: calculated risks fr m exposures to benzene during smoking and chloroform during showering. (Both of these exposures could not be measured using the personal monitors employed in the TEAM Studies; however, they could be estimated using breath measurements for smokers[7] and models for exposure during showers[8]).

An additional six VOCs have been added. Two of these (styrene and 1,1,1-trichloroethane) are prevalent and their concentrations are well characterized, but their carcinogenicity is in doubt. A third chemical (methylene chloride) is probably prevalent but very few measurements of personal exposure have been made due to its high volatility. The remaining three chemicals (vinylidene chloride, 1,2-dichloroethane, and 1,2-dibromoethane) are well-established animal carcinogens but are much less prevalent. They have been measured in only a few percent of the personal and outdoor air samples collected in the TEAM Studies. Thus the risk estimates for these six VOCs are more uncertain than the estimates for the original six VOCs.

The pesticide exposures are the unweighted mean of the seasonally-averaged values for each city. Since Jacksonville, FL was chosen as a high-use area and Springfield, MA as a low-use area, the average of the two may represent a closer approach to actual mean exposures than either one separately.

Risks are calculated as a simple multiple of the exposur and the potency. If the potency is given in units of $[mg/kg\text{-}day]^{-1}$, the exposure is translated to a daily dose by assuming 20 m^3 inspired air per day and a body weight of 70 kg.

RESULTS

The mean measured exposures, carcinogenic potencies and calculated upper-bound cancer risks are displayed for 12 VOCs in Table I and for 23 pesticides in Table II. The percent of total airborne exposure due to indoor sources is also provided.

Six VOCs and four pesticides exceeded the de minimus or negligible risk level of 10^{-6} by a factor of 10 or more. The six VOCs were benzene, vinylidene chloride, p-dichlorobenzene, chloroform, methylene chloride, and carbon tetrachloride. The four pesticides were chlordane, heptachlor, aldrin and dieldrin. All four have been banned by EPA. Despite the bans, exposures remain high, due perhaps to their long life in the soil and their movement into homes after being injected in the soil as termiticides. Carbon tetrachloride has also been banned from consumer products, but its long life in the atmosphere has led t a global background that is sufficiently high to result in a nonnegligible risk.

Three additional VOCs and four additional pesticides are at or above the 10^{-6} risk level, but by less than an order of magnitude. The three VOCs are tetrachloroethylene, trichloroethylene, and ethylene dichloride. The four pesticides (or degradation products) are dichlorvos, α-BHC. γ-BHC (lindane), and heptachlor epoxide.

Two additional VOCs and four additional pesticides had upper-bound risks between 10^{-6} and 10^{-7}: styrene, 1,1,1-trichloroethane, hexachlorobenzene, propoxur, DDT, and DDE.

Nine chemicals (all pesticides) were below the 10^{-7} risk level for airborne exposure.

Finally, one VOC and one pesticide were seldom or never detected. However, detection limits were high for ethylene dibromide and for pentachlorophenol, leaving open the possibility that these chemicals could represent non-negligible lifetime risks of cancer. Additional exposure studies with lower detection limits for these chemicals are necessary before a more trustworthy estimate of their risks is possible.

Indoor sources accounted for the great majority (80-100%) of the total airborne risk associated with most of these chemicals. Carbon tetrachloride is the only one of the target chemicals for which outdoor sources account for a majority of the airborne risk, indicating the effectiveness of the ban on its use in consumer products.

DISCUSSION

Uncertainty of Estimates

Great uncertainty accompanies most risk estimates. Th major uncertainties involved in potency calculations are well known: the extrapolation from animals to man and from high dose to low dose. These uncertainties are such that a given chemical may not cause human cancer at all--the actual cancer risk may be exactly zero. Even if the risk is not zero, the estimates could easily be wrong by factors of 10, 100, or more, depending on th shape of the dose-response curve, the possible existence of a threshold due to DNA repair or other mechanisms, and many other factors.

Considerably less uncertainty is associated with some of the exposure estimates. The overall mean VOC exposure in eight cities was usually withi a factor of 3 of the extremes for an individual city, whether the city was rural, suburban, urban, or heavily industrialized. The reason for this predictability appears to be the relative importance of consumer products, personal activities, and building materials to human exposure; such factors do not vary greatly across the country. Of c urse, personal activities can result in very high exposures for sh rt periods, but we are concerned here with long-term exposures. Somewhat more variation was noted for pesticides, with differences of a factor of 10 in exposure noted for a number f pesticides in Jacksonville and Springfield.

For the more prevalent VOCs and pesticides, relatively little error is associated with the estimates of the relative contribution of indoor and outdoor sources. The reason is that the same instrumentation was used to measure both indoor and outdoor air--even if the instruments were biased, the relative proportions would remain nearly unchanged. For the more rarely found chemicals (e.g., vinylidene chloride, ethylene dichloride), the indoor/outdoor ratio is less certain.

Only one of these chemicals is considered a human carcinogen: benzene. Therefore the risk estimate associated with benzene is on more solid ground than any of the others. Benzene is also the only one of these chemicals with human epidemiological studies showing a possible influence of environmental levels of exposure on cancer risk: two studies show that children of smokers die of leukemia at two or more times the rate of children of nonsmokers[9,10]. The higher mortality rate is consistent with the measured elevated levels of benzene in the breath of smokers (suggesting exposure of the fetus in the womb of the pregnant smoker). Elevated levels of benzene in th air of homes have also been documented by the TEAM Study and by a study in West Germany[11]; however, the increase (on the order f 50% in both studies) does not seem enough to explain the increase in the mortality rate unless children are more susceptible to benzene-induced leukemia at some point in the first 8-9 years of life.

The case of vinylidene chloride requires further comment. This chemical is highly volatile and therefore "breaks thr ugh" the Tenax monitor after only a portion of the monitoring period. The concentration is calculated on the basis of the "breakthrough volume" rather than the actual sampling volume, and depending on the pattern of exposure during the monitoring period, may be either an over- or under-estimate of the actual concentration. Also, because the sampling volume of 20 liters is effectively reduced to the breakthrough volume of only a few liters, the sensitivity is reduced by the same factor. The limits of detection for vinylidene chloride ranged from 3 to 14 $\mu g/m^3$, about an order of magnitude worse than for most of the other target VOCs. Out of 1085 personal air samples collected from 355 New Jersey residents over three different seasons, only 77 (7%) had measurable concentrations of vinylidene chloride. (An ther 107 (10%) showed trace concentrations.)

The population risk for such rarely detected chemicals can be calculated, but the interpretation of the risk presents difficulties. For example, the single highest measured exposure to VDC was 120,000 $\mu g/m^3$, which was incurred by a cabinetmaker. (The second highest value of 14,000 $\mu g/m^3$ was also measured for this same person in a different season.) Taken together, these two values accounted for more than 80% of the total calculated exposure (and therefore the risk) for the population. If we include these values in our calculation of risk, then vinylidene chloride exposures average 150 $\mu g/m^3$, and the upper-bound risk is 7.5×10^{-3}--greater than the risks from radon and passive smoking combined! If we drop these two values, the population exposure decreases to 28 $\mu g/m^3$, and the risk to 1.4 X 10^{-3}--still very large. However, four other exposures exceeded 1000 $\mu g/m^3$. If these values are also dropped from the risk calculation, the average exposure decreases to 6.5 $\mu g/m^3$, and the associated upper-bound risk to 3.2 X 10^{-4}. Thus the population average of 150 $\mu g/m^3$ for all 355 persons is actually composed of an average of 6.5 $\mu g/m^3$ for about 350 persons, and an average of about 30,000 $\mu g/m^3$ for about 5 persons. This corresponds to a difference in risk of the two groups of a factor of 5000!

Because of the tremendous effect of a few measurements, the calculated upper-bound risk for vinylidene chloride must be considered tentative. Only additional data on personal exposures (collected by methods with a sensitivity of 1 $\mu g/m^3$ or less) will provide the information necessary for an adequate risk assessment.

Exposures Through Other Routes

Exposures through routes other than air and water may be possible for some of the chemicals. All of the above chemicals were measured in drinking water, and found to present less than 1% of the risk due to airborne exposures with a single exception: chloroform. All of the VOCs were also measured in food and beverages; again chloroform was the only VOC found in significant amounts in food. Many of the pesticides have been measured in food by the FDA for years; however, exposures in food account for only a small proportion of total exposure to the four pesticides of highest risk through airborne routes. Food exposures outweigh air exposures for some of the other pesticides (e.g, Captan).

House dust may provide an important reservoir for any or all of the pesticides, and possibly also for the least volatile of the VOCs (p-dichlorobenzene, tetrachloroethylene). Ingestion of the dust by toddlers could be an important additional source of risk.

Comparison with Other Environmental Risks

Several organic chemicals of interest were not monitored in the TEAM Studies. Among these are formaldehyde and 1,3-butadiene. Risk estimates for these chemicals may be compared to the risks calculated above.

The carcinogenicity of formaldehyde is controversial, due to the unusual metabolic pathway associated with its carcinogenicity in rodents. Estimates for the cancer risk of formaldehyde[12] range over extreme limits, from zero to 10^{-3}. Employing an intermediate potency factor (unit risk of 1.3 X 10^{-5} $(\mu g/m^3)^{-1}$) and measured values of 40 $\mu g/m^3$ for normal (non-mobile home) housing stock results in a risk of about 5 X 10^{-4}.

Recent animal studies of 1,3-butadiene have resulted in revising its potency upward by nearly three orders of magnitude. Although no personal exposure data are available for this chemical, a recent study has measured the level in sidestream smoke at about 400 µg per cigarette[13]. This is approximately the level of benzene in sidestream smoke (330 µg per cigarette); therefore, if 1,3-butadiene is not too reactive, we can calculate that it will be elevated by about 4 $\mu g/m^3$ in smoking homes and by 13 $\mu g/m^3$ in workplaces allowing smoking. Using the revised unit risk value of 2.8 X 10^{-4}, and assuming 38 million homes with smokers averaging three residents each, 75 million workers in workplaces allowing smoking, and 26 million non-workers exposed to cigarette smoke, we arrive at an upper-bound risk associated with 1,3-butadiene of 6 X 10^{-4}.

Thus the individual risks for formaldehyde and 1,3-butadiene are greater than the airborne risk of any of the other 35 VOCs and pesticides considered in this report. However, the great uncertainty in the carcinogenic potency of formaldehyde, including the uncertainty as to whether it is a human carcinogen at all, and the lack of exposure data for 1,3-butadiene make the risk estimates for these two chemicals particularly speculative.

The combined upper-bound risk of about 10^{-3} associated with these 37 predominantly indoor organic chemicals appears to be similar to the risks associated with the most severe environmental hazards (radon and passive smoking). For example, the risk associated with nonsmokers' exposure to radon has been estimated to be about 10^{-3} and that with passive smoking has been estimated[14] at 2 X 10^{-3}. It should be noted, however, that the risk estimates for radon and passive smoking are based on human epidemiology studies, and are therefore on firmer ground than all of the risk estimates for the organic chemicals with the exception of benzene.

The risk estimates for these organic chemicals are considerably higher than the risks associated with some EPA regulations (NESHAPS chemicals and Superfund clean-up criteria). Similar conclusions regarding the importance of indoor air pollution compared to other environmental hazards have been reached by EPA Headquarters[15] and by three Regions[16]; both of these reports rank indoor air pollution as among the top two or three environmental threats to public health.

Two other risk estimates for personal exposure to VOCs hav been published[12,17]. McCann arrived at similar risk estimates for most of the chemicals; Tancrede estimated 5-10 times higher risks, due partly to using a different method for calculating potencies from animal data and partly to considering explicitly several additional sources of uncertainty. No previous risk estimates for most of these pesticides have been possible, du to the lack of exposure information.

Actions to avoid these risks may be taken by individuals. Since the sources of the risks are often personal activities (smoking, using air fresheners), these activities can be halted or modified. (For example, smokers could establish a room in the home with separate ventilation.) Exposures from chloroform could be reduced by drinking bottled water or using an activated carbon filter on the water supply. Exposures from petroleum-based products could be reduced by discarding or storing used paint cans and sprays in a detached garage or tool shed. Dry-clean d clothes could be hung outdoors for a day-- ne study indicates that 20-30% of tetrac..loroethy ene residues on the clothes will outgas during the first day.

The reason for the large number of pesticides observed in indoor air in the latest TEAM Study is not well understood. Termiticides, like radon gas, may be entering the basement du t soil gas movement; it may be that the same techniques to contr l radon (sealing the foundation, providing separate ducting at the entrance points) may also control termiticide entry. Other pesticides, particularly the long-lived chlorinated hydrocarb ns such as DDT, may be entering the home by being tracked in on people's shoes. If so, removing shoes before entering the home, and reducing or eliminating the use of carpets or rugs (which collect large amounts of dust containing pesticides and metals as well), should reduce pesticide exposures.

SUMMARY AND CONCLUSIONS

Measured personal exposures to 12 VOCs and 23 pesticides in EPA's TEAM Studies have been used to arrive at upper-bound lifetime cancer risk estimates. Seven VOCs and seven pesticides have upper-bound risks ranging from 10^{-6} to 10^{-4}. The combined upper-bound risk of about 10^{-3} from these organic indoor air pollutants is nearly comparable to the estimates of risk from radon and environmental tobacco smoke. (However, the latter two estimates are based on human epidemiology studies, and are therefore subject to far less uncertainty.) These upper-b und risks are much greater than the health risks associated with most other environmental problems.

Despite the recognized large uncertainty in these risk estimates, these findings provide additional support for the conclusion of two recent comparative rankings of environmental risk by EPA: that indoor air pollution is one of the greatest threats to public health of all environmental problems.

REFERENCES

1. Integrated Risk Information System (IRIS), US EPA, Washington, DC.

2. Engler, R., memorandum to Office of Pesticide Programs Division Directors, USEPA, October 27, 1989.

3. Wallace, Lance A., Pellizzari, E.D., Hartwell, T.D., Sparacino, C., Whitmore, R., Sheldon, L., Zelon, H. and Perritt, R., "The TEAM Study: personal exposures to toxic substances in air, drinking water, and breath of 400 residents of New Jersey, North Carolina, and North Dakota" Environmental Research 43: 290-307 (1987).

4. Wallace, L. A., Pellizzari, E.D., Hartwell, T.D., Whitmore, R., Zelon, H., Perritt, R. and Sheldon, L. "The California TEAM Study: breath concentrations and personal exposures to 26 volatile compounds in air and drinking water of 188 residents of Los Angeles, Antioch, and Pittsburg, CA." Atmos. Environ. 22: 2141-2163 (1988).

5. Immerman, F.W. and Schaum, J.L. Final Report of the Nonoccupational Pesticide Exposure Study (NOPES), US Environmental Protection Agency, Research Triangle Park, NC (1990).

6. Wallace, Lance A. "Cancer risks from organic chemicals in th home", in Environmental Risk Assessment: Is Analysis Useful?, APCA Specialty Conference Proceedings, Air Pollution Control Assoc., Pittsburg, PA. SP-55 (1986).

7. Wallace, Lance A., Pellizzari, E., Hartwell, T., Perritt, K., and Ziegenfus, R., "Exposures to benzene and other volatile organic compounds from active and passive smoking," Archives of Environmental Health 42: 272-279 (1987).

8. McKone, T.E. "Human exposure to VOCs in household tap water: The indoor inhalation pathway", Env. Sci. Tech. 21: 1194-1201 (1987).

9. Sandler, D.P., Everson, R.B., Wilcox, A.J., and Browder, J.P. "Cancer risk in adulthood from early life exposure to parents' smoking". Am. J. Public Health, 75:467 (1985).

10. Stjernfeldt, M., Berglund, K., Lindsten, J., and Ludvigsson, J. "Maternal smoking during pregnancy and risk of childhood cancer", Lancet, June 14, 1986:1350-52 (1986).

11. Krause, C., Mailahn, W., Nagel, R., Schulz, C., Seifert B., Ullrich D. "Occurrence of volatile organic compounds in the air of 500 Homes in the Federal Republic of Germany", in Proceedings of The 4th International Conference on Indoor Air Quality and Climate, Vol. 1, pp. 102-106. Institute for Soil, Water, and Air Hygiene. Berlin (West) (1987).

12. McCann, J. Horn, L., Girman, J. and Nero, A.V. Potential risks from exposure to organic compounds in indoor air. Lawrence Berkeley Laboratory, report # LBL-22473. Berkeley, CA (1986).

13. Lofroth, G. Burton, B., Forehand, L., Hammond, S.K., Seila, R., Zweidinger, R., and Lewtas, J., "Characterization of environmental tobacco smoke," Env. Sci. and Tech. : (1989).

14. Repace, J., "A quantitative estimate of nonsmokers' lung cancer risk from passive smoking," Env. Int. 11:3-22 (1985).

15. USEPA, Unfinished Business: A Comparative Assessment of Environmental Problems, Washington DC (1987).

16. USEPA, Comparing Risks and Setting Environmental Priorities: Overview of Three Regional Projects, Office of Policy, Planning, and Evaluation, Washington, DC (1989).

17. Tancrede, M, Wilson, R. Zeise, L. and Crouch, E.A.C., "The carcinogenic risk of organic vapors indoors: a theoretical survey," Atmos. Env. 21: 2187-2205 (1987).

Table I. Upper-Bound Lifetime Cancer Risks of 12 VOCs Measured in the TEAM Studies (1980-87)

Chemical	Exposure[a] ($\mu g/m^3$)	Potency $(\mu g/m^3)^{-1}$ ($X\ 10^{-6}$)	Risk ($X10^{-6}$)	Indoor Air Contribution[b] (percent)
Benzene				
Air	15	8	120[c]	60
Smokers	90	8	720[c]	--
Vinylidene chloride	6.5[d]	50	320	86
Chloroform				
Air	3	23	70	80
Showers (Inhalation)	2	23	50	100
Water	30[e]	2.3[e]	70	--
Food & Beverages	30[e]	2.3[e]	70	--
p-Dichlorobenzene	22	4	90	98
1,2-Dibromoethane	0.05	510	25	40
Methylene chloride	6[f]	4	24	60
Carbon tetrachloride	1	15	15	20
Tetrachloroethylene	15	0.6	9	80
Trichloroethylene	7	1.3	9	80
Styrene				
Air	1	0.3[g]	0.3	70
Smokers	6	0.3	2	--
1,2-Dichloroethane	0.5	7	4	60
1,1,1-Trichlorethane	30	0.003	0.1	70

---

[a] Arithmetic means based on 24-hour average exposures of ≈ 750 persons in six urban areas measured in the TEAM Studies

[b] Based on backyard measurements in 175 homes in six urban ar as

[c] The risk estimates for benzene are based on human epidemiology and are therefore mean as opposed to upper-bound estimates

[d] Six measurements exceeding 1000 $\mu g/m^3$ dropped from the calculation; inclusion of the measurements leads to an average exposure of 150 $\mu g/m^3$.

[e] These figures are in $\mu g/L$ or ppb rather than $\mu g/m^3$

[f] Based on only eight 24-hour indoor measurements in 1987.

[g] Source: US EPA (1983) Review and Evaluation of Evidence for Cancer Associated with Air Pollution, EPA 450/5-83-006.

Table II. Upper-bound Lifetime Cancer Risks from Airborne Exposures to 23 Pesticides Measured in the NOPES TEAM Study

Pesticide	Exposure[a] (ng/m^3)	Potency (kg-d/mg)	Risk (X10^{-6})	Indoor Air Contribution[b] (percent)
Banned Termiticides				
Heptachlor	71	4.5	90 (19)[c]	90
Chlordane	198	1.3	70 (15)	94
Aldrin	13	17	60 (13)	100
Dieldrin	3	16	14 (3)	92
Heptachlor Epoxide	0.4	9.1	1 (0.2)	80
DDE	2.2	0.34	0.2 (0.04)	100
DDT	0.7	0.34	0.1 (0.02)	88
Other Pesticides				
Dichlorvos	33	0.29	2.7	100
γ-BHC (Lindane)	6.6	1.3	2.5	100
α-BHC	0.5	6.3	1	100
Propoxur	100	0.0079	0.2	98
Hexachlorobenzene	0.3	1.67	0.1	90
Dicofol	2.6	0.34	0.05	100
o-Phenylphenol	58	0.0016	0.02	99
2,4-D	0.6	0.019	0.003	77
Atrazine	0.05	0.22	0.003	100
cis-Permethrin	0.4	0.022	0.003	100
trans-Permethrin	0.1	0.022	0.001	100
Chlorothalonil	0.7	0.011	0.002	28
Folpet	0.5	0.0035	0.0005	50
Captan	0.1	0.0023	0.00007	100
DDD	<4	0.34	<0.4	---
Pentachlorophenol	<730	0.013	<3	---

[a] Arithmetic mean of population-weighted and seasonally-weighted average personal exposures measured for 173 persons in Jacksonville, FL and 85 persons in Springfield/Chicopee, MA.

[b] Percent of total airborne exposure only, based on outdoor measurements at each home in the two cities.

[c] All risks calculated assuming 70-year lifetime exposure at the measured levels. For banned pesticides, whose environmental concentrations should decrease over time, an alternative calculation of risk (in parentheses) assuming a 10-year half-life in soil is provided.

Comparison of Risks from Outdoor and Indoor Exposure to Toxic Chemicals

Lance A. Wallace
Atmospheric Research and
Exposure Assessment Laboratory
U.S. Environmental Protection Agency
Bldg 166, Bicher Road
Vint Hill Farms Station
Warrenton, VA 22186-5129

For presentation at
EPA-ILSI Joint Symposium on
Methodology for Assessing Health Risks
from Complex Mixtures in Indoor Air
Arlington, VA
April 17-19, 1990

Proceedings Not Yet
Published
9/4/90

INTRODUCTION

For the last decade, EPA's TEAM Studies have been providing data on personal exposures (including indoor and outdoor concentrations) to organic chemicals for more than 1000 persons representing more than 1,000,000 residents of ten U. S. cities. About 35 of these chemicals cause cancer in animals and may cause cancer in man. In this paper, we calculate the upper-bound lifetime risk associated with airborne exposures to each chemical. We also try to apportion the risk between indoor and outdoor sources. Although the absolute magnitudes of these upper-bound risks are very uncertain, the relative rankings of the chemicals and their sources may be useful in focusing our attention on efficient ways to reduce exposure.

METHODS

The calculation of ca risk requir two factors: carcinogenic potencies of icals and me exposures of people. Chemical potencies re taken from EPA sources[1,2]. Exposure measurements (including some overnight indoor air measurements) for 12 volatile organic compounds (VOCs) are taken from TEAM Studies carried out in eight U.S. cities between 1980 and 1987[3,4]. Personal exposures and indoor air concentrations for 23 carcinogenic pesticides were measured in two cities between 1986 and 1988[5]. Outdoor air measurements were made for all chemicals in the back yards of the residents; therefore an estimate can be made of the relative contribution of outdoor and indoor air to total airborne exposure to all the target VOCs and pesticides.

In a previous study of cancer risks of six prevalent VOCs[6], the TEAM cities were divided into "metropolitan" and "non-metropolitan" categories. The mean exposures calculated for each city were averaged to provide a risk associ ed with each of the two categories. Assuming that the TEAM cit es represented typical values, U.S. Census figures were employed to calculate a risk for the U.S. population. The results from that study have been reproduced here, with two additions: calculated risks fr m exposures to benzene during smoking and chloroform during showering. (Both of these exposures could not be measured using the personal monitors employed in the TEAM Studies; however, they could be estimated using breath measurements for smokers[7] and models for exposure during showers[8]).

An additional six VOCs have been added. Two of these (styrene and 1,1,1-trichloroethane) are prevalent and their personal exposures and outdoor air concentrations are well characterized, but their carcinogenicity is in doubt. A third chemical (methylene chloride) is probably prevalent but very few measurements of personal exposure have been made due to its high volatility. The remaining three chemicals (vinylidene chloride, 1,2-dichloroethane, and 1,2-dibromo thane) are well-established animal carcinogens but are much less prevalent. They have been measured in only a few percent of the personal and outdoor air samples collected in the TEAM Studies. Thus the risk estimates for these six VOCs are more uncertain than the estimates for the original six VOCs.

The pesticide exposures are the unweighted means of the seasonally-averaged values for each city. Since Jacksonville, FL was chosen as a high-use area and Springfield, MA as a low-use area, the average of the two may represent a closer approach to actual mean exposures than either one separately.

Risks are calculated as a simple multiple of the exposur and the potency. If the potency is given in units of [mg/kg-day]$^{-1}$, the exposure is translated to a daily dose by assuming 20 m^3 inspired air per day and a body weight of 70 kg.

RESULTS

The mean measured exposures, outdoor air concentrations, carcinogenic potencies and calculated upper-bound cancer risks are displayed for 12 VOCs in Table I and for 23 pesticides in Table II.

Seven VOCs and four pesticides exceeded the de minimus or negligible risk level of 10^{-6} by a factor of 10 or more. The seven VOCs were benzene, vinylidene chloride, *p*-dichlorobenzene, chloroform, ethylene dibromide, methylene chloride, and carbon tetrachloride. The four pesticides were chlordane, heptachlor, aldrin and dieldrin. All four have been banned by EPA. Despite the bans, exposures remain high, due perhaps to their long life in the soil and their movement into homes after being injected in the soil as termiticides. Carbon tetrachloride has also been banned from consumer products, but its long life in the atmosphere has led to a global background that is sufficiently high to result in a nonnegligible risk.

Three additional VOCs and four additional pesticides are at or above the 10^{-6} risk level, but by less than an order of magnitude. The three VOCs are tetrachloroethylene, trichloroethylene, and ethylene dichloride. The four pesticides (or degradation products) are dichlorvos, α-BHC. γ-BHC (lindane), and heptachlor epoxide.

Two additional VOCs and four additional pesticides had upper-bound risks between 10^{-6} and 10^{-7}: styrene, 1,1,1-trichloroethane, hexachlorobenzene, propoxur, DDT, and DDE.

Nine chemicals (all pesticides) were below the 10^{-7} risk level for airborne exposure.

Finally, one pesticide--pentachlorophenol (PCP)--was never detected. However, the detection limit was very high for PCP, leaving open the possibility that this pesticide could represent non-negligible lifetime risks of cancer. Additional exposure studies with lower detection limits for pentachlorophenol, and also for some of the less prevalent VOCs such as ethylene dibromide and vinylidene chloride, are necessary before a more trustworthy estimate of their risks is possible.

Indoor sources accounted for the great majority (80-100%) of the total airborne risk associated with most of these chemicals. Carbon tetrachloride is the only one of the target chemicals for which outdoor sources account for a majority of the airborne risk, indicating the effectiveness of the ban on its use in consumer products.

DISCUSSION

Upper-bound vs. "best-guess" estimates of potency

The potencies employed in the risk calculations above are, with one exception (benzene), "upper-bound" potencies calculated from animal experiments. This raises the question of what the "best-guess" potency might be. Unfortunately, the EPA has chosen not to calculate "best" estimates of potency, arguing that such estimates are inherently more unstable than the upper-bound estimates. Without arguing this point, it may still be instructive to calculate "best" estimates of potency. For example, are such "best" estimates lower by a factor of 3, 10, or 100? Could the "best" estimate sometimes be zero? To answer such questions, we may examine the approach developed by a gr up at Harvard University[9]. In this approach, a median potency is calculated from the animal studies, together with an estimate of the natural logarithm of the geometric standard deviation (σ_x) of that animal potency. A median potency in humans is then calculated, based on an assumed uncertainty connected with extrapolation from animal to man ($\sigma_y = 1.5$), and an uncertainty (if necessary) associated with converting from oral dose to inhalation dose ($\sigma_t = 1.6$). Assuming log-normal distributions for all the sources of uncertainty, estimates of mean and upper-bound potencies can be calculated using standard relationships of log-normal distributions:

$$\text{Mean} = \text{Median} \times \exp(\sigma^2/2)$$

$$\text{where } \sigma^2 = \sigma_x^2 + \sigma_y^2 + \sigma_t^2$$

When these calculations are carried out, using values for the median potencies and their uncertainties as calculated by the Harvard group, we find that 95% upper-bound potencies are typically about 7 times mean potencies. This finding is in general agreement with the results of a study[10] that compared upper-bound potency estimates from animal studies with observed potencies from human epidemiology of about 20 chemicals that are both animal and human carcinogens. That study found that, in general, the upper-bound estimates from the animal data were about an order of magnitude higher than the "best" estimates from the human data.

Therefore, the upper-bound risk estimates in Tables I and II (except for benzene) may be divided by a factor of 7 or 10 to provide "best" estimates of risk based on mean potencies and mean exposures. This results in benzene emerging as the single chemical with the highest risk of all 35 considered.

For some of these chemicals, it is also possible to argue that the "best" estimate of risk is 0. For example, if a chemical causes cancer in animals but not in man, its carcinogenic risk in humans is zero by definition. Now suppose that a chemical is "more likely than not" to be a noncarcinogen--what is the "best" estimate of its carcinogenic risk? One way to answer this question is to say that there is a better than 50% probability that it is not a carcinogen, and therefore the "best" (median) estimate of risk is zero. Another approach might be to assign a probability that it is a carcinogen, calculate the risk on the basis of the animal studies, and then "dilute" that risk by multiplying by the assigned probability. Thus, if the animal studies are ambiguous, and we assign only a 10% probability that the chemical is a human carcinogen, the risk calculation would be multiplied by 0.10 to arrive at a "best" estimate of risk. In the case of the chemicals considered here, most are classified as "B2" (probable human carcinogens), but some are classified as "C" (possible human carcinogens). (Still others, such as tetrachloroethylene and p-dichlorobenzene, wobble back and forth between the two classifications.) If the "probables" were assigned a likelihood factor greater than 50%, while the "possibles" were assigned a factor less than 50%, one of the above two approaches could be employed to further adjust the risks calculated in Tables I and II.

Assigning risk to population subgroups

The risk calculations above have generally been made on the basis of the entire population studied in the TEAM Studies, or on extrapolating those results to the U.S. population. One exception has been the risks to smokers of benzene and styrene, which apply only to the 50 million active smokers in the U.S. However, if it were possible to identify sources of exposure, and characterize population subgroups on the basis of their exposure to those sources, it would be possible to refine our estimates of risk. In particular, we would find that risks are higher among the exposed subgroups, and lower (perhaps zero) among the less exposed or unexposed subgroups.

As an exampl of the above points, we may consider the pesticide dichlorvos. This pesticide was found in about a third of Jacksonville homes and only 2% of Springfield homes. Outdoor concentrations were negligible in both regions. If the personal exposures came mostly from use of a consumer product containing dichlorvos, it seems reasonable to calculate a risk based on exposures to the users only. Thus the calculated risk to Jacksonville users would be three times the risk averaged over the entire population, and the calculated risk to Springfield users would be 50 times the risk averaged over the population. Since the calculated risk for the Jacksonville population was 5×10^{-6}, the risk to users would be 15×10^{-6}; the calculated risk for the Springfield population was 3×10^{-7}, leading to a risk to users of 15×10^{-6}--identical to the risk to Jacksonville users. This calculation has not changed the total population risk in either case--it has simply apportioned the risk across the populations. Thus we have gained a sharper definition of the risk. It is also interesting to note that the risks calculated in this way show that in both areas, upper-bound risks to users exceed the one in a million level by a factor of 15; the previous calculation indicated that the upper-bound risks averaged across the population were very close to this dividing line. This approach does not appreciably change the risks calculated for the more prevalent VOCs and pesticides, but has the potential for order-of-magnitude increases in calculated risks for the less prevalent chemicals.

For example, applying this approach to other pesticides, we find that α-BHC, which was found in only 27% of Jacksonville homes and only 2% of Springfield homes, has a calculated average risk in these homes of 8×10^{-6} and 20×10^{-6}, respectively, compared to the value of only 1×10^{-6} averaged across the population. Other pesticides whose lifetime upper-bound risks in exposed homes approached or exceeded one in a million included 2,4-D, DDE, hexachlorobenzene, and dicofol.

Uncertainty of Estimates

Great uncertainty accompanies most risk estimates. The major uncertainties involved in potency calculations are well known: the extrapolation from animals to man and from high d se to low dose. These uncertainties are such that a given chemical may not cause human cancer at all--the actual cancer risk may be exactly zero. Even if the risk is not zero, the estimates could easily be wrong by factors of 10, 100, or more, depending on the shape of the dose-response curve, the possible existence of a threshold due to DNA repair or other mechanisms, and many other factors.

Considerably less uncertainty is associated with some f the exposure estimates. The overall mean VOC exposure in eight cities was usually within a factor of 3 of the extremes for an individual city, wh ther the city was rural, suburban, urban, or heavily industrialized. The reason for this predictability appears to be the relative importance of consumer products, personal activities, and building materials to human exposure; such factors do not vary greatly across the country. Of course, personal activities can result in very high exposures for short periods, but we are concerned here with long-term exposures. Somewhat more variation was noted for pesticides, with differences of a factor of 10 in exposure noted for a number of pesticides in Jacksonville and Springfield.

For the nine prevalent VOCs, about 2000 measurements have been made of 12-hour average personal exposure, and more than 500 have been made of outdoor concentrations. Thus, for these more prevalent VOCs and pesticides, relatively little error is associated with the estimates of the relative contribution of indoor and outdoor sources. The reason is that the same instrumentation was used to measure both indoor and outdoor air--even if the instruments were biased, the relative proportions would remain nearly unchanged.

However, two other VOCs were measurable in only a small percentage of samples: vinylidene chloride (7%) and ethylene dibromide (2%). (The 12th VOC, ethylene dichloride, was measured in about 20% of the samples, but with most measurements hovering close to the detection limit.) For these more rarely found chemicals, the indoor/outdoor ratios are less certain, as are the risk estimates.

Some chemicals were prevalent but do not have sufficient animal studies to establish their carcinogenicity. Among these are toluene and xylenes, not found to be carcinogenic in two-year rat and mouse studies conducted by the National Toxicology Program (NTP), but found to be carcinogenic in natural-lifetime rat and mouse studies carried out in Italy (Maltoni, personal communication). Limonene (used in lemon-scented products and also as a food additive) was the single VOC at the highest average concentrations in people's homes; a recent 2-year NTP study found clear evidence of carcinogenicity in one sex-speci s combination, but no evidence in the other three sex-species combinations. Two pesticides that were prevalent at relatively high concentrations indoors were chlorpyrifos (Dursban) and diazinon. Health studies of these pesticides do not completely rule out their possible carcinogenicity. Because of the prevalence and high concentrations of these VOCs and pesticides, further health studies are indicated.

Several of the chemicals with the highest associated risks will be discussed separately.

Benzene

Only one of these chemicals is considered a human carcinog n: benzen . Therefore the risk stimate associated with benzene is on more solid ground than any of the others. Benzene is also the only one of these chemicals with human epidemiological studies showing a possible influence of environmental levels of exposure on cancer risk: two studies show that children of smokers die of leukemia at two or more times the rate of children of nonsmokers[11,12]. The higher mortality rate is consistent with the measured elevated levels of benzene in the breath of smokers (suggesting exposure of the fetus in the womb of the pregnant smoker). Elevated levels of benzene in the air of homes have also been documented by the TEAM Study and by a study in West Germany[13]; however, the increase (on the order of 50% in both studies) does not seem enough to explain the incr ase in the mortality rate unless children are more susceptible to benzene-induced leukemia at some point in the first 8-9 years of life.

Major sources of exposure to benzene appear to be activ and passive smoking, driving and other personal activities associated with automobiles, use of attached garages for parking cars, storing gasoline and kerosene, and the use of certain consum r products (marking pens, paints, glues, rubber products). The major outdoor source is auto exhaust; emissions from stationary sources account for only a few percent of nationwide exposures.

Vinylidene chloride

This chemical is highly volatile and therefore "breaks through" the Tenax monitor after only a portion of the monitoring period. The concentration is calculated on the basis of the "breakthrough volume" rather than the actual sampling volume, and depending on the pattern of exposure during the monitoring period, may be either an over- or under-estimate of the actual concentration. Also, because the sampling volume of 20 liters is effectively reduced to the breakthrough volume of only a few liters, the sensitivity is reduced by the same factor. The limits of detection for vinylidene chloride ranged from 3 to 14 $\mu g/m^3$, about an order of magnitude worse than for most of the other target VOCs. Out of 1085 personal air samples collected from 355 New Jersey residents over three different seasons, only 77 (7%) had measurable concentrations of vinylidene chloride. (Another 107 (10%) showed trace concentrations.)

The population risk for such rarely detected chemicals can be calculated, but the interpretation of the risk presents difficulties. For example, the single highest measured exposure to VDC was 120,000 $\mathrm{ɟ}/m^3$, which was incurred by a cabinetmaker (The second highe alue of 14,000 $\mu g/m^3$ was also measured fc this same person different season.) Taken together, the two value accoun for more than 80% of the total calculat exposure and the re the risk) for the population. If we

SL 064731

include these values in our calculation f risk, then vinylidene chloride exposures average 150 $\mu g/m^3$, and the upper-bound risk is 7.5×10^{-3}--greater than the risks from radon and passive smoking combined! If we drop these two values, the population exposure decreases to 28 $\mu g/m^3$, and the risk to 1.4×10^{-3}--still very large. However, four other exposures exceeded 1000 $\mu g/m^3$. If these values are also dropped from the risk calculation, the average exposure decreases to 6.5 $\mu g/m^3$, and the associated upper-bound risk to 3.2×10^{-4}. Thus the population average of 150 $\mu g/m^3$ for all 355 persons is actually composed of an average of 6.5 $\mu g/m^3$ for about 350 persons, and an average of about 30,000 $\mu g/m^3$ for about 5 persons. This corresponds to a difference in risk of the two groups of a factor of 5000!

Because of the tremendous effect of a few measurements, the calculated upper-bound risk for vinylidene chloride must be considered tentative. Only additional data on personal exposures (c llected by methods with a sensitivity of 1 $\mu g/m^3$ or less) will provide the information necessary for an adequate risk assessment.

p-Dichlorobenzene

This chemical has two major uses: as a moth repellent, it is a registered pesticide; as an air "freshener", it is an additiv (ften unlabeled) to consumer products. From the TEAM Study results, it appears that about a third of homes employ p-DCB. Therefore the average risk to users is about 3 times the risk shown in the table. The risk to non-users should be no more than that associated with average outdoor concentrations, or about 1% of the risk to users.

Chloroform

This chemical is unique among the target VOCs in having many routes of exposure: air, food, water, beverages. Indoor and outdoor air levels and levels in drinking water have been documented in all the TEAM Studies. The pilot TEAM Study of 1980-82 also documented chloroform levels of 15-56 ppb ($\mu g/L$) in milk, butter, cheese, and ice cream; and 9-178 ppb in soft drinks[14]. A recent Japanese study (Y. Sato: personal communication) of seven housewives indicated that they were exposed to chloroform through all routes, but that the diet provided more exposure (10.7 $\mu g/day$) than the air (2.0) and water routes (2.4 $\mu g/day$).

Methylene chloride

This chemical is too volatile to be collected on Tenax; therefore few personal exposure measurements have been made. Several indoor and outdoor measurements were made in the 1987 TEAM Study in Los Angeles using evacuated canisters--these are the values on which the risk estimate has been based. However, the existing data is s sparse that the estimate must be considered very speculative.

Short-term exposures at the high ppm level from using paint strippers have been documented[15]. Such an exposure for one day would equal the lifetime exposure to ambient concentrations of methylene chloride. Therefore the population risk from this chemical might better be calculated from data on the number of people who use paint strippers and the amount of time they use them.

Ethylene dibromide (1,2-dibromoethane)

This chemical is widely used as a fungicide, particularly on grain, and therefore the main risk is thought to be through food. However, the potency is so high that even the very low airborne exposures measured in the TEAM Study produce a non-negligible risk. Only 15 of 621 personal air samples (2.4%) exceeded th quantifiable limit of 0.05 $\mu g/m^3$. Another 61 samples (12.2%) showed trace amounts. If we assume a value of 0 for the 545 non-detected samples, and the lowest possible value of 0.05 for the 61 trace samples, the average exposure is 0.014 $\mu g/m^3$. Assuming the maximum values for the trace samples (0.24 $\mu g/m^3$) and th non-detects (0.05 $\mu g/m^3$) results in an average exposure of 0.087 $\mu g/m^3$. A value between these two extremes is 0.05 $\mu g/m^3$, resulting in a risk estimate of 25 X 10^{-6}. Only 3 of 282 outdoor air samples were measurable (and only 5 were at trace levels), and the range of average values using the same assumptions as above is 0.003 to 0.06 $\mu g/m^3$. The maximum personal exposure was only 0.97 $\mu g/m^3$, so that the risk calculations are not extensively skewed by a few samples as they were in the case f vinylidene chloride.

Chlordane and Heptachlor

These pesticides were recently withdrawn (April 1988), after wide use as termiticides (≈85% of the market). They were applied primarily as a liquid poured or injected into soil around building foundations. Therefore their appearance in the NOPES study as airborne vapors may indicate widespread intrusion of soil gas into the home through cracks or drains in the basement or ground floor.

Aldrin and Dieldrin

These pesticides were withdrawn from use in the U.S. in the early 1980's. They were used mainly as termiticides (about 10% of the market). Their appearance in the NOPES study is further indication of a long half-life in soil coupl d with some mechanism allowing intrusion into the h me.

Dichlorvos (DDVP)

This insecticide was widely used on "pest strips" before such a use was banned. Measurements during different seasons in the two cities ranged from 98-99% not detectable in Springfield, and from 65-89% not detectable in Jacksonville. As discussed above, although the risk averaged over the entire population is close to the 10^{-6} level of risk, when averaged over the smaller population of users, the risk climbs to about 15×10^{-6} in both cities.

Exposures Through Other Routes

All of the above chemicals (both VOCs and pesticides) were measured in drinking water, and found to present less than 1% of the risk due to airborne exposures with the single exception of chloroform. All of the chlorinated VOCs were also measured in food and beverages; again chloroform was the only VOC found in significant amounts in food.

Exposures through routes other than air and water have been documented for some of the pesticides. Many of the pesticides have been measured in food by the FDA for years; however, exposures in food account for only a small proportion of total exposure to the four pesticides of highest risk through airborne routes. Food exposures outweigh air exposures for some of the ther pesticides (e.g, Captan).

House dust may provide an important reservoir for any or all of the pesticides, and possibly also for the least volatile of the VOCs (p-dichlorobenzene, tetrachloroethylene). DDT was found in house dust in five of eight homes in the NOPES study. Since this chemical has been banned for nearly two decades, its appearance in house dust is troubling. Ingestion of the dust by toddlers could be an important additional source of risk. The DDT may be tracked in on people's shoes from outdoor soil.

Comparison with Other Environmental Risks

Several organic chemicals of interest were not monitored in the TEAM Studies. Among these are formaldehyde and 1,3-butadiene. Risk estimates for these chemicals may be compared to the risks calculated above.

The carcinogenicity of formaldehyde is controversial, due to the unusual metabolic pathway associated with its carcinogenicity in rodents. Estimates for the cancer risk of formaldehyde[16] range over extreme limits, from zero to 10^{-3}. Employing an intermediate potency factor (unit risk of 1.3×10^{-5} $(\mu g/m^3)^{-1}$) and measured values of 40 $\mu g/m^3$ for normal (non-mobil home) housing stock results in a risk of about 5 X 10^{-4}. (Average outdoor concentrations of formaldehyde have been about 4 $\mu g/m^3$ corresponding to a risk of about one tenth of this level.)

Recent animal studies of 1,3-butadiene have resulted in revising its potency upward by nearly three orders of magnitude. Although no personal exposure data are available for this chemical, a recent study has measured the level in sidestream smoke at about 400 μg per cigarette[17]. This is approximately th level of benzene in sidestream smoke (330 μg per cigarette); therefore, if 1,3-butadiene is not too reactive, we can calculat that it will be elevated by about 4 $\mu g/m^3$ in smoking homes and by 13 $\mu g/m^3$ in workplaces allowing smoking. Using the revised unit risk value of 2.8 X 10^{-4}, and assuming 38 million homes with smokers averaging three residents each, 75 million workers in workplaces allowing smoking, and 26 million non-workers exposed to cigarette smoke, we arrive at an upper-bound risk associated with exposure to 1,3-butadiene in environmental tobaccos smoke of 6 X 10^{-4}. No data exist on personal exposures or indoor concentrations of 1,3-butadiene. (Outdoor concentrations of this chemical have been estimated to lie within a range of 0.3-1.6 $\mu g/m^3$, corresponding to a risk of about 1-4 X 10^{-4}.)

Thus the individual risks for formaldehyde and 1,3-butadiene are greater than the airborne risk of any of the other 35 VOCs and pesticides considered in this report. However, the great uncertainty in the carcinogenic potency of formaldehyde, including the uncertainty as to whether it is a human carcinogen at all, and the lack of exposure data for 1,3-butadiene make the risk estimates for these two chemicals particularly speculative.

The combined upper-bound risk of about 10^{-3} associated with these 37 predominantly indoor org nic chemicals appears to be similar to the risks associated w h the most severe environmental hazards (radon and ssive smoking). For example, the risk associated with nonsmoke exposure to radon has been estimated to be about 10^{-3} and that with passive smoking has b en estimated[18] at 2 X 10^{-3}. It should be noted, however, that the risk estimates for radon and passive smoking are based on human epidemiology studies, and are therefore on firmer ground than all of the risk estimates for the organic chemicals with the exception of benzene.

The risk estimates for these organic chemicals are considerably higher than the risks associated with some EPA regulations (NESHAPS chemicals and Superfund clean-up criteria). Similar conclusions regarding the importance of indoor air pollution compared to other environmental hazards have been reached by EPA Headquarters[19] and by three Regions[20]; both of these reports rank indoor air pollution as among the top two or three environmental threats to public health.

Two other risk estimates for personal exposure to VOCs have been published[9,16]. McCann arrived at similar risk estimates for most of the chemicals; Tancrede estimated 5-10 times higher risks, due partly to using a different method for calculating potencies from animal data and partly to considering explicitly several additional sources of uncertainty. No previous risk estimates for most of these pesticides have been possible, due to the lack of exposure information.

Actions to avoid these risks may be taken by individuals. Since the sources of the risks are often personal activities (smoking, using air fresheners), these activities can be halted or modified. (For example, smokers could establish a room in th home with separate ventilation.) Exposures from chloroform could be reduced by drinking bottled water or using an activated carb n filter on the water supply. Exposures from petroleum-based products could be reduced by discarding or storing used paint cans and sprays in a detached garage or tool shed. Dry-clean d clothes could be hung outdoors for a day--one study indicates that 20-30% of tetrachloroethylene residues on the clothes will outgas during the first day.

The reason for the large number of pesticides observed in indoor air in the latest TEAM Study is not well understood. Termiticides, like radon gas, may be entering the basement due to soil gas movement; it may be that the same techniques to control radon (sealing the foundation, providing separate ducting at th entrance points) may also control termiticide entry. Other pesticides, particularly the long-lived chlorinated hydrocarbons such as DDT, may be entering the home by being tracked in on people's shoes. If so, removing shoes before entering the home, and reducing or eliminating the use of carpets or rugs (which collect large amounts of dust containing pesticides and metals as well), should reduce pesticide exposures.

SUMMARY AND CONCLUSIONS

Measured personal exposures to 12 VOCs and 23 pesticides in EPA's TEAM Studies have been used to arrive at upper-bound lifetime cancer risk estimates. Seven VOCs and seven pesticides have upper-bound risks ranging from 10^{-6} to 10^{-4}. The combined upper-bound risk of about 10^{-3} from these organic indoor air pollutants is nearly comparable to the estimates of risk from radon and environmental tobacco smoke. (However, the latter two estimates are based on human epidemiology studies, and are therefore subject to far less uncertainty.) These upper-bound risks are much greater than the health risks associated with most other environmental problems.

Several chemicals for which we have inadequate information, either on exposure or potency, to calculate risk were identified: //

vinylidene chloride, methylene chloride, 1,3-butadiene, formald hyde, ethylene dibromide, chlorpyrifos, and diazinon. //

Despite the recognized large uncertainty in these risk estimates, these findings provide additional support for the conclusion of two recent comparative rankings of environmental risk by EPA: that indoor air pollution is one of the greatest threats to public health of all environmental problems.

REFERENCES

1. Integrated Risk Information System (IRIS), US EPA, Washington, DC.

2. Engler, R., memorandum to Office of Pesticide Programs Division Directors, USEPA, October 27, 1989.

3. Wallace, Lance A., Pellizzari, E.D., Hartwell, T.D., Sparacino, C., Whitmore, R., Sheldon, L., Zelon, H. and Perritt, R., "The TEAM Study: personal exposures to toxic substances in air, drinking water, and breath of 400 residents of New Jersey, North Carolina, and North Dakota" Environmental Research 43: 290-307 (1987).

4. Wallace, L. A., Pellizzari, E.D., Hartwell, T.D., Whitmore, R., Zelon, H., Perritt, R. and Sheldon, L. "The California TEAM Study: breath concentrations and personal exposures to 26 volatile compounds in air and drinking water of 188 residents of Los Angeles, Antioch, and Pittsburg, CA." Atmos. Environ. 22: 2141-2163 (1988).

5. Immerman, F.W. and Schaum, J.L. Final Report of the Nonoccupational Pesticide Exposure Study (NOPES), US Environmental Protection Agency, Research Triangle Park, NC (1990).

6. Wallace, Lance A. "Cancer risks from organic chemicals in the home", in Environmental Risk Assessment: Is Analysis Useful?, APCA Specialty Conference Proceedings, Air Pollution Control Assoc., Pittsburg, PA. SP-55 (1986).

7. Wallace, Lance A., Pellizzari, E., Hartwell, T., Perritt, K., and Ziegenfus, R., "Exposures to benzene and other volatile organic compounds from active and passive smoking," Archives of Environmental Health 42: 272-279 (1987).

8. McKone, T.E. "Human exposure to VOCs in household tap water: The indoor inhalation pathway", Env. Sci. Tech. 21: 1194-1201 (1987).

9. Tancrede, M, Wilson, R. Zeise, L. and Crouch, E.A.C., "The carcinogenic risk of organic vapors indoors: a theoretical survey," Atmos. Env. 21: 2187-2205 (1987).

10. Allen, B., Crump, K. and Shipp, A., "Correlation between Carcinogenic Potency of Chemicals in Animals and Humans," Risk Anal. 8:531-544 (1988).

11. Sandler, D.P., Everson, R.B., Wilcox, A.J., and Browder, J.P. "Cancer risk in adulthood from early life exposure to parents' smoking". Am. J. Public Health, 75:467 (1985).

12. Stjernfeldt, M., Berglund, K., Lindsten, J., and Ludvigsson, J. "Maternal smoking during pregnancy and risk of childhood cancer", Lancet, June 14, 1986:1350-52 (1986).

13. Krause, C., Mailahn, W., Nagel, R., Schulz, C., Seifert B., Ullrich D. "Occurrence of volatile organic compounds in the air of 500 Homes in the Federal Republic of Germany", in Proceedings of The 4th International Conference on Indoor Air Quality and Climate, Vol. 1, pp. 102-106. Institute for Soil, Water, and Air Hygiene. Berlin (West) (1987).

14. Entz, R.C., Thomas, K. and Diachenko, G. "Residues of Volatile Halocarbons in Foods Using Headspace Gas Chromatography," J. Agric. Food Chem. 30:846-849 (1982).

15. Gi in, J. R. and Hodgson, A.T. "Exposure to methylene chlori from controlled use of a paint remover in a residence" paper presented at 80th annual meeting of the Air Pollution Control Association, New York, NY June 21-26, 1987.

16. McCann, J. Horn, L., Girman, J. and Nero, A.V. Potential risks from exposure to organic compounds in indoor air. Lawrence Berkeley Laboratory, report # LBL-22473. Berkeley, CA (1986).

17. Lofroth, G. Burton, B., Forehand, L., Hammond, S.K., Seila, R., Zweidinger, R., and Lewtas, J., "Characterization of environmental tobacco smoke," Env. Sci. and Tech. : (1989).

18. Repace, J., "A quantitative estimate of nonsmokers' lung cancer risk from passive smoking," Env. Int. 11:3-22 (1985).

19. USEPA, Unfinished Business: A Comparative Assessment of Environmental Problems, Washington DC (1987).

20. USEPA, Comparing Risks and Setting Environmental Priorities: Overview of Three Regional Projects, Office of Policy, Planning, and Evaluation, Washington, DC (1989).

SL 064739

Table I. Upper-Bound Lifetime Cancer Risks of 12 VOCs Measured in the TEAM Studies (1980-87

Chemical	Exposure[a] ($\mu g/m^3$)	Potency $(\mu g/m^3)^{-1}$ (X 10^{-6})	Risk (X10^{-6})	Outdoor Air Concentration[b] ($\mu g/m^3$)
Benzene				
Air	15	8	120[c]	6
Smokers	90	8	720[c]	--
Vinylidene chloride	6.5[d]	50	320	<1
Chloroform				
Air	3	23	70	0.6
Showers (Inhalation)	2	23	50	--
Water	30[e]	2.3[e]	70	--
Food & Beverages	30[e]	2.3[e]	70	--
p-Dichlorobenzene	22	4	90	0.6
1,2-Dibromoethane	0.05	510	25	0.03
Methylene chloride	6[f]	4	24	2[f]
Carbon tetrachloride	1	15	15	0.6
Tetrachloroethylene	15	0.6	9	3
Trichloroethylene	7	1.3	9	1
Styrene				
Air	1	0.3[g]	0.3	0.3
Smokers	6	0.3	2	--
1,2-Dichloroethane	0.5	7	4	0.2
1,1,1-Trichlorethane	30	0.003	0.1	7

---

[a] Arithmetic means based on 24-hour average exposures of ≈ 750 persons in six urban areas measured in the TEAM Studies

[b] Based on backyard measurements in 175 homes in six urban areas

[c] The risk estimates for benzene are based on human epidemiology and are therefore mean as opposed to upper-bound estimates

[d] Six measurements exceeding 1000 $\mu g/m^3$ dropped from the calculation; inclusion of the measurements leads to an average exposure of 150 $\mu g/m^3$.

[e] These figures are in $\mu g/L$ or ppb rather than $\mu g/m^3$

[f] Based on only eight 24-hour measurements in 1987.

[g] Source: US EPA (1983) Review and Evaluation of Evidence for Cancer Associated with Air Pollution, EPA 450/5-83-006.

Table II. Upper-bound Lifetime Cancer Risks from Airborne Exposur s 23 Pesticides Measured in the NOPES TEAM Study

Pesticide	Exposure[a] (ng/m^3)	Potency (kg-d/mg)	Risk (X10^{-6})	Outdoor Air Concentration[b] (ng/m^3)
Banned Termiticides				
Heptachlor	71	4.5	90 (19)[c]	7
Chlordane	198	1.3	70 (15)	14
Aldrin	13	17	60 (13)	0.1
Dieldrin	3	16	14 (3)	0.2
Heptachlor Epoxide	0.4	9.1	1 (0.2)	0.1
DDE	2.2	0.34	0.2 (0.04)	ND[d]
DDT	0.7	0.34	0.1 (0.02)	ND
Other Pesticides				
Dichlorvos	33	0.29	2.7	ND
γ-BHC (Lindane)	6.6	1.3	2.5	0.4
α-BHC	0.5	6.3	1	ND
Propoxur	100	0.0079	0.2	2.5
Hexachlorobenzene	0.3	1.67	0.1	0.1
Dicofol	2.6	0.34	0.05	ND
o-Phenylphenol	58	0.0016	0.02	0.6
2,4-D	0.6	0.019	0.003	0.1
Atrazine	0.05	0.22	0.003	ND
cis-Permethrin	0.4	0.022	0.003	ND
trans-Permethrin	0.1	0.022	0.001	ND
Chlorothalonil	0.7	0.011	0.002	0.5
Folpet	0.5	0.0035	0.0005	0.2
Captan	0.1	0.0023	0.00007	ND
DDD	<4	0.34	≤0.4	ND
Pentachlorophenol	<730	0.013	≤3	ND

[a] Arithmetic mean of population-weighted and seasonally-weighted average personal exposures measured for 173 persons in Jacksonville, FL and 85 persons in Springfield/Chicopee, MA.

[b] Based on outdoor measurements at each home in the two cities.

[c] All risks calculated assuming 70-year lifetime exposure at the measured levels. For banned pesticides, whose environmental concentrations should decrease over time, an alternative calculation of risk (in parentheses) assuming a 10-year half-life in soil is provided.

[d] Not Detected

REGULATORY PROGRAM OF THE UNITED STATES GOVERNMENT

APRIL 1, 1990 - MARCH 31, 1991

Supporters of this system argue that it woulc provide an incentive for better estimates of the costs of legislative proposals and a basis for an explicit discussion of the costs and tradeoffs of such proposals. High cost ceilings would focus attention on the expected benefits of the program, and alternative approaches; cost ceilings that were too low would prevent agencies from issuing implementing regulations. Such an approach would, needless to say, give agencies an incentive to choose regulatory approaches that would produce the greatest benefits at the lowest costs.

ISSUES AND AREAS FOR FURTHER STUDY

While the fiscal budget process provides a continuous record of actual expenditures, there is no comparable record of the cost of meeting regulatory requirements.[30] Members of Congress and the past two Administrations have considered developing an accounting framework to record direct regulatory expenditures, but more work needs to be done to solve the practical accounting problems inherent in measuring the private expenditures that Federal regulations mandate. These include:

- Developir a record of actual expenditures while minimizi the recordkeeping burden on the private sector;
- Identifying an appropriate "baseline," recognizing that some costs would be incurred even in the absence of Federal regulation; and
- Estimating the costs of forgoing certain products where Federal regulation prohibits production or distribution.

Each of these raises difficult issues in designing an effective regulatory budget process. For example, the costs of banning a product are not directly measurable and can only be estimated using complex statistical mod·'s. However, measuring only the rect compli- ar ›sts for oversight purposes creates a bias toward be g substances and products instead of control- lir em.

. a first step in determining the feasibility of the regulatory budget concept, OMB has begun systematically to collect the costs of all significant published regulatory actions. Analysis of these data should aid in the development of ways to overcome the problems of regulatory budgeting, uncover unforeseen problems in developing cost estimates, and more fully refine a workable regulatory budgeting process.

Current Regulatory Issues in Risk Assessment and Risk Management

Many Federal agency regulatory decisions are intended to reduce risks to human life and health. Government regulations control which agricultural chemicals may be used to reduce insect damage, increase farm yields, and improve the quality of food products. Other rules govern hazards in the Nation's workplaces and emissions from its factories. There are regulations directing the way in which automobiles must be manufactured, commercial aircraft maintained, and trains operated. Hardly any widespread human activity that entails risk is free of some degree of social control, often achieved through government regulation.

Regulatory decisions involving risk require agencies to address questions such as, "How safe is 'safe'?" and "How clean is 'clean'?" When government agencies promulgate regulations intended to reduce a risk or mitigate a hazard, they are engaging in what has become known as *risk management*. These policy choices inevitably involve consideration of both the risks entailed by the underlying activity and the social consequences of regulatory intervention. Thus, the first challenge of risk management is to set priorities to determine which risks are worth reducing and which are not.

For government to carry out its risk-management responsibilities, there must be an extensive investment in the careful assessment and quantification of risks. The term *risk assessment* means the application of credible scientific principles and statistical methods to develop estimates of the likely effects of natural phenomena and human activities.

The need to keep risk assessment and risk management separate has long been the objective of responsible public officials. In 1983, the National Academy of Sciences (NAS) studied the process of managing risk

[30] Researchers, using different methods, assumptions, and time periods, have formed incomplete estimates by adding up the cost individual regulations. These estimates accordingly show considerable variation for current annual costs ranging from $60 billion to $1 billion a year—5 to 15 percent of current Federal outlays.

SL 064743

in the Federal Government and offered the following recommendations, among others:

Recommendation 1: Regulatory agencies should take steps to establish and maintain a clear conceptual distinction between assessment of risks and the consideration of risk management alternatives; that is, the scientific findings and policy judgments embodied in risk assessments should be explicitly distinguished from the political, economic, and technical considerations that influence the design and choice of regulatory strategies.[31]

Recommendation 2: Before an agency decides whether a substance should or should not be regulated as a health hazard, a detailed and comprehensive written risk assessment should be prepared and made publicly available. This written assessment should clearly distinguish between the scientific basis and the policy basis for the agency's conclusions.[32]

The belief that risk assessment and risk management should be kept separate enjoys widespread support among professional risk-assessment practitioners and risk-management officials.[33] Others have emphasized the importance of ensuring that policy biases do not distort the analysis of alternative risk-management choices.[34] The NAS principles have also have been endorsed by a number of Federal agencies, including the Office of Science and Technology Policy (OSTP), the Environmental Protection Agency (EPA), and the Department of Health and Human Services (HHS).[35]

Unfortunately, risk-assessment practices continue to rely on conservative models and assumptions that effectively intermingle important policy judgments within the scientific assessment of risk. Policymakers must make decisions based on risk assessments in which scientific findings cannot be readily differentiated from embedded policy judgments. This policy environment makes it difficult to discern serious hazards from trivial ones, and distorts the ordering of the Government's regulatory priorities. In some cases, the distortion of priorities may actually increase health and safety risks.

This section explores some of the continuing difficulties that plague the practice of risk assessment, and describes briefly their policy implications. It can be summarized in three observations:

The continued reliance on conservative (worst-case) assumptions distorts risk assessment, yielding estimates that may overstate likely risks by several orders of magnitude. Many risk assessments are based on animal bioassays utilizing sensitive rodent species dosed at extremely high levels. Conservative statistical models are used to predict low-dose human health risks, based on the assumption that human biological response mimics that observed in laboratory animals. Worst-case assumptions concerning actual human exposure are commonly used instead of empirical data, further exaggerating predicted risk levels.

Conservative biases embedded in risk assessment impart a substantial "margin of safety". The choice of an appropriate margin of safety should remain the province of responsible risk-management officials, and should not be preempted through biased risk assessments. Estimates of risk often fail to acknowledge the presence of considerable uncertainty, nor do they present the extent to which conservative assumptions overstate likely risks. Analyses of risk-management alternatives routinely ignore these uncertainties and treat the resulting upper-bound estimates as reliable guides to the likely consequences of regulatory action. Decisionmakers and the general public often incorrectly infer a level of scientific precision and accuracy in the risk-assessment process that does not exist.

Conservatism in risk assessment distorts the regulatory priorities of the Federal Government, directing societal resources to reduce what are often trivial carcinogenic risks while failing to address more substantial threats to life and health. Distortions are probably most severe in the area of cancer-risk assessment, because many conservative models and assumptions were developed specifically for estimat-

[31] National Academy of Sciences, *Risk Assessment in the Federal Government: Managing the Process*, Washington, DC: National Academy Press, 1983 (hereinafter, *NAS Risk Management Study*), p. 151.

[32] Ibid., p. 153.

[33] For representative views of risk-assessment practitioners see, e.g., Lester B. Lave, *The Strategy of Social Regulation: Decision Frameworks for Policy*, Washington, DC: Brookings, 1981; Lester B. Lave, "Methods of Risk Assessment," Chapter 2 in *Quantitative Risk Assessment in Regulation*, Lester B. Lave, ed., Washington, DC: Brookings, 1982, esp. pp. 52–54. For representative views of risk-management officials see, e.g., William D. Ruckelshaus, "Science, Risk, and Public Policy," *Vital Speeches of the Day*, Volume 49, No. 20, August 1, 1983, pp. 612-615.

[34] See, e.g., Howard Kunreuther and Lisa Bendixen, "Benefits Assessment for Regulatory Problems," and Baruch Fischhoff and Louis Anthony Cox, Jr., "Conceptual Framework for Regulatory Benefits Assessment," Chapters 3 and 4, respectively, in *Benefits Assessment: The State of the Art*, Judith D. Bentkover, Vincent T. Covello, and Jeryl Mumpower, eds., Dordrecht, Netherlands: D. Reidel, 1986, pp. 44–45, 59–61.

[35] See U.S Office of Science and Technology Policy, "Chemical Carcinogens: A Review of the Science and Its Associated Principles," Principle 29 (50 FR 10378, March 14, 1985, hereinafter, *OSTP Risk Assessment Guidelines*); U.S. Environmental Protection Agency, "Guidelines for Carcinogen Risk Assessment," 51 FR 34001 (September 24, 1986, hereinafter, *EPA Carcinogen Risk Assessment Guidelines*); U.S. Department of Health and Human Services, *Risk Assessment and Risk Management of Toxic Substances*, April 1985, p. 20.

ing upper bounds for these risks. Risk-assessment methods with similar conservative biases are less common elsewhere, particularly in those areas where real-world data are available, or where the mechanism by which injury or illness occurs is better understood.

A renewed commitment to the NAS recommendations is clearly warranted. As quantitative risk assessment plays an increasingly significant role in risk management, the need to separate science from policy becomes ever more important, if either process is to maintain public confidence. As former EPA Administrator William D. Ruckelshaus has noted:

> Risk assessment...must be based on scientific evidence and scientific consensus *only*. Nothing will erode public confidence faster than the suspicion that policy considerations have been allowed to influence the assessment of risk.[36]

ALTERNATIVE RISK-ASSESSMENT METHODOLOGIES

Risk assessments of chemical substances in general (and of possible carcinogens in particular) involve a mixture of facts, models, and assumptions. There is considerable debate concerning the scientific merits of the models and assumptions commonly used in risk assessments. In some cases, a scientific consensus has developed to support a particular model or assumption. In other instances, however, certain models and assumptions are relied upon because they reflect past practices rather than the leading edge of science. Furthermore, a scientific basis for several of the most critical models and assumptions simply does not exist.

Most scientists agree that these models and assumptions impart a conservative bias: that is, they lead to risk projections that the actual (but unknown) risk is very unlikely to exceed. These "upper-bound" estimates are often useful as a screening device, to exclude from regulatory concern potential hazards that are insignificant even under worst-case conditions. Unfortunately, upper-bound risk estimates are routinely employed for altogether different purposes, such as estimating the likely benefits of regulatory actions. Policymakers are required to act on the basis of biased representations of both the magnitude of the underlying hazard and the extent to which Government action will ameliorate it.

Contemporary risk assessment relies heavily upon animal bioassay and epidemiology. Each approach has theoretical advantages and disadvantages. In practice, both can be misused to bolster preestablished conclusions. The following discussion emphasizes problems in carcinogenic risk assessment, because the prevention and cure of cancer plays such a major role in policy issues involving risks to life and health.

Animal Bioassay

Animal testing enables scientists to estimate risks ex ante, before human health effects materialize, whereas epidemiological studies can only detect such effects ex post. In addition, animal tests can be conducted under tightly controlled laboratory conditions, which provide more reliable estimates of exposure and avoid many of the confounding factors that often plague epidemiological investigations. The relatively short lifetimes of experimental mammals (such as rats and mice) allow scientists to ascertain the possible effects of long-term exposure in just a few years.

Animal testing suffers serious limitations, however, arising from certain critical assumptions. Despite its routine application, there is no accepted scientific basis for the assumption that results can be meaningfully extrapolated from test animals to humans.[37] Some scientists believe that animal data should not be used in assessing human health risks.[38]

Another critical limitation is the reliance on very high doses to generate adverse effects in test animals.[39] A mathematical model must be used to bridge the gap between these high-dose exposures and the low-dose exposures more typically faced by people. Many different mathematical models can be constructed to fit the data at high doses. These models often vary enormously, however, in their predictions of risk at low doses.

Beyond these unavoidable methodological constraints, the results of animal bioassays may be subject to conflicting scientific interpretation or strongly influenced by the choice of research method.

[36] William D. Ruckelshaus, (*op. cit.*), p. 614.

[37] *OSTP Guidelines*, Guideline 8, p. 10376.

[38] See, e.g., Bruce Ames, Renae Magaw, and Lois Swirsky Gold, "Ranking Possible Carcinogenic Hazards," *Science*, Vol. 236, April 17, 1987; Gio Batta Gori, "The Regulation of Carcinogenic Hazards," *Science*, Vol. 208, April 18, 1980.

[39] *OSTP Guidelines*, Guideline 11, p. 10377.

SL 064745

Tissue preparation and histology present obvious opportunities for error, as experts may disagree as to how slides should be interpreted.[20] This problem generally is not significant at high doses, where malignancies are often obvious. At low doses, however, pathologists often differ in how they distinguish tumors from hyperplasia. Subjectivity cannot be avoided where such interpretations of the data must be made.[21]

Epidemiology

Epidemiology is attractive because it largely avoids these two problems. It focuses on observable human health effects instead of on hypothesized outcomes based on animal experimentation, and it relies upon real-world exposures to generate empirical data. Many of the serious problems associated with animal studies can be avoided, allowing researchers to develop risk estimates that are directly related to human health.

Unfortunately, epidemiological research suffers from its own set of limitations. For example, retrospective studies often have difficulty correlating morbidity and mortality with exposure to specific substances. Exposure data are commonly lacking, incomplete, imprecise, or affected by systematic recall or selection biases. Furthermore, the risks these studies seek to detect are often very small relative to background, thus making statistically significant effects difficult to observe. When health effects are latent, correlating exposures to illness is even harder.

Besides these unavoidable methodological limitations, epidemiological studies often suffer from outright bias. Many studies employ scientifically questionable procedures aimed at demonstrating positive relationships between specific substances and human illness.[22] Some researchers use inappropriate statistical procedures to "mine" existing databases in search of associations. One result of these practices is that epidemiological studies often display contradictory results.[23]

Despite these constraints, properly conducted animal bioassays and epidemiological studies both have useful roles to play in quantitative risk assessment. Indeed, they are complementary. The usual weaknesses of epidemiological investigations—unreliable exposure data, confounding effects—are readily avoided in laboratory experiments on animals. The weaknesses of animal bioassays—high- to low-dose extrapolation, animal-to-man conversion—do not arise in epidemiological studies. Careful risk assessment incorporates both types of analysis to ensure that the emerging picture of human health risk is as complete as possible, and that inferences derived from this picture are themselves internally consistent.

ISSUES IN RISK ASSESSMENTS DERIVED LARGELY FROM ANIMAL BIOASSAYS

Animal bioassays tend to dominate current risk assessments. An important reason for this is that the derivation of dose-response relationships is a critical regulatory motive for performing quantitative risk assessment. Animal studies are ideally suited to serve this purpose by virtue of the controlled conditions under which dose and response can be calibrated. Epidemiological studies often are relegated to providing merely a "reality check" to ensure that the implications of animal bioassays are plausibly consistent with real-world experience. Because of this heavy emphasis on animal testing, the focus here is on several major problems that arise with respect to risk assessments primarily based on the results of animal bioassays.

The Use of Sensitive Test Animals

To enhance the power of animal tests, scientists typically rely on genetically sensitive test animals. It

[20] In the original analysis of the rat bioassay used to derive the dose-response function for dioxin, 9 of 85 controls were said to develop liver tumors. An independent review of this data resulted in 16 of the 85 controls being classified as having such tumors. See U.S. Environmental Protection Agency, *A Cancer Risk-Specific Dose Estimate for 2, 3, 7, 8-TCDD, Appendix A*, EPA/600/6-88/007Ab, June 1988 (hereinafter, *Dioxin Risk Assessment Appendix A*), pp. 2–3.

[21] Colin N. Park and Ronald D. Snee, "Quantitative Risk Assessment: State-of-the-Art for Carcinogenesis," Chapter 4 in *Risk Management of Existing Chemicals*, Rockville, MD: Government Institutes, 1983, p. 56.

[22] Alvan R. Feinstein, "Scientific Standards in Epidemiological Studies of the Menace of Daily Life," *Science*, Vol. 242, December 2, 1988, pp. 1257–1263.

[23] Linda C. Mayes, Ralph I. Horowitz, and Alvan R. Feinstein, "A Collection of 56 Topics with Contradictory Results in Case-Control Research," *International Journal of Epidemiology*, Vol. 17, No. 3 (1988), pp. 680–685.

is unclear whether these species accurately mimic biological responses in humans.

Some test species are extremely sensitive. For example, approximately one-third of all male B6C3F1 mice, a common test species, spontaneously develop liver tumors.[34] The same phenomenon occurred in an important bioassay concerning dioxin using female Sprague-Dawley (Spartan) rats. Tumors observed in dosed animals were predominantly located in the liver. However, approximately one-fifth of the animals in the control group also developed liver tumors.[35] The relevance of elevated liver tumors in hypersensitive species has been questioned by scientists and is not universally considered probative evidence of carcinogenicity. Nevertheless, cancer risk assessments often proceed on the assumption that these data are sufficient to conclude that a substance is indeed a carcinogen.[36]

The reliance on sensitive test animals also biases risk assessments in a more subtle way. It establishes powerful incentives to search for and develop increasingly sensitive test species. As test animals become more sensitive, repeated testing using identical protocols will tend to result in higher and higher estimates of risk even if all other factors are held constant.

Selective Use of Alternative Studies

In their respective risk-assessment guidelines, both OSTP and EPA recommend that relevant animal studies should be considered irrespective of whether they indicate a positive relationship.[37] In practice, however, studies that demonstrate a statistically significant positive relationship routinely receive more weight than studies that indicate no relationship at all.[38] For example, the plant growth regulator daminozide (Alar) and its metabolite unsymmetrical 1,1-dimethylhydrazine (UDMH) recently received B2 classifications ("probable human carcinogen"). Each of these classifications was based on a single positive animal bioassay.[39] Overcoming such a classification requires, at a minimum, two "essentially identical" studies showing no such relationship.[40] In the case of Alar and UDMH, however, a more stringent test was apparently applied: Three high-quality negative studies showed no significant effects; these studies appear to have received little or no weight in the classification decision.[41]

Selective Interpretation of Results

Risk-assessment guidelines generally give the greatest weight to the most sensitive test animals. Thus, if a substance has been found to cause cancer in one

[34] Ames *et al.*, *(op. cit.)*, p. 276.

[35] *Dioxin Risk Assessment Appendix A*, pp. 2–3.

[36] See Ames *et al.*, *(op. cit.)*, p. 276 (arguing that such data are irrevelant); *OSTP Guidelines* Guideline 9, p. 10377 (concluding that such data "must be approached carefully"); and *EPA Carcinogen Risk Assessment Guidelines*, p. 33995 (making the policy judgment that such data are sufficient evidence of carcinogenesis). Liver tumors dominated in EPA's dioxin risk assessment. See *Dioxin Risk Assessment*, appendix A, pp.2–3.

[37] See *OSTP Guidelines*, Guideline 25, p. 10378; *EPA Carcinogen Risk Assessment Guidelines*, p. 33995.

[38] See *EPA Carcinogen Risk Assessment Guidelines*, p. 33999–34000. A single animal test that shows a positive result "to an unusual degree" (p. 33999) is sufficient to warrant at least a B2 classification ("probable human carcinogen"), even if this result occurs in a species known to have a high rate of spontaneous tumors. A strong animal bioassay or epidemiological study showing no evidence of carcinogenic effect cannot overcome this presumption (p. 34000).

[39] See "Second Peer Review of Daminozide (Alar) and UDMH (Unsymmetrical 1,1-dimethylhydrazine)," Memorandum from John A. Quest to Mark Boodee, U.S. Environmental Protection Agency, OPTS, May 15, 1989 (hereinafter, *Alar/UDMH Internal Peer Review No. 2*). This internal OPTS panel reviewed several recent studies on Alar and UDMH.

One study of Alar yielded a statistically significant increase in common lung tumors in mice, but only for one of three dosage levels. Results were not statistically significant at the higher and two lower dosages, and controls also displayed unusually high tumor incidence. 90% of the lung tumors in dosed mice were benign, versus 89% in the controls.

One study of UDMH yielded statistically significant increases in common lung and uncommon liver tumors in mice, but only for the higher of two dosages. 97% of the lung tumors in dosed mice were benign, versus 100% in the controls. 29% of the liver tumors in dosed mice were benign; no tumors were observed in the controls.

Prior studies that purported to show a carcinogenic response had been judged inadequate by EPA's Scientific Advisory Panel, an external peer review group. The Office of Pesticides and Toxic Substances (OPTS) panel noted that a different internal EPA risk-assessment panel (the Carcinogen Assessment Group) considered these studies sufficient to justify B2 classifications when it evaluated them for EPA's Office of Solid Waste and Emergency Response. Despite the scientific controversy, the OPTS panel interpreted these prior studies as "supporting evidence" under EPA's risk-assessment guidelines.

[40] See *EPA Carcinogen Risk Assessment Guidelines*, p. 33995 (establishing the need for replicate identical studies showing no effect), and p. 33999 (establishing the minimum requirement of two well-designed studies showing no increased tumor incidence to warrant a "no evidence" determination).

[41] *Alar/UDMH Internal Peer Review No. 2*, pp. 6, 8, 9. EPA's scheme for carcinogen classification is itself an issue among scientists. See, e.g., U.S. Environmental Protection Agency, Risk Assessment Forum, *Workshop Report on EPA Guidelines for Carcinogen Risk Assessment*. EPA/625/3-89/015, Washington, DC: March 1989, pp. 21–26.

species or gender but shown to exhibit no effects elsewhere, the results pertaining to the sensitive species or gender typically will be used to develop estimates of human-health risks. For example, if male mice develop cancer from a substance but female mice and rats of both genders do not, then the results from the male mouse often will be used to derive estimates of cancer risks to humans.[42]

Once a positive result has been obtained in an animal bioassay, a substance often will be provisionally classified as a probable human carcinogen. The statistical burden of proof then shifts to the no-effect hypothesis. Because it is logically impossible to prove a negative, however, this practice establishes a virtually irrebuttable presumption in favor of the carcinogenesis hypothesis.

Severe Testing Conditions

Current risk-assessment protocols require the use of very high doses. Unfortunately, high doses are often toxic for reasons unrelated to their capacity to cause cancer. A common procedure is to use what is called the maximum tolerated dose (MTD), which is the most that can be administered to a test animal without causing acute toxicity. At such exposure levels, substances often cause severe inflammation and chronic cell killing. For example, formaldehyde causes nasal tumors in rats when administered in high doses. However, MTD administration severely inflames nasal passage tissues. It is therefore unclear whether the cancers induced are caused by formaldehyde per se or by the toxic effects of high doses.

Results such as these have caused some scientists to question the validity of rodent tests performed at the MTD for estimating human health risks that arise from exposure at low doses.[43] By combining very high doses with highly sensitive test subjects, some animal bioassays are predisposed to discover apparent carcinogenic effects.

Relevance of Animal Bioassay Results

An important reason why animals vary in their sensitivity is that they have different physiologies, metabolic processes, reproductive cycles, and a host of other species-specific characteristics that largely result from unique evolutionary paths. Each of these factors needs to be carefully considered in evaluating the significance of animal data with respect to human health. This is recognized in both the OSTP and EPA guidelines, but it is often neglected when the guidelines are applied to specific substances.[44]

The most important assumption in this regard is that animal test results can be meaningfully extrapolated to humans. A recent study of chemicals tested under the auspices of the U.S. National Toxicology Program shows that this assumption can lead to the erroneous classification of many chemicals as probable human carcinogens.[45] Positive associations have been obtained in either rats or mice for half of 214 chemicals tested. However, results were consistent across these two genetically similar species only 70 percent of the time. If it is assumed that rodent bioassays have the same sensitivity and selectivity with respect to human carcinogens as they do between rodent species, and it is further assumed that 10 percent of all chemicals are in fact human carcinogens, then 27 of every 100 randomly selected chemicals would be misclassified as probable human carcinogens. Only three chemicals would be misclassified as noncarcinogens. Thus, "false positives" would be 9 times more common than "false negatives."[46]

Of course, this ratio of false positives to false negatives reflects highly conservative "upper-bound" assumptions concerning sensitivity and selectivity. Given the high degree of similarity between rats and mice and the limited resemblance between rodents and humans, the sensitivity of rodent bioassays with respect to human carcinogenicity is probably much lower than 70 percent. Furthermore, other research indicates that selectivity may be as low as 5 percent.

[42] See *EPA Carcinogen Risk Assessment Guidelines*, p. 33997 (data from long-term animal studies showing the greatest sensitivity should generally be given the greatest emphasis).

[43] See, e.g., Ames *et al.*, *(op. cit.)*, pp. 276–277.

[44] *OSTP Guidelines*, Guideline 25, p. 10378; *EPA Carcinogen Risk Assessment Guidelines*, p. 34003 (responding to comments on the draft guidelines and affirming agreement with OSTP Guideline 25).

[45] Lester B. Lave, Fanny K. Ennever, Herbert S. Rosenkranz, and Gilbert S. Omenn, "Information Value of the Rodent Bioassay," *Nature*, Vol. 336 (December 15, 1988), pp. 631–633.

[46] *False negatives* occur when a test fails to detect effects when they are in fact present. *Sensitivity* refers to the capacity of a test to minimize false negatives. *False positives* occur when a test appears to detect effects that in fact are absent. *Selectivity* refers to a test's ability to minimize false positives. The 9 to 1 ratio of false positives to false negatives calculated by Lave *et al.* assumes that both selectivity and sensitivity equal about 70%.

Adjusting only for this lower selectivity suggests that false positives are almost 30 times more common than false negatives. This raises serious questions concerning the practical utility of the current approach to animal bioassays for the purpose of quantitative risk assessment.[47]

Other factors should also be considered when relying upon animal bioassay results as the primary basis for quantitative risk assessments. For example, certain substances are toxic or even carcinogenic by one pathway but not by others. Nevertheless, animal bioassay protocols often emphasize the most sensitive pathway. As long as human exposure is likely to arise the same way, then this choice may be reasonable. However, the pathway to which the test species is sensitive sometimes reflects an exposure route that is implausible or irrelevant for humans. For example, formaldehyde causes nasal tumors in rats at 12 times the rate observed in the next most sensitive animal species. This extreme sensitivity may be related to the fact that rats breathe only through the nose.

There may be important differences between animals and humans that make specific tumors irrelevant. For example, some chemicals cause cancer in the zymbal gland of the rat; because humans lack such a gland it is unclear whether these results matter in estimating human health risk. Other substances induce cancer through biochemical mechanisms not found in humans.

A greater controversy surrounds the question whether the same weight should be given to benign and malignant tumors. The scientific consensus is that benign and malignant tumors should be aggregated only when it is scientifically defensible to do so.[48] In practice, however, benign and malignant tumors are routinely aggregated unless a strong case can be made *against* the practice.[49] The difference between these default assumptions is significant: One approach counts only carcinomas that *are* present, whereas the other counts tumors that *might become* carcinomas. In an extreme case, a substance that promotes benign tumors but never causes cancer could be classified as a probable human carcinogen simply because benign and malignant tumors are treated equally.

In addition, tumor incidence is commonly pooled across sites to obtain a total estimate of carcinogenic effects.[50] This implicitly assumes that cancer induction is independent across sites and not the result of either metastasis or the same biological mechanism. Given the extreme sensitivity of test species and the regular use of MTD administration, other explanations for tumors occurring at multiple sites appear just as plausible.

The Choice of Dose-Response Model

No single mathematical model is accepted as generally superior for extrapolating from high to low doses.[51] Consequently, Federal agencies often use a variety of different models. Rather than being a scientific footnote to the risk-assessment process, however, the choice of model is actually an important policy issue. The multistage model appears to be the most commonly used method for estimating low-dose risks from chemicals, and there are two major sources of bias embedded in this choice: its inherent conservatism at low doses, and the routine use of the "linearized" *form* in which the 95 percent upper bound is used instead of the unbiased estimate.

The *multistage model* essentially involves fitting a polynomial to a data set, with the number of "stages" identified by the number of terms in the polynomial. Since animal bioassays rarely have more than three dose levels, it is unusual to see applications of the multistage model with more than two stages. Although the multistage model enjoys some scientific support because it is compatible with multistage theories of carcinogenesis, in practice the model fails to include enough stages, due to the absence of sufficient alternative exposure cohorts.

The multistage model typically yields low-dose risk estimates that are higher than most other models. For example, when five different dose-response models were analyzed in a recent risk assessment of cadmium, estimates of cancer risks at moderate doses varied by a factor of 100. This difference among

[47] Lave *et al.*, *(op. cit.)*, p. 631. Adjusting also for less sensitivity reduces the ratio of false positives to false negatives. For example, if sensitivity is only 10 percent and all other parameters remain unchanged, then this ratio declines to 9.5 to 1. However, this implies that both types of statistical errors are rampant, which raises questions concerning the practical utility of animal bioassays. This is, in fact, precisely the concern raised by Lave *et al.*, *(op. cit.)*, who conclude that such tests are cost-effective investments in information only under extraordinary conditions.

[48] *OSTP Guidelines*, p. 10376.

[49] *EPA Carcinogen Risk Assessment Guidelines*, p. 33997.

[50] *Id.*

[51] *OSTP Guidelines*, Guideline 26, p. 10378; Ames *et al.*, *(op. cit.)*, p. 276.

estimates widened as doses declined toward the very low levels within the range of regulatory concern. At very low doses, two of the five models predicted excess lifetime cancer risks greater than one in one thousand (10^{3}), a risk oftentimes regarded by policymakers as unacceptable. However, two other equally plausible models predicted essentially no excess cancer risk at all. Since none of the five models offers a scientifically superior basis for deriving low-dose risks, the choice of model is therefore a pivotal policy decision. The accepted practice under these circumstances is to develop a subjectively-derived "best" estimate while fully informing decisionmakers as to the extent of uncertainty surrounding it.[52] In the cadmium case, as in most others, this practice was not followed: Estimates of the number of statistical cancers that would be prevented by regulation were presented based only on the multistage model.[53]

The *linearized multistage model (LMS)* is a special version of the multistage model in which the 95 percent upper confidence limit of the linear term is used instead of the unbiased estimate. That is, the model identifies the largest value for the linear term that cannot be rejected at the 95 percent confidence level and uses it in place of the unbiased estimate. Assuming that the model has been correctly specified, there is only a 5 percent chance that the true risk exceeds this level.

The LMS has become the preferred statistical approach because estimates derived from it appear to be more "stable" than estimates obtained from the ordinary multistage model. The "stability" issue originally arose because unbiased estimates of low-dose risks are very sensitive to the maximum-likelihood estimate (MLE) of the value of the linear term. When the MLE of the linear term is positive, it dominates estimated risks at low doses. In some instances, however, the MLE of the linear term is zero, and low-dose risk estimates decline precipitously. Using the 95 percent upper confidence limit ensures that the linear term is always positive, thus eliminating the inherent "instability" of low-dose risk estimates derived from the multistage model. [54]

Another often-cited advantage of the LMS procedure is that it provides a "yardstick" for comparing potencies across chemicals.[55] A uniform risk-assessment procedure such as the LMS, it is argued, enables policymakers to better understand the relative significance of a broad array of chemical hazards and set regulatory priorities accordingly.

Finally, the LMS is often defended on the ground that it is prudent to err on the side of caution when dealing with potentially carcinogenic chemicals. Because the LMS generates upper-bound risk estimates, policymakers can be confident that actual risks are likely to be lower.

None of these purported advantages of the LMS approach has a sound statistical basis. It is a fundamental axiom of statistics that unbiased estimates are generally preferred to biased ones. Using the upper confidence limit instead of the unbiased estimate exaggerates underlying specification errors instead of eliminating them. "Instability" is overcome, but at the cost of greater errors in specification.

The inherent instability of the multistage model reflects a generalized misspecification of dose-response—that is, the real human dose-response relationship is often very different from what the multistage model constrains it to be. The model is extremely sensitive to small differences in observed tumor incidence, which can cause dramatic changes in estimated low-dose risks. The LMS procedure eliminates this sensitivity without remedying the underlying specification error. Proper statistical procedure requires correcting model misspecification, not masking its symptoms behind biased parameter estimates.

The LMS procedure inflates low-dose risk estimates by a factor of two or three when the MLE of the linear term is positive. However, it increases low-dose risk estimates by orders of magnitude when the MLE of the linear term is zero.[56] This means that the degree of hidden conservative bias is substantially greater for what are demonstrably lower risks.

By its very nature, the LMS cannot serve as a useful yardstick for comparing the relative risk of a variety of potential carcinogens. If a given statistical procedure generated identical biases across substances tested,

[52] See, e.g., *OSTP Guidelines*, Guidelines 27, 29, and 31, p. 10378; *EPA Carcinogen Risk Assessment Guidelines*, pp. 33999, 34003.

[53] Occupational Safety and Health Administration, "Occupational Exposure to Cadmium; Proposed Rule," 55 FR 4076 (February 6, 1990).

[54] Albert L. Nichols and Richard J. Zeckhauser, "The Dangers of Caution: Conservatism in Assessment and the Mismanagement of Risk," Chapter 3 in *Advances in Applied Micro-Economics, Volume 4: Risk, Uncertainty, and the Valuation of Benefits and Costs*, V. Kerry Smith, ed., Greenwich, CT: JAI Press, 1986, pp 55–82, esp. pp. 62–63. A nontechnical version of this paper is available by the same authors as "The Perils of Prudence: How Conservative Risk Assessments Distort Regulation," *Regulation*, November/December 1986, pp 13–24.

[55] U.S. Environmental Protection Agency, *A Cancer Risk-Specific Dose Estimate for 2,3,7,8-TCDD*, EPA/600/6-88/007Aa, June 1988 (hereinafter, *Dioxin Risk Assessment*), pp. 45–46.

[56] Nichols and Zeckhauser, *op. cit.*, pp. 62–63.

then it would still yield an accurate rank-ordering of theoretical hazards. Similarly, if the procedure added a stochastic bias from a uniformly distributed random variable, the resulting rank-ordering would still be accurate on an expected-value basis. The problem with the LMS is that it generates biases that intensify with the degree to which the multistage model misspecifies the true dose-response relationship. Even if the multistage model provided an accurate rank-ordering of hazards, the LMS could not do so, because it injects biases that are systematic with statistical misspecification.

The LMS procedure (and the multistage model itself) is also fatally flawed as a yardstick for regulatory priority setting because it fails to take account of human exposure in the calculation of unit risks. Regardless of the procedure's capacity to accurately rank-order *hazards*, failing to adjust unit risks by relative human *exposure* virtually guarantees that regulatory priorities will be misordered. Resources tend to be focused on reducing the greatest theoretical *hazards* rather than the most significant human health *risks*.[57]

Finally, the "margin of safety" argument in favor of the LMS unequivocally contradicts the widely recognized need to distinguish science from policy.[58] The LMS introduces into each risk assessment a conservative bias of varying but unknown magnitude. This practice fundamentally alters regulatory decisionmaking. Instead of leaving policy decisions to policymakers, the LMS disguises fundamental policy decisions concerning the appropriate margin of safety behind the veil of science.

In summary, the LMS cannot be justified as a method of scientific risk assessment. The "yardstick" defense implicitly asserts that scientific advancements in risk-assessment methodology should take a back seat to the preservation of an outdated and misguided statistical procedure. The "margin of safety" argument tacitly usurps from policymakers the authority and responsibility for risk-management decisions. Finally, the statistical "instability" overcome by the LMS is an artifact of specification error, not any scientific theory of human carcinogenesis that warrants the intentional use of biased parameter estimates. The habitual reliance upon either the multistage model or its LMS descendant cannot be supported by sound scientific principles.

Alternative models are available, of course, and they have been applied in many quantitative risk assessments. Because proper model specification is the foundation of applied statistical methodology, alternatives to the multistage model should be expected and encouraged. Indeed, innovation is the hallmark of scientific inquiry; policies that institutionalize any particular model specification effectively stifle scientific advancement.

Unfortunately, models other than the multistage model are often discouraged in practice.[59] Agencies may require substantial scientific evidence in support of an alternative model before allowing it to be used. Alternative models thus face a burden of demonstrating scientific plausibility that the multistage model cannot satisfy. Even in the extraordinary case in which this burden can be satisfied, estimates may be required from the linearized multistage model anyway.[60]

The potential hur in health threat posed by dioxins provides an excelle example of the problem of model selection. Using the ame linearized multistage model, EPA, the Centers for Disease Control (CDC), and the Food and Drug Administration (FDA) have arrived at upper-bound risk estimates that span an order of magnitude.[61] Depending on the data and assumptions used, the linearized multistage model predicts unit risk factors that vary by as much as 1,200, with the

[57] Some scientists have attempted to devise alternative indexes of relative human health risk that explicitly account for variations in human exposure. Ames *et al.*, (*op. cit.*), pp. 272–273, describe one such alternative (the Human Exposure/Rodent Potency index, or HERP) and rep index values for 36 substances. Because the HERP index is based on a relative rather than absolute scale, the distorting effect of conservative biases embedded in the underlying risk assessments has been significantly reduced. Many substances suspected of being environmental carcinogens rank very low on the HERP index, suggesting that regulatory priorities have been seriously misdirected.

[58] See, *e.g.*, *NAS Risk Management Study*, p. 161; *OSTP Risk Assessment Guidelines*, Principle 29, p. 10378; and *EPA Carcinogen Risk Assessment Guidelines*, p. 34001.

[59] See, *e.g.*, Ames *et al.*, (*op. cit.*), p. 276 (continued reliance on linear models despite the accumulation of evidence against linearity); and Lester B. Lave, "Health and Safety Risk Analysis: Information for Better Decisions," *Science*, Vol. 236, April 17, 1987, pp. 291–295, esp. p. 292 (agencies often resist modeling improvements and a that yield lower risk estimates).

[60] *EPA Carcinogen Risk Assessment Guidelines*, pp. 33997–33998. "In the absence of adequate information to the contrary, the linearized multistage procedure will be employed. . . . Considerable uncertainty will remain concerning responses at low doses; therefore, in most cases, an upper-limit risk estimate using the linearized multistage procedure should also be presented."

[61] *Dioxin Risk Assessment Appendix A*, p. 13. Unbiased risk estimates vary by a similar factor.

SL 064751

three risk estimates mentioned earlier clustered at the high end of the range.[62] Risk assessments based on different models have led other governments to establish unit risk factors that are a thousand times less stringent than the most commonly used of these three; one study suggests that this particular estimate overstates the most likely risk estimate by a factor of almost 5,000.[63]

Conversion from Animals to Humans

Once risk has been extrapolated to low doses in rodents, scientists must convert them to human dose-equivalents. The two most common approaches involve the use of body-weight or surface-area conversions, and there are scientific reasons for choosing either approach in individual cases. The surface-area approach leads to estimates of risk that are between 7 and 12 times greater than those based on the body-weight method, depending upon the test species. Despite the ambiguity of the underlying science, the more conservative surface-area method is often applied reflexively.[64]

ISSUES ARISING FROM HUMAN EXPOSURE ESTIMATES

In addition to developing estimates of the dose-response function, agencies must estimate the likely level of human exposure. This section examines some of the issues and problems that arise in conducting an exposure assessment.

It is a generally accepted principle of exposure assessment that estimates should be based on the most likely scenario, with appropriate consideration of uncertainty.[65] Nevertheless, agencies often use conservative assumptions for exposure when real-world data are unavailable. When each of these assumptions tends to overstate likely human risks, the multiplicative effect of even a small overstatement at each stage in an exposure assessment will yield a substantial overestimate of actual exposure. For example, the multiplicative effect of overstating risk by a factor of two at five different points in an exposure assessment will overstate actual risk by a factor of thirty-two.

Worst-Case Environmental Conditions

When data are available they often relate to unusually sensitive environments or highly contaminated conditions. When estimating regional or nationwide exposures, agencies often use data from these local "hot spots" in developing more general national estimates of health risks. However, such data are never representative and estimates extrapolated from them are generally unreliable and misleading.

In addition, chemicals often degrade naturally after they have been released to the environment. In some cases, degradation occurs very quickly, whereas in others the process may take many years or even decades. A common practice in exposure assessment modeling is to assume that exposures remain constant over time—that is, chemicals are assumed never to degrade, or degradation by-products are assumed to pose identical risks.

The Maximum-Exposed Individual

In addition to estimating the amount of a substance that may actually be present in the environment, a risk analysis must also consider the conditions under which humans may be exposed. Actual risks vary considerably depending on location, mobility, and a host of other factors. Nevertheless, estimates often are based on the upper-bound lifetime cancer risk to the maximum-exposed individual (MEI), the hypothetical person whose exposure is greater than all others. Sometimes, risks to the entire population are estimated by assuming that everyone is exposed at the MEI level. Because environmental regulations are often justified using MEI-based risk assessments, actual risks may be substantially lower than what decisionmakers and the general public perceive them to be.

---

[62] *Dioxin Risk Assessment*, pp. 46–49. 10^{-6} risk-specific doses (RsDs) derived from the linearized multistage model span the range from 0.001 to 1.2 picogram/kg/day. The RsDs of EPA, CDC, and FDA are 0.006, 0.03, and 0.06 pg/kg/day, respectively.

[63] *Dioxin Risk Assessment*, p. 4.

[64] *EPA Carcinogen Assessment Guidelines*, p. 33998. "EPA will continue to use this [surface area] scaling factor unless data on a specific agent suggest that a different scaling factor is justified."

[65] EPA guidance documents have historically called for unbiased estimates of exposure. See, e.g., U.S. Environmental Protection Agency, "Guidelines for Exposure Assessment," 50 FR 34042–34054 (September 24, 1986, hereinafter, *EPA Exposure Assessment Guidelines*); U.S. Environmental Protection Agency, *Superfund Public Health Evaluation Manual*, OSWER Directive 9285.4–1, October 1986, and U.S. Environmental Protection Agency, *Superfund Exposure Assessment Manual* (Revised Draft), OSWER Directive 9285.5–1, December 1986. EPA recently abandoned the calculation of unbiased exposure estimates for Superfund sites on the ground that it was insufficiently conservative. EPA's new protocol requires the estimation of "reasonable maximum exposure" instead of the average and upper-bound estimates. Reasonable maximum exposure constitutes a new term of art that EPA intends to be "well above the average case" but not as extreme as the upper-bound. It provides a new opportunity for embedding conservative assumptions into exposure assessment and exaggerating estimates of actual human-health risk at Superfund sites. See *Risk Assessment Guidance for Superfund, Volume I: Human Health Evaluation Manual (Part A), Interim Final*, EPA/540/1-89/002, December 1989, Chapter 6, pp. 5, 47–50.

In developing the MEI risk level, analyses invariably assume that the level of exposure is continuous over a 70-year lifetime. This assumption overstates actual risks, because people are mobile, encounter a constantly changing portfolio of daily risks to life and health, and can take actions that reduce risk.

Assumptions vs. Real-World Exposure Data

The thread that connects these exposure assessment issues is that simple constructs which overstate exposure are typically used in lieu of real-world data, often because such data are unavailable. The risk estimates generated by these models depend on the validity of their assumptions; even small biases in exposure assessment assumptions can result in a substantial overstatement of risk.

For example, regulatory agencies may not have statistically reliable real-world data on pesticide residues in agricultural products. They also may not know the proportion of a given crop that has been treated with a particular pesticide. A common resolution of these uncertainties is to assume that residues are equal to the regulatory "tolerance"—the maximum level allowed to be present in food sold in interstate commerce—and that 100 percent of the relevant crop has been treated. Both assumptions overstate actual exposure, but are encouraged by agency guidance as a way to instill conservatism in risk assessment.[66] When data are available, however, the extent of this conservative bias becomes evident. In a recent special review for the pesticide Captan, for example, EPA reduced its earlier upper-bound lifetime cancer risk estimate by two orders of magnitude when it replaced the original conservative assumptions with real-world data. Even with these improvements, EPA still reported that upper-bound risks were probably overstated. For example, field tests were performed based on applications at the maximum legal rate and as close to harvest as the label permits. Similarly, feeding studies assumed that animal diets were dominated by feedstuffs that happened to contain high residues relative to other feedstuffs, such as almond hulls and raisin waste. As EPA noted, even if these assumptions accurately represented typical animal diets, they would do so only for portions of California where these crops are grown; nationwide extrapolations based on these "hot-spots" would very likely overstate exposure.[67] Since two of the highest product-specific risks were attributed to milk and meat, these remaining conservative biases can be expected to be significant.

IMPLICATIONS OF CONSERVATIVE RISK ASSESSMENT FOR RISK MANAGEMENT AND REGULATORY DECISIONMAKING

The primary purpose of risk assessment is to provide data as a basis for risk management decisions. Providing useful data requires the synthesis of information concerning risks and exposure levels into a coherent package that can be used to develop regulatory options. Decisionmakers then can use these risk estimates in evaluating regulatory alternatives. Unfortunately, the way in which risk information is characterized tends to overstate risks, making them appear much greater than they are likely to be. As a result, decisionmakers may make regulatory choices that are very different from the ones they would make if they were fully informed.

Quantification of Uncertainty

In accordance with the recommendations of the National Academy of Sciences, the *OSTP Guidelines* explicitly call for the quantification of uncertainty, particularly as it arises in the selection of dose-response models and exposure assumptions.[68] Unfortunately, Federal regulatory proposals that utilize risk assessment rarely provide this information, nor do they analyze the implications of uncertainty for decisionmaking. Instead, many risk assessments only identify a lifetime upper-bound level of risk.[69]

The differences between upper-bound and expected-value estimates may be considerable. As we indicated earlier, the upper-bound risk estimate for dioxin may be 5,000 times greater than the most likely estimate. Plausible risk estimates for perchloroethylene (the primary solvent used in dry cleaning) vary by a factor of about 35,000.[70]

In some instances, decisionmakers may not be informed that risk estimates differ because of policy choices hidden in the risk-assessment methodology. In EPA's proposed rule limiting emissions from coke

[66] *EPA Exposure Assessment Guidelines*, p. 34053. "When there is uncertainty in the scientific facts, it is Agency policy to err on the side of public safety."

[67] See, e.g., U.S. Environmental Protection Agency, "Captan: Intent to Cancel Registrations; Conclusion of Special Review," 54 FR 8127–8126 (February 24, 1989).

[68] *OSTP Guidelines*, (Guideline 27), p. 10378.

[69] See, e.g., *EPA Carcinogen Risk Assessment Guidelines*, p. 33998.

[70] Nichols and Zeckhauser, *(op. cit.)*, pp. 64–65.

ovens, for example, cancer risks were estimated based on the LMS model—a model that is designed to yield upper-bound estimates of risk. In previous rules involving similar types of risks, however, EPA used the unbiased maximum likelihood estimate. To the extent that decisionmakers were not informed that the higher estimate of risk was largely due to a different low-dose extrapolation procedure, regulatory decisions based on this risk assessment were likely to reflect misunderstanding rather than science.[71]

Plausible estimates of likely cancer risk can often be found buried in regulatory background documents. However, *Federal Register* rulemaking notices seldom present such estimates alongside upper-bound estimates. This practice overstates baseline human health threats, as well as the amount of risk reduction that may be accomplished by regulation. Policymakers and the public are misled because they typically see only the upper-bound estimates of the threat.

The prevalent Federal agency practice is to calculate the benefits of Federal regulatory initiatives based solely on upper-bound estimates of risk and exposure. In a recent proposal to reduce occupational exposure to cadmium, for example, the Occupational Safety and Health Administration (OSHA) developed risk estimates based on five alternative models for animal data, and two alternative models for human data. Across these seven data/model combinations, estimated excess lifetime cancer risk at the least stringent of the two proposed exposure standards varied from 0 to 153 cases per 10,000 workers occupationally exposed for 45 years. OSHA based its proposed exposure standards on *one* of these data/model combinations—the multistage model applied to animal data. This data/model combination predicted an excess lifetime cancer risk of 106 per 10,000 exposed workers, and was used to estimate aggregate cancer incidence and the risk-reduction benefits attributable to the new standard. Uncertainties in the underlying risk assessment, which span several orders of magnitude, were not carried forward through the exposure assessment and benefit calculation stages. This analytic error effectively obscured the uncertainty surrounding the true incidence of cadmium-induced lung cancer, and resulted in benefit estimates that may exceed actual reductions in occupational illness by several orders of magnitude.[72]

Misordered Priorities, Perverse Outcomes

Logically, one would expect that the routine overstatement of likely risks would lead to inefficient regulatory choices. Decisionmakers, convinced that a certain substance or activity poses a significant threat to public health, might well take actions that they would otherwise resist. Alternatively, they might take actions that address the wrong real-life risks.

To the extent that risk assessments differ in the degree to which they adopt conservative assumptions, it is difficult to determine which activities pose the greatest risks and hard to establish reasonable priorities for regulatory action. Because conservatism in risk assessment is especially severe with respect to carcinogens, it is reasonable to expect that other health and safety risks tend to receive relatively less attention and weight. As a result, society may actually incur greater total risk, because of misordered priorities caused by conservative biases in cancer risk assessment.[73]

A perverse and unfortunate outcome of using upper-bound estimates based on compounded conservative assumptions is that the practice may actually increase risk, even in situations where cancer is the only concern. Regulatory actions taken to address what are in fact insignificant threats may implicitly tolerate or ignore better known, documented risks that are far more serious. For example, before it was banned, ethylene dibromide (EDB) was used as a grain and soil fumigant to combat vermin and molds. Vermin transmit disease, and molds harbor the natural and potent carcinogen aflatoxin B. The estimated human cancer risk from the aflatoxin contained in one peanut butter sandwich is about 75 times greater than a full day's dietary risk from EDB exposure. On this basis alone, it might have been appropriate to accept a small increase in cancer risk from EDB to reduce the much larger cancer risk from aflatoxin. By eliminating the relatively small hazard from EDB, Federal risk managers may have intensi-

[71] Letter from Wendy Gramm (Administrator of the Office of Information and Regulatory Affairs) to Lee Thomas (Administrator of the Environmental Protection Agency), August 12, 1986, p. 3.

[72] Occupational Safety and Health Administration, "Occupational Exposure to Cadmium; Proposed Rule," 55 *Federal Register* 4076, 4080, 4093.

[73] This is precisely the policy issue raised by Nichols and Zeckhauser, (*op. cit.*), pp. 69–71, who note that EPA's 1985 decision to limit lead in gasoline was threatened by concerns about potential increases in benzene exposure. Any tradeoff between lead and benzene risks would have been biased against lead, as estimates of benzene risks are more conservative simply because it is a carcinogen, whereas lead is not.

fied the relatively potent threat of aflatoxin associated with an increase in the prevalence of mold contamination.[74]

The emphasis on risks faced by the maximum-exposed individual may also cause a perverse result by increasing overall population risks. For example, EPA's proposed regulation of the disposal of sewage sludge would probably create more public health risk than it eliminates. The proposal outlines a regulatory scheme that would shift disposal from generally safe practices to relatively risky alternatives. Thus, setting sludge quality standards to achieve an MEI upper-bound lifetime cancer risk of one in 100,000 (10^{-5}) would prevent 0.2 statistical cancer cases resulting from monofilling and land application. However, it would cause 2.0 additional statistical cancers by forcing a shift away from these disposal approaches toward incineration.[75]

These problems can be addressed by providing decisionmakers with the full range of information on the risks of a substance or an activity. Thus, decisionmakers should be given the likely risks as well as estimates of uncertainty and the outer ranges of the potential risk. Then, if regulatory decisionmakers want to choose a very cautious risk management strategy, they can do so and a margin of safety can be applied explicitly in the final decision. This approach is superior to one in which the expected risk and an unknown margin of safety are hidden behind the veil of a succession of upper-bound estimates adopted at key points in the risk-assessment process.

The public and affected parties also benefit from knowing both the expected risk and the margin of safety rather than being given upper-bound estimates that are probably very different from actual risks. People are likely to have a better intuitive understanding of the significance of averages than they have of unlikely extremes. To the extent that a margin of safety is appropriate—perhaps to protect unusually sensitive subpopulations—the magnitude of this margin can be more readily communicated if made explicit. In addition, providing information in this way should help improve public confidence in quantitative risk assessment as the basis for decisionmaking.

AVOIDING CONSERVATIVE BIASES IN RISK ASSESSMENT

Risk assessment remains a powerful and useful scientific tool for estimating many of the risks that arise in a technologically advanced society. Unfortunately, it is also susceptible to hidden biases that may undermine its scientific integrity and the basis for policymakers' reliance on such information in risk management decisions. For policymakers and the public to continue to rely on risk assessment in the development of regulatory initiatives, a renewed effort must be made to separate science from policy and provide risk information that is both meaningful and reliable.

Expected Value Estimates

Perhaps the most important current need in regulatory decisionmaking is for carefully prepared and scientifically credible estimates of the likely risks involved. Relying on worst-case analysis based on extremely conservative risk assessment and exposure models leads to widespread misunderstanding on the part of both Government officials and individual citizens. Decisionmakers at all levels need unbiased and impartial risk information so they can focus their attention on significant problems and avoid being distracted by minutiae.[76]

Weight-of-Evidence Determinations

Similar procedures are needed for assigning weights to each relevant study in the risk-assessment literature. Current practice gives undue weight to studies that show positive relationships. Resulting risk classifications are thus conservatively biased estimates derived from samples of similarly biased bservations.

Full Disclosure

Efficient and responsible decisionmaking requires that policymakers and the public be fully informed about the implications of the regulatory alternatives among which they must choose. Meeting this requirement demands a careful discrimination between science and policy. When risk estimates depend on assumptions and judgments instead of data, the meaning and implications of these nonscientific parameters must be clearly articulated.

Avoiding Perverse Outcomes

Careful attention needs to be paid to the likely results of regulatory alternatives, with an eye toward avoiding choices that have the perverse effect of increasing net risk. All human activity involves risk.

---

[74] Ames *et al.*, *(op. cit)*, p. 273.

[75] U.S. Environmental Protection Agency, "Standards for the Disposal of Sewage Sludge; Proposed Rule," 54 FR 5746–5902 (February 6, 1989).

[76] Nichols and Zeckhauser, *op. cit.*, pp. 72–76.

Decisionmakers need to be sure that specific actions taken in the name of risk-reduction in one area do not make matters worse elsewhere. Quantitative risk assessment can help in this regard so long as the methods applied are not inherently biased in a way that undermines comparisons across alternatives, each of which entails some degree of risk.

Our discussion has covered only the highlights of risk-assessment methods, yet we have identified several independent places at which conservative assumptions are commonly used. Individually, each of these assumptions might appear to be prudent responses to scientific uncertainty. In combination, however, they result in a distortion equal to the product of the individual conservative biases. To illustrate, suppose that there are ten independent steps in a risk assessment and prudence dictates assumptions that in each instance result in risk estimates two times the expected value. Such a process would yield a summary risk estimate that is more than 1,000 times higher than the most likely risk estimate. Because there are usually many more than ten steps, and many of them will incorporate conservative biases that exceed an order of magnitude, risk estimates based on such practices will often exceed the most likely value by a factor of one million or more.

When risk assessments contain hidden value judgments, their scientific credibility is inevitably compromised. To the extent that policymakers and the public fail to understand the magnitude of the margin of safety embedded in quantitative risk assessments, policy choices are distorted from the course that would have been selected if decisionmakers had been better informed of the actual risks. *Ironically*, these policy decisions may actually increase total societal risk. Too much attention is focused on relatively small hazards that have been exaggerated by conservative risk assessments, leaving alone larger risks that have been estimated using unbiased procedures.

Information as an Alternative Regulatory Strategy

Federal regulation was initiated to deal with economic problems caused by monopoly and so-called "excess competition." Subsequent events have shown that, in general, economic regulation—fixing prices, establishing restrictive terms of trade, and erecting barriers to entry—is usually inefficient and detrimental to innovation. In response to these lessons, Federal regulation of this type has been under increasing criticism. As indicated above, however, much more needs to be done to reform economic regulation and restore competition.

Federal regulation has more recently been initiated to deal with what economists call externalities, situations in which participants in voluntary market transactions do not bear the full costs or capture all of the benefits of these exchanges. Common examples of externalities include environmental pollution and traffic congestion, common property resources such as fisheries and public forests, and "public goods" such as basic scientific research. In each of these instances, regulation may be an appropriate mechanism to modify or restore distorted market processes, or to stablish markets where heretofore they have not existed, to maximize net social benefits (including environmental, health, and safety benefits). The key ingredient is the determination that existing markets are, in some significant manner, failing to perform efficiently.

The traditional regulatory approach to externalities has been the promulgation of standards. Because this approach often remedies existing externalities by creating new ones, economic incentive instruments are becoming an increasingly popular alternative to standards. The principal attraction of economic incentives is that they rely on market forces rather than attempt to suppress them.

This section explores another alternative regulatory strategy—the production, provision, or mandated disclosure of information. The first subsection briefly summarizes the economics of information as it relates to regulatory decisionmaking. Three points stand out in this discussion. First, because information is costly to acquire and the capacity to process it is limited, there is an optimal level of information for every market transaction. Second, differences in the amount and quality of information between buyers and sellers are normal and do not necessarily indicate market failure. Rather, these differences generally reflect variations in the costs and benefits that are attributable to information. Third, competitive markets provide powerful incentives for buyers and sellers to reveal relevant information. Market processes, not government regulations, provide the dominant motivation for generating, acquiring, and disclosing information. The role of government regulation thus should be to supplement these processes when they prove to be inadequate, not to supplant them when they work well.

The second subsection identifies three rationales for government intervention in the production or mandated disclosure of information. Two of these are economic—the public-good character of some types of

APPENDIX V

Regulatory Impact Analysis Guidance

A Regulatory Impact Analysis (RIA) should demonstrate that a proposed regulatory action satisfies the requirements of Section 2 of Executive Order No. 12291. To do so, it should show that:

- There is adequate information concerning the need for and consequences of the proposed action;
- The potential benefits to society outweigh the potential costs; and
- Of all the alternative approaches to the given regulatory objective, the proposed action will maximize net benefits to society.

The fundamental test of a satisfactory RIA is whether it enables independent reviewers to make an informed judgment that the objectives of Executive Order No. 12291 are satisfied. An RIA that includes all the elements described below is likely to fulfill this requirement. Although variations consistent with the spirit and intent of the Executive Order may be warranted for some rules, most RIAs should include these elements.

The guidance in this document is not in the form of a mechanistic blueprint, for a good RIA cannot be written according to a formula. Competent professional judgment is indispensable for the preparation of a high-quality analysis. Different regulations may call for very different emphases in analysis. For one proposed regulation, the crucial issue may be the question of whether a market failure exists, and much of the analysis may need to be devoted to that key question. In another case, the existence of a market failure may be obvious from the outset, but extensive analysis might be necessary to estimate the magnitude of benefits to be expected from proposed regulatory alternatives. The amount of analysis (whether scientific, statistical, or economic) that a particular issue requires depends on how crucial that issue is to determine the best alternative and on the complexity of the issue.

Regulatory analysis inevitably involves uncertainties and requires informed professional judgments. Whenever an agency has questions about such issues as the appropriate analytical techniques to use or the alternatives that should be considered, it should consult with the Office of Management and Budget as early in the analysis stage as possible.

This document is written primarily in terms of proposed regulatory changes. However, it is equally applicable to the review of existing regulations. In the latter case, the regulation under review should be compared to a baseline case of no regulation and to reasonable alternatives.

Elements of a Regulatory Impact Analysis

Preliminary and final Regulatory Impact Analyses of major rules should contain five elements. They are: (1) a statement of the potential need for the proposal, (2) an examination of alternative approaches, (3) an analysis of benefits and costs, (4) the rationale for choosing the proposed regulatory action, and (5) a statement of statutory authority. These elements are explained in Sections I–V below.

I. STATEMENT OF POTENTIAL NEED FOR THE PROPOSAL

In order to establish the potential need for the proposal, the analysis should demonstrate that (a) market failure exists that is (b) not adequately resolved by measures other than Federal regulation.

A. Market Failure

The analysis should determine whether there exists a market failure that is likely to be significant. Once such market failure has been identified, the analysis should show how adequately the regulatory alternatives to be considered address the specified market failure. The three major types of market failure are externality, natural monopoly, and inadequate information.

1. *Externality.* An externality occurs when one party's actions impose uncompensated benefits or costs on another outside the marketplace. Environmental problems are a classic case of externality. Another example is the case of common property resources that may become congested or overused, such as fisheries or the broadcast spectrum. A third example is a "public good," such as defense or scientific research, whose distinguishing characteristic is that it is inefficient, or impossible, to exclude individuals from its benefits.

2. *Natural monopoly.* Natural monopoly exists where a market can be served at lowest cost only if production is limited to a single producer. Local telephone, gas, and electricity services are examples.

3. *Inadequate information.* The optimum, or ideal, level of information is not necessarily the maximum possible amount, because information, like other goods, should not be produced when the costs of doing so exceed the benefits. The free market does not

necessarily supply an optimal level of information, because information, once generated, can be disseminated at little or no marginal cost, and because it is commonly infeasible to exclude nonpayers from reaping benefits from the provision of information by others. Where market failure due to inadequate information is the rationale for government intervention, a regulatory action to improve the availability of information will ordinarily be the preferred alternative.

The current state of knowledge about the economics of information is not highly developed. Therefore, regulatory intervention to address an information problem should only be undertaken where there is substantial reason to believe that private incentives to provide information are seriously inadequate and that the specific regulatory intervention proposed will provide net benefits for society.

In many circumstances, the availability of information, while perhaps not optimal, is reasonably adequate, so that attempts to regulate information are as likely to make things worse as to make them better. Information about a particular characteristic of a product, for example, would be reasonably adequate if buyers could determine the existence of the characteristic by inspection of the product before purchase or (in the case of a frequently purchased product) by use of the product. Even if the characteristic could not be determined by buyers, government intervention would not be warranted where sellers have incentives to reveal the existence of the characteristic to buyers. Sellers will have substantial incentives to supply information about any characteristic that is important to buyers and valued positively by them, particularly if the level of the characteristic varies between the products of one seller and another. In these circumstances, sellers whose products rank highly in the valued characteristic can increase their sales by informing buyers of the superiority of their products. If the level of the characteristic does not vary between the products of one seller and another, individual sellers have less incentive to inform buyers about the characteristic. Even so, the incentives of individual sellers or of a trade association to supply information may be substantial.

Sellers are least likely to supply adequate information about a particular characteristic of their product where the characteristic is negatively valued by consumers and the level of the characterstic does not vary between the products of one seller and those of another (e.g., cholesterol in eggs). Even in such circumstances, substantial information about the characteristic may be available to buyers. For example, sellers of rival products may supply the information (e.g., while sellers of butter may have no incentive to tell buyers about cholesterol in butter and its possible consequences, sellers of margarine do have such an incentive). Where the negative characteristic involves a health or safety hazard, the threat of future product liability lawsuits may give sellers adequate incentives to reveal information about the potential hazard. News media, consumer groups, public health agencies, and similar services may supply information not supplied by sellers. In summary, while it is possible to identify situations in which market failure due to inadequate information is more likely to warrant regulatory intervention, each situation must be examined on a case-by-case basis.

There should be a presumption against the need for certain types of regulatory actions, except in special circumstances. A particularly demanding burden of proof is required to demonstrate the potential need for any of the following types of regulations:

- Price controls in competitive markets
- Controls on production or sales in competitive markets
- Mandatory uniform quality standards for goods or services, unless they have hidden safety or other defects and the problem cannot be adequately dealt with by voluntary standards or information disclosing the hazard to potential buyers or users
- Controls on entry into employment or production, except (a) where indispensable to protect health and safety (e.g., FAA tests for commercial pilots) or (b) to manage the use of common property resources (e.g., fisheries, airwaves, Federal lands, and offshore areas).

B. Alternatives to Federal Regulation

Even where a market failure exists, there may be no need for Federal regulatory intervention if other means of dealing with the market failure resolve the problem adequately or better than the proposed Federal regulation would. Among the alternative means that may be applicable are the judicial system (particularly liability cases to deal with health and safety), antitrust enforcement, and workers' compensation systems.

An important alternative that may often be relevant is regulation at the State or local level. In determining whether there exists a potential need for a proposed Federal regulation, the analysis should examine whether regulation at the Federal level is more appropriate than regulation at the State or local level. This analysis may support regulation at the Federal level where rights of national citizenship (such as legal equality among the races) or considerations of interstate commerce are involved. If interstate commerce is involved the analysis should attempt to determine whether the burdens on interstate commerce arising from different State and local regulations are so great that they outweigh the advantages of diversity and local political choice. In some cases, the nature of the market failure may itself suggest the most appropriate governmental level of regulation. For example, pollution that spills across state lines (such as acid rain whose precursors are transported widely in the atmosphere) is probably best controlled by Federal regulation, while localized pollution (such as garbage truck noise) is probably more efficiently handled by local government regulation.

In general, because demands among localities for different governmental services differ and because competition among governmental units for taxpayers and citizens may encourage efficient regulation, the smallest unit of government capable of correcting the market failure should be chosen. This must, however, be balanced against the possibility of higher costs because national firms would be required to comply with more than one set of regulations and because administering similar regulations in more than one governmental unit involves some costs of duplication. Thus, some analysis may be necessary to determine which level of government can most efficiently regulate a specific market failure.

If the analysis does suggest a potential need for a Federal action, it should also consider alternatives of nonregulatory Federal measures. For example, as an alternative to requiring an action or the use of a particular product, it may be more efficient to subsidize it. Similarly, a fee or charge may be a preferable alternative to banning or restricting a product or action. An example would be an effluent discharge fee, which has been recommended as an efficient way to limit pollution, because it causes pollution sources with different marginal costs of abatement to control effluents in an efficient manner. In addition, legislative measures that make use of economic incentives, such as changes in insurance provisions or changes in property rights, should be considered.

II. AN EXAMINATION OF ALTERNATIVE APPROACHES

The RIA should show that the agency has considered the most important alternative approaches to the problem and must provide the agency's reasoning for selecting the proposed regulatory change over such alternatives. Ordinarily, it will be possible to eliminate some alternatives by a preliminary analysis, leaving a manageable number of alternatives to be evaluated by quantitative benefit-cost analysis according to the principles to be described in Section III. The number and choice of alternatives to be selected for detailed benefit-cost analysis is unavoidably a matter of judgment. There must be some balance between thoroughness of analysis and practical limits to the agency's capacity to carry out analysis.

Alternative regulatory actions that should be explored include the following:

1. *More performance-oriented standards for health, safety, and environmental regulations.* Performance standards are generally to be preferred to engineering or design standards because they allow the regulated parties to achieve the regulatory objective in the most cost-effective way. In general, a performance standard should be preferred wherever that performance can be measured or reasonably imputed. Performance standards should also be applied as broadly as possible without creating too much variation in regulatory benefits; for example, by setting emission standards on a plant-wide or firm-wide basis rather than source by source. It is misleading and inappropriate, however, to characterize a standard as a performance standard if it is set so that there is only one feasible way to meet it; as a practical matter, such a standard is a design standard.

2. *Different requirements for different segments of the regulated population.* For example, there might be different requirements for large and small firms. If such a differentiation is made, it should be based on perceptible differences in the costs of compliance or in the benefits to be expected from compliance. For example, some worker safety measures may exhibit economies of scale, that is, lower costs per worker protected in large firms than in small firms. A heavier burden should not be placed on one segment of the regulated population on the grounds that it is better able to afford the higher cost; this is a sure formula for loading disproportionate costs on the most productive sectors of the economy.

3. *Alternative levels of stringency.* In general, both the benefits and costs associated with a regulation will increase with the level of stringency (although costs will eventually increase more rapidly than benefits). It is important to consider alternative levels of stringency to better understand the relationship between stringency and benefits and costs. This approach will increase the information available to the decisionmaker on the option that maximizes net benefits.

4. *Alternative effective dates of compliance.* The timing of a regulation may also have an important effect on its net benefits. For example, costs of a regulation may vary substantially over different compliance dates for an industry that requires a year or more to plan its production runs efficiently. In this instance, a regulation whose requirements provide

SL 064759

sufficient lead time is likely to achieve its goals at a much lower overall cost than a regulation that is effective immediately.

5. *Alternative methods of ensuring compliance.* Compliance alternatives include the appropriate entity (local, State, or Federal) enforcing compliance, whether compliance is enforced by on-site inspection or periodic reporting, and structuring compliance penalties so that they provide the most appropriate incentives.

6. *Informational measures.* Measures to improve the availability of information include government establishment of a standardized testing and rating system (the use of which could be made mandatory or left voluntary), mandatory disclosure requirements (e.g., by advertising, labeling, or enclosures), and government provision of information (e.g., by government publications, telephone hot-lines, or public interest broadcast announcements). If intervention is necessary to address a market failure arising from inadequate information, informational remedies will generally be the preferred approaches. As an alternative to a mandatory standard, a regulatory measure to improve the availability of information has the advantage of being a more market-oriented approach. Thus, providing consumers information about concealed characteristics of consumer products gives consumers a greater choice than banning these products (for example, consumers are likely to benefit more from information on energy efficiency than from a prohibition on sale of appliances or automobiles falling below a specified standard of energy efficiency).

Except for prohibiting indisputably false statements (whose banning can be presumed beneficial), specific informational measures must be evaluated in terms of their benefits and costs. Paradoxically, the current state of knowledge does not generally permit the benefits and costs of informational remedies to be measured very accurately. Nonetheless, it is essential to consider carefully the costs and benefits of alternative informational measures, even if they cannot be quantified very precisely. Some effects of informational measures can easily be overlooked. For example, the costs of a mandatory disclosure requirement for a consumer product include not only the obvious cost of gathering and communicating the required information, but also the loss of any net benefits of information displaced by the mandated information, the cost of any inaccurate consumer interpretation of the mandated information, and any inefficiencies arising from the incentive that mandatory disclosure of a particular characteristic gives to producers to overinvest in improving that specific characteristic of their products.

Where information on the benefits and costs of alternative informational measures is insufficient to provide a clear choice between them, as will often be the case, the least intrusive alternative, sufficient to accomplish the regulatory objective, should be chosen. For example, it will often be sufficient for government to establish a standardized testing and rating system without mandating its use, because firms that score well according to the system will have ample incentive to publicize the fact.

7. *More market-oriented approaches.* In general, alternatives that provide for more market-oriented approaches, with the use of economic incentives replacing command-and-control requirements, should be explored. Market-oriented alternatives that may be considered include fees, subsidies, penalties, marketable rights or offsets, changes in liabilities or property rights, and required bonds, insurance or warranties (in many instances, implementing these alternatives will require legislation).

III. ANALYSIS OF BENEFITS AND COSTS

A. General Principles

The preliminary analysis called for by Sections I and II should have narrowed the number of alternatives to be considered by quantitative benefit-cost analysis to a workable number. Ordinarily, one of the alternatives will be to promulgate no regulation at all, and this alternative will commonly serve as the base from which increments in benefits and costs are calculated for the other alternatives. Even if alternatives such as no regulation are not permissible statutorily, it is often desirable to evaluate the benefits and costs of such alternatives to determine if statutory change would be desirable. Departments and agencies bear a similar burden when they perform environmental impact statements in which alternatives that lie outside their statutory authority must be considered.

In some cases, the desirability of specific alternatives outside the scope of the agency's regulatory authority may be determined by use of basic economic concepts in light of the principles enumerated in Section I. In other instances, however, only a quantitative benefit-cost analysis can resolve the question, and such alternatives will need to be included in the analysis of this section. In addition, alternative forms of agency regulation will need to be evaluated by quantitative benefit-cost analysis.

1. *Evaluation of Alternatives.* Except where prohibited by law, the primary criterion for choice among alternatives is expected net benefit (benefits minus costs). Other criteria may sometimes produce equivalent results, but they must be used with care to avoid

the potentially serious pitfalls to be explained in Part B of this section and in Section IV. Both benefits and costs should be expressed in discounted constant dollars. Appropriate discounting procedures are discussed in the following section.

The distinction between benefits and costs in benefit-cost analysis is somewhat arbitrary, since a positive benefit may be considered a negative cost, and vice versa, without affecting the net benefit (benefits minus costs) decision criterion. This implies that the considerations applicable to benefit estimates also apply to costs and vice versa. The different issues are considered separately under benefits or costs in Sections B and C below according to where they most often arise.

If the proposed regulation is composed of a number of distinct provisions, it is important to evaluate the benefits and costs of the different provisions separately. The interaction effects between separate provisions (such that the existence of one provision affects the benefits or costs arising from another provision) may complicate the analysis but does not eliminate the need to examine provisions separately. In such a case, the desirability of a specific provision may be appraised by determining the net benefits of the proposed regulation with and without the provision in question. Where the number of provisions is large and interaction effects are pervasive, it is obviously impractical to analyze all possible combinations of provisions in this way. Some judgment must be used select the most significant or suspect provisions for such analysis.

2. *Discounting*. The monetary values of benefits and costs occurring in different years should be discounted to their present values so that they are comparable. This is not the same as correcting for inflation. An inflation adjustment is made with a price index, whereas discounting to present value is done with a discount rate. Benefits and costs expressed in constant (i.e., unaffected by inflation) dollars must further be discounted to present values before benefits and costs in different years can be added together to determine overall net benefits. As an equivalent alternative to discounting nonmonetized benefits, the RIA may use the discount rate to annualize (amortize) costs over a period that corresponds to the occurrence of the benefits. Regardless of the discounting procedure selected, the RIA must contain a schedule indicating when the benefits and costs occur.

Discounting takes account of the fact that resources (goods or services) in a given year are worth more than identical resources in a later year. The underlying reason for this is that resources can be invested so as to return more resources later. Partly because of this productivity of investment, individuals value consumption in earlier years higher than consumption in later years.

Modern analysis of discounting for public programs stresses the distinction between two rates of return:

- The *before-tax* rate, also known as the opportunity cost of capital. This is the real rate of return to marginal private investments. Estimates of the opportunity cost of capital in the U.S. economy vary substantially. The 10 percent discount rate specified by OMB Circular A-94 for use in evaluating government programs is intended to represent the opportunity cost of capital.
- The *after-tax* rate, also known as the consumption rate of interest. This represents the rate at which consumers would be willing to exchange present for future consumption, that is, the rate at which consumers must be compensated for postponing their consumption. As with the opportunity cost of capital, alternative estimates of the consumption rate of interest vary significantly. A rate of 4 percent is reasonably representative of the range of alternative estimates and consistent with a 10 percent before-tax rate of return.

The basic concept underlying the academic literature on public-sector discounting is that economic welfare is ultimately determined by consumption and only indirectly by investment. Therefore, the value of investment must be measured by the value of the subsequent increase in consumption it permits. Any effect that a government program has on investment must be converted to an equivalent time-stream of consumption before being discounted. In practice, this results in a complex procedure that uses the before-tax and after-tax discount rates, a "shadow price of capital," and the impacts of benefits and costs on investment. It is recommended that agencies continue to use the well-understood procedure of discounting by a single rate (as specified by OMB Circular A-94) and, when appropriate, perform additional analysis using the more complex shadow-price-of-capital methodology.

There are two circumstances when it is important to perform sensitivity analysis using the shadow price of capital approach:

(a) Where the costs of the regulation are almost entirely current costs borne by consumers. In such circumstances, a low rate close to 4 percent is called for. (This assumes, as is normally the case, that the benefits are all in the form of disposable income or other benefits directly to individuals.)

(b) Where some of the costs are capital costs financed out of saving *and* there is a long period between the time when most costs are incurred and the time when most benefits accrue. In general, the *smaller* the fraction of costs that are capital costs financed out of saving and the *longer* the time period between costs and benefits, the greater the likelihood that the shadow price of capital approach will be correct.

It is conceptually incorrect to adjust the discount rate as a device to account for the uncertainty of expected future benefits and costs. This procedure will virtually never lead to a correct adjustment of benefits and costs. Therefore, risk and uncertainty should be dealt with according to the principles in Section 3 below and not by changing the discount rate.

3. *Treatment of Risk and Uncertainty.* Where uncertainties exist about important parameters affecting the expected benefits or costs of an alternative under consideration, it is essential to carry out a *sensitivity analysis* to determine the effect on net benefits of plausible variations in the value of the parameters. One form of sensitivity analysis involves calculation of the "switch-point" value of the parameter under examination, that is, the value of the parameter at the break-even point at which the net-benefit decision criterion switches over from favoring one alternative to favoring another. When this break-even point of the parameter value is determined, the analysis may then consider the probability that the true parameter value is above or below the break-even value. For example, if the major uncertainty about a proposed regulation were its cost, the analysis could calculate how high the cost would need to be in order to reduce the net benefit of the proposal to zero. If it is judged to be highly unlikely that the actual cost would be that high or higher, it may be concluded that the choice of the proposed alternative is not sensitive to uncertainties about its cost.

A primary objective of sensitivity analysis is to identify where additional analysis may be most needed. If the choice of a specific regulatory action is sensitive to alternative parameter values that are about equally likely to be true, more research to better determine the true parameter value could be very valuable.

Wherever parameter estimates are uncertain, for either benefits or costs, expected-value estimates should be presented. Hypothetical best-case or worst-case estimates may be presented as alternatives for sensitivity analysis. Where possible, information about the probability distribution of the parameter estimate should be presented.

A common situation that arises in estimating both benefits and costs is that a number of different studies may exist which together provide a range of different estimates for a particular parameter. In general, it is not appropriate to use the midpoint of the range of extreme values provided by the studies. Such a technique ignores the information provided by all studies except those providing the extreme values, which may be the least reliable. The preferred approach to deriving an expected-value estimate of a particular parameter in this situation would be to derive it as a weighted average of the estimates of the individual studies, with the weight of each estimate being based on the reliability (in the best judgment of the agency) of the study that produced it.

Where expected future benefits or costs are uncertain, their value to those who receive them may be different from their value if they were certain. (Often, but not always, a certain future benefit is worth more to people than an uncertain future benefit with the same expected value.) As noted in the previous section, it is incorrect to adjust the discount rate as a device to account for the riskiness of future benefits or costs. Any allowance for risk should be made by adjusting the monetary values (for the year in which they occur) of the uncertain benefits and costs so that they are expressed in terms of their "certainty-equivalents."

For an uncertain benefit in future year X, the certainty-equivalent is the number of certain dollars in year X that the uncertain benefit is worth to its recipient. For example, suppose that a particular regulation reduces the probability of fire in a particular type of facility. As part of a benefit-cost analysis for this regulation, the dollar value of the expected reduction in fire loss would be calculated. The owners of the protected facilities place a higher dollar value on the risk of a fire than the expected dollar value of the loss. This is demonstrated by their willingness-to-pay for fire insurance. Therefore, their relative net cost (the percentage difference between insurance premiums and insurance company claim payments) for fire insurance can be used to increase the expected dollar value of the reduction in fire loss to its certainty-equivalent value.

In the example of the preceding paragraph, the adjustment for risk would involve an increase in the value of the benefit, whereas uncertainty of a benefit is normally thought to reduce its certainty-equivalent value. The reason is that even though this benefit by itself is uncertain, it acts to reduce the overall level of risk that would prevail in the absence of the regulation. This illustrates the important principle that what matters is not the variability or riskiness of a regulation's net benefits by themselves but the regulation's effect on risk and uncertainty overall.

While an adjustment to account for risk may be called for in the fire-risk example given, a similar adjustment for the value of reductions in fatalities and injuries would not be appropriate. Assuming that the values of fatalities and injuries have been derived by the willingness-to-pay methodology recommended in Section B.2 below, they would already represent the certainty-equivalent value of the uncertain risk. This is because the estimated dollar values represent the certain dollar amounts that individuals would sacrifice to reduce these risks.

Probably, in most cases, it will not be advisable to adjust for risk and uncertainty. As a theoretical matter, no adjustment for risk is necessary wherever the net benefits are widely dispersed among many individuals and are not correlated with disposable income. And in cases where this does not apply, risk may be relatively unimportant or may already be taken into account by use of the willingness-to-pay methodology. In other cases, there may be no practical way to quantify the value of changes in risk.

4. *Assumptions.* Where benefit or cost estimates are heavily dependent on certain assumptions, it is essential to make these assumptions explicit and, where alternative assumptions are plausible, to carry out sensitivity analyses based on plausible alternative assumptions. If the decision criterion proves to be sensitive to alternative plausible assumptions, this may necessitate further research to develop more evidence on which of the alternative assumptions is the most appropriate. Because the adoption of a particular estimation methodology sometimes implies major hidden assumptions, it is important to analyze estimation methodologies carefully to make hidden assumptions explicit.

5. *International Trade Effects.* In calculating the benefits and costs of a proposed regulatory action, generally no explicit distinction needs to be made between domestic and foreign resources. If, for example, compliance with a proposed regulation requires the purchase of specific equipment, the opportunity cost of that equipment is ordinarily best represented by its domestic cost in dollars, regardless of whether the equipment is produced domestically or imported. The relative value of domestic and foreign resources is correctly represented by their respective dollar values, as long as the foreign exchange value of the dollar is determined by a free exchange market. Nonetheless, an awareness of the role of international trade may be quite useful for assessing the benefits and costs of a proposed regulatory action. For example, the existence of foreign competition usually makes the demand curve facing a domestic industry more elastic than it would be otherwise. Elasticities of demand and supply frequently can significantly affect the magnitude of the benefits or costs of a regulation.

A regulation that discriminates unjustifiably against foreign exporters is a form of economic protectionism. The economic loss to the U.S. due to the fact that protectionism is economically inefficient will be reflected in the net benefit estimate of any properly conducted benefit-cost analysis. However, a benefit-cost analysis will generally not be able to measure the potential U.S. loss from the threat of future retaliation by foreign governments. Therefore, special attention should be given to any possibility that a regulation would unjustifiably discriminate between domestic and foreign producers and consumers—both discrimination against foreigners and discrimination in favor of foreigners.

The fact that a regulation has a differential effect on foreigners as compared to Americans does not necessarily constitute discrimination. If, for example, an automobile safety standard could be complied with less expensively by large cars than by small cars, such a standard would be more favorable to American car producers, who produce relatively more large cars compared to the fleet mix of foreign producers. Nonetheless, such a differential effect would not be discriminatory if the difference in compliance cost between large and small cars was necessary to achieve legitimate regulatory objectives in the most efficient way.

If a regulation has an adverse differential effect on foreign producers or consumers relative to domestic producers and consumers that is not necessary to realize regulatory goals efficiently, then a discriminatory effect on foreign trade exists. The RIA should identify any substantial differential effect on international trade and explain why it is necessary to achieve legitimate regulatory goals in the most efficient way. One means for reducing the likelihood of international discrimination would be for a U.S. product standard for an internationally traded good to be based on an international standard, wherever an international standard exists and is compatible with the health, safety, or environmental needs of the U.S. International harmonization can be beneficial for regulations directly setting standards for internationally traded goods or services. For example, it would be appropriate to consider international harmonization in setting safety standards for automobiles. There is no similar advantage to international harmonization where a regulation does not directly affect the quality of an internationally traded good or service, even if it indirectly affects its costs (e.g., environmental controls for automobile plants).

6. *Distributional Effects.* Those who bear the costs of a regulation and those who enjoy its benefits often are not the same persons. Benefits and costs of regulation may also be distributed unevenly over time, perhaps spanning several generations. There is no generally accepted way to monetize potential

distributional effects. Attempts to incorporate distributional concerns in benefit-cost analysis require the establishment of unequal weights for different groups in society. Because positive economics treats equally the willingness-to-pay of all individuals, any alternative weighting would undermine the objective character of the analysis. Policymakers may wish, however, to take account of the distributional effects of various regulatory alternatives. Therefore, where there are potentially important differences between those who stand to gain and those who stand to lose under alternative regulatory options, the RIA should identify these groups and indicate the nature of the differential effects. The RIA should also present information on the streams of benefits and costs over time as well as present value estimates, particularly where intergenerational effects are concerned.

B. Benefit Estimates

The RIA should state the beneficial effects of the proposed regulatory change and its principal alternatives. In each case, there should be an explanation of the mechanism by which the proposed action is expected to yield the anticipated benefits. An attempt should be made to quantify all potential real incremental benefits to society in monetary terms to the maximum extent possible. A schedule of monetized benefits should be included that would show the type of benefit and when it would accrue; the numbers in this table should be expressed in constant, undiscounted dollars. Any expected incremental benefits that cannot be monetized should be explained.

The RIA should identify and explain in detail the data or studies on which benefit estimates are based. Where benefit estimates are derived from a statistical study, the RIA must provide sufficient information so that an independent observer can determine the representativeness of the sample, whether it was extrapolated from properly in developing aggregate estimates, and whether the results are statistically significant.

For regulations addressing health and safety risks, the calculation of potential benefits should derive from the agency's estimate of the mean expected value of the reduction in risk attributable to the standard. Estimates of the prevailing level of risk and of the reduction in risk to be anticipated from a proposed standard should be unbiased expected-value estimates rather than hypothetical worst-case estimates. Extreme safety or health results should be weighted (along with intermediate results) by the probability of their occurrence to estimate the expected result implied by the available evidence. In addition, to the extent possible, the distribution of probabilities for various possible results should be presented separately, so as to allow for an explicit margin of safety, where required, in final decisions. If a margin of safety is to be provided, the proper place for it is the final stage of the decision-making process, not by adjusting the risk or benefit estimates in a conservative direction at the information-gathering or analytical stages of the process. Conservative estimates should be presented as alternatives to best estimates for sensitivity analysis but should not substitute for them.

It is important to guard against double-counting of benefits. For example, if a regulation improved the quality of the environment in a community, the value of real estate in the community might rise, reflecting the greater attractiveness of living in the improved environment. It would ordinarily be incorrect to include the rise in property values among the benefits of the regulation. Ordinarily, the value of environmental benefits (e.g., reduced health risks, scenic improvements) will already be included among the benefits. The rise in property values reflects the capitalized value of these improvements. Therefore, to count as benefits both the value of the environmental improvements and the corresponding increase in property values is to count the same benefits twice. Only where a direct estimate of the benefits has not been included would it be appropriate to include the increase in property values among the benefits.

1. *General Considerations*. The concept of "opportunity cost" is the appropriate construct for valuing both benfits and costs. The principle of "willingness-to-pay" captures the notion of opportunity cost by providing an aggregate measure of what individuals are willing to forgo so as to enjoy a particular benefit. Market transactions provide the richest database for estimating benefits based on willingness-to-pay, so long as the goods and services affected by a potential regulation are traded in markets. Estimation problems arise in a variety of instances, of course, where prices or market transactions are difficult to monitor. Markets may not even exist in some instances, forcing regulatory analysts to develop appropriate proxies that simulate market exchange. Indeed, the analytical process of deriving benefit estimates by simulating markets may suggest alternative regulatory strategies that create such markets.

Willingness to pay always provides the preferred measure of benefits. Estimates of willingness-to-pay based on observable and replicable behavior deserve the greatest level of confidence. Considerably less confidence should be conferred on benefit estimates that are neither derived from market transactions nor based on behavior that is observable or replicable. Of course, innovative benefit estimation methodologies may be necessary in some cases, and should be encouraged. However, reliance upon such methods intensifies the need for quality control to ensure that estimates derived conform as closely as possible to what would be observed if markets existed.

2. *Principles for Valuing Directly Observable Benefits*. Ordinarily, goods and services are to be valued at their market prices. However, in some instances, the market value of a good or service may not reflect its true value to society. If a regulatory alternative involves changes in such a good or service, its monetary value for purposes of benefit-cost analysis should be derived using an estimate of its true value to society (often called its "shadow price"). For example, suppose a particular air pollutant damages crops. One of the benefits of controlling that pollutant will be the value of the crop saved as a result of the controls. If the price of that crop is held above the free-market equilibrium price by a government price-support program it will overstate the value of the benefit of controlling the pollutant if the crop saved were valued at the market price established by the support program. The social value of the benefit should be calculated using a shadow price for crops subject to price supports. The estimated shadow price should reflect the value to society of marginal uses of the crop (e.g., the world price if the marginal use is for exports). If the marginal use is to add to very large surplus stockpiles, the shadow price would be the value of the last units released from storage minus storage cost. Therefore, where stockpiles are large and growing, the shadow price is likely to be low and could well be negative.

3. *Principles for Valuing Benefits that are Indirectly Traded in Markets*. In some important instances, a benefit corresponds to a good or service that is indirectly traded in the marketplace. Important examples include reductions in the health-and-safety risks, the use-value of environmental amenities and scenic vistas, and savings in time. To estimate the monetary value of such an indirectly traded good, the willingness-to-pay valuation methodology is still conceptually superior, because the amount that people are willing to pay for a good or service is the best measure of its value to them. As noted in Sections 4 and 5 immediately following, alternative methods may be used where there are practical obstacles to the accurate application of direct willingness-to-pay methodologies.

A variety of methods have been developed for estimating indirect benefits. Generally, these methods apply statistical techniques to distill from observable market transactions the portion of willingness-to-pay that can be attributed to the benefit in question. Examples include estimates of the value of environmental amenities derived from travel-cost studies, hedonic price models that measure differences or changes in the value of land, and statistical studies of occupational-risk premiums in wage rates.

Contingent-valuation methods have become increasingly popular for estimating indirect benefits, but they suffer from the fact that survey instruments have a limited capacity to simulate real-world market behavior. Benefit estimates derived from contingent-valuation studies thus have a greater burden of analytical care to ensure that they represent in an unbiased manner what actually occurs in the marketplace.

4. *Principles and Methods for Valuing Benefits that are Not Traded Directly or Indirectly in Markets*. Some types of goods, such as the social benefit of preserving environmental amenities apart from their use and direct enjoyment by people, are not traded directly or indirectly in markets. The practical obstacles to accurate measurement are similar to (but generally more severe than) those arising with respect to indirect benefits, principally because there are not market transactions to provide data for willingness-to-pay estimates.

Contingent-valuation methods provide the only analytical approaches currently available for estimating the benefits of such untraded goods. The absence of observable and replicable behavior with respect to the benefit in question, combined with the difficulties of avoiding bias in contingent-valuation studies, argues for great care and circumspection in the use of such methods. This means, for example, that estimates of willingness-to-pay must incorporate the variety of alternative means individuals have of expressing value for untraded goods. Moreover, analyses must faithfully capture individuals' budget constraints, which restrict their willingness-to-pay for untraded as well as traded goods and services. Benefit analyses derived from contingent valuation and similar methods thus require considerable analytic rigor in design and careful execution. Absent such efforts, analyses based heavily on the benefits of untraded goods and services ordinarily would fail the test of a satisfactory RIA.

5. *Methods for Valuing Health and Safety Benefits*. For health and safety benefits, a distinction should be made between risks of nonfatal illness or injury and fatality risks.

(a) *Nonfatal illness and injury*. Although the willingness-to-pay approach is conceptually superior, the current state of empirical research in the area is not sufficiently advanced to assure that estimates derived by this method are necessarily superior to direct-cost valuations of reductions in risks of nonfatal illness or injury. Any injury-value estimate from a willingness-to-pay study is necessarily an average over a specific combination of injuries of varying severity. If the average injury severity in such a study is greatly different from that for the regulatory action under study, then the study's estimated injury value may not be appropriate for evaluating that action. Accordingly, the agency should use whichever approach it considers most appropriate for the decision at hand. The primary components of the direct-cost approach are medical costs and the value of lost production. Possibly important costs that may be omitted by the use of the direct-cost approach are the value of pain and suffering and the value of time lost from leisure and other activities that are not economically directly productive.

(b) *Fatality.* Reductions in fatality risks are best monetized according to the willingness-to-pay approach. The value of changes in fatality risk is sometimes expressed in terms of the "value of life." This is something of a misnomer since the value of a life really refers to the sum of many small reductions in fatality risk. For example, if the annual risk of death is reduced by one in a million for each of two million people, that represents two "statistical lives" saved per year (two million × one millionth = two). If the annual risk of death is reduced by one in 10 million for each of 20 million people, that also represents two statistical lives saved. The conclusion that the fatality risk reductions in these two cases are equivalent implies an assumption. The implicit assumption—that equal increments in risk are valued equally—allows different risk increments to be added together and compared directly. As a different example, suppose there are two alternative reductions in the annual risk faced by an individual:

A: from $.10 \times 10^{-4}$ to $.09 \times 10^{-4} = .01 \times 10^{-4}$
B: from 1.00×10^{-4} to $.99 \times 10^{-4} = .01 \times 10^{-4}$

Since in both cases the reduction in annual risk is the same ($.01 \times 10^{-4}$), the value of A and B should be considered the same.

The assumption that equal increments in fatality risk are of equal value is a legitimate one, so long as the level of fatality risk is below 10^{-4} annually. There is evidence that the willingness-to-pay value for increments in fatality risk does not change significantly over a wide range of risk exposure below 10^{-4} annually.

For levels of annual risk exposure of 10^{-4} and above it cannot be assumed that equal increments of risk are valued equally. At these higher risk levels, it is particularly important to distinguish between situations of voluntary risk assumption and those of involuntary risk. Where the high risk is involuntary, it is appropriate to value reductions in risk from that high level more highly than equal risk reductions at lower risk levels. In general, the greater the risk that an individual bears, the higher will be the value the individual places on marginal changes in risk. On the other hand, where a high risk is chosen voluntarily those assuming the risk tend to be persons who place a relatively low value on averting safety risks. Empirical studies of risk premiums in high-risk occupations suggest that reductions in voluntarily assumed high risks should be valued less than equal risk reductions at ordinary risk levels.

Estimates of the value of fatality risks refer only to changes in an uncertain risk of death. They have no application to the certain prevention of the death of an identifiable individual.

6. *Alternative Methodological Frameworks for Estimating Health and Safety Benefits.* Several alternative ways of incorporating fatality risks into the framework of benefit-cost analysis may be appropriate. These may involve either explicit or implicit valuation of fatality risks.

One acceptable explicit valuation approach would be for the agency to select a single value for reductions in fatality risk at ordinary risk levels (below 10^{-4} annually) and use this value consistently for evaluating all its programs that affect ordinary fatality risks. Another acceptable explicit valuation approach would be to use a range of values for reductions in fatality risk and apply sensitivity analysis as with other parameters that have alternative plausible values. The range of alternative values should be a reasonable one, not one that includes the most extreme upper and lower values of fatality risk reduction that have been estimated. Extreme values are more appropriate for instances of extraordinarily high risks (above 10^{-4} annually), with the extreme low values being appropriate where voluntary assumption of high risk leads to self-selection and the extreme high values being appropriate where the high risk is involuntarily assumed.

Where the analysis uses a range of alternative values for reductions in fatality risk, it may be useful to calculate break-even values, as in other sensitivity analyses. This requires calculating the borderline value of reductions in fatality risk at which the net benefit decision criterion would switch over from favoring one alternative to favoring another (i.e., the value of fatality risk at which the net benefits of the two alternatives are equal). This method will frequently be infeasible because of its computational demands or because alternatives are continuous rather than discrete (e.g., alternative stringencies for exposure levels), but where appropriate, it is a useful supplement to the sensitivity analysis.

SL 064766

An implicit valuation approach could entail calculations of the cost per unit of reduction in fatality risk (cost per "statistical life saved"), with costs defined as costs minus monetized benefits. This must be used with care since there is a serious potential pitfall: It is *not* correct to choose between two mutually exclusive alternatives by selecting the alternative with lowest cost per statistical life saved. The alternative with higher cost per life saved may nonetheless be the alternative with the higher net benefit to society.

The way to avoid this pitfall while retaining the implicit valuation approach is to make all calculations of cost per life saved in terms of increments between alternatives. Alternatives should be arrayed in order of their total reduction in expected fatalities and the incremental cost per life saved calculated between each adjacent pair of alternatives. In contrast to explicit valuation approaches, this avoids the necessity of specifying in advance a value for reductions in fatality risks. However, a range of values will be implied by the final selection of an alternative. This range should be consistent with estimated values of reductions in fatality risks calculated according to the willingness-to-pay methodology.

Another way of expressing reductions in fatality risks is in terms of life-years saved. For example, if a regulation protected individuals whose average remaining life expectancy was 40 years, then a risk reduction of one fatality would be expressed as 40 life-years saved. Such a refinement may be desirable for regulations that disproportionately protect young people (e.g., motor vehicle safety regulations) or elderly people (e.g., regulations controlling carcinogens). To derive the value of a life-year saved from an estimate of the value of life, first determine the average remaining life expectancy of the sample population in the study from which the estimate was drawn. Assuming that the average age of the sample population is known, the average remaining life expectancy may be derived from actuarial tables giving life expectancy in relation to age. Using standard compound interest tables, the value of a life-year saved can then be determined as the estimated value of life annualized over a period equal to the number of years of remaining average life expectancy.

C. Cost Estimates

1. *General Considerations*. The opportunity cost of an alternative is the value of the benefits foregone as a consequence of that alternative. For example, the opportunity cost of banning a product (e.g., a drug, food additive, or hazardous chemical) is the foregone net benefit of that product. It is measured by changes in producers' and consumers' surpluses. (Producers' surplus is the difference between the amount a producer is paid for a unit of a good and the minimum amount the producer would accept to supply that unit. It is measured by the distance between the price and the supply curve for that unit. Consumers' surplus is the difference between what a consumer pays for a unit of a good and the maximum amount the consumer would be willing to pay for that unit. It is measured by the distance between the price and the demand curve for that unit.) As another example, even if a resource required by regulation does not have to be paid for because it is already owned by the regulated firm, nonetheless, the use of that resource to meet the regulatory requirement has an opportunity cost equal to the net benefit it would have provided in the absence of the requirement. Any such foregone benefits for an alternative should be monetized wherever possible and either added to the costs or subtracted from the benefits of that alternative. Any costs that are averted as a result of an alternative should be monetized wherever possible and either added to the benefits or subtracted from the costs of that alternative.

All costs calculated should be incremental, that is, they should represent changes in costs that would occur if the regulatory alternative is chosen compared to costs in the base case (ordinarily no regulation or the existing regulation). Future costs that would be incurred even if the regulation is not promulgated, as well as costs that have already been incurred (sunk costs), are not part of incremental costs. If marginal cost is not constant for any component of costs, incremental costs should be calculated as the area under the marginal cost curve over the relevant range.

Costs include private-sector compliance costs, government administrative costs, and costs of reallocating workers displaced as a result of the regulation. Costs that are not monetary outlays must be included and should be attributed a monetary value wherever possible. Such costs may include the value (opportunity cost) of benefits foregone, losses in consumers' or producers' surpluses, discomfort or inconvenience, and loss of time. A schedule of monetized costs should be included that would show the type of cost and when it would occur; the numbers in this table should be expressed in constant, undiscounted dollars. Any expected incremental costs that cannot be monetized should be explained. An important type of cost that often cannot be quantified is a slowing in the rate of innovation or of adoption of new technology. For example, regulations requiring a costly and time-consuming approval process for new products or new facilities may have such costs, as may regulations setting much more stringent standards for new facilities than existing ones.

Two accounting cost concepts that should not be counted as costs in benefit-cost analysis are interest and depreciation..The time value of money is already accounted for by the discounting of benefits and costs. Depreciation is already taken into account by the time distribution of benefits and costs; the only legitimate use for depreciation calculations in benefit-cost analysis is to estimate the salvage value of a capital investment.

2. *Real Costs versus Transfer Payments.* An important, but sometimes difficult, problem in cost estimation is to distinguish between real costs and transfer payments. Transfer payments are not genuine costs but payments for which no real good or service is received in return. Several examples of problems that may arise from the confusion between transfer payments and real costs (or benefits) may help to identify situations in which further analysis of the problem may be warranted. Monopoly profits, insurance payments, government subsidies and taxes, and distribution expenses are four potential problem areas.

(a) *Monopoly profits.* If, for example, sales of a competitively produced product were restricted by a government regulation so as to raise prices to consumers, the resulting monopoly profits are not a benefit of the rule, nor is their payment by consumers a cost. The real benefit-cost effects of the regulation would be represented by changes in producers' and consumers' surpluses.

(b) *Insurance payments.* Potential pitfalls in benefit-cost analysis may also arise in the case of insurance payments, which are transfers. Suppose, for example, a worker safety regulation, by decreasing employee injuries, led to reductions in firms' insurance premium payments. It would be incorrect to count the amount of the reduction in insurance premiums as a benefit of the rule. The proper measure of benefits is the value of the reduction in worker injuries, monetized as described previously, plus any reduction in real costs of administering insurance (such as the time of insurance company employees needed to process claims) due to the reduction in worker insurance claims. Reductions in insurance premiums that are matched by reductions in insurance claim payments are changes in transfer payments, not benefits.

(c) *Indirect taxes and subsidies.* A third instance where special treatment may be needed to deal with transfer payments is the case of indirect taxes (tariffs or excise taxes) or subsidies on specific goods or services. Suppose a regulation requires firms to purchase a $10,000 piece of imported equipment, on which there is a $1,000 customs duty. For purposes of benefit-cost analysis the cost of the regulation for each firm ordinarily would be $10,000, not $11,000, since the $1,000 customs duty is a transfer payment from the firm to the Treasury, not a real resource cost. This approach, which implicitly assumes that the equipment is supplied at constant costs, should be used except in special circumstances. Where the taxed equipment is not supplied at constant cost, the technically correct treatment is to calculate how many of the units purchased as a result of the regulation are supplied from increased production and how many from decreased purchases by other buyers. The former units would be valued at the price without the tax and the latter units would be valued at the price including tax. This calculation is usually difficult and imprecise because it requires estimates of supply and demand elasticities, which are often difficult to obtain and inexact. Therefore, this treatment should only be used where the benefit-cost conclusions are likely to be sensitive to the treatment of the indirect tax. While costs ordinarily should be adjusted to remove indirect taxes on specific goods or services as described here, similar treatment is not warranted for other taxes, such as general sales taxes applying equally to most goods and services or income taxes.

(d) *Distribution expenses.* The treatment of distribution expenses is also a source of potential error. For example, suppose a particular regulation raises the cost of a product by $100 and that wholesale and retail distribution expenses are on average 50 percent of the factory-level cost. It would ordinarily be incorrect to add a $50 distribution markup to the $100 cost increase to derive a $150 incremental cost per product for benefit-cost analysis. Most real resource costs of distribution do not increase with the price of the product being distributed. In that case, either distribution expenses would be unchanged or, if they increased, the increase would represent distributor monopoly profits. Since the latter are transfer payments, not real resource costs, in neither case should additional distribution expenses be included in the benefit-cost analysis. However, increased distribution expenses should be counted as costs to the extent that they correspond to increased real resource costs of the distribution sector as a result of the change in the price or characteristics of the product.

D. Expenditure Rules

Regulations establishing terms or conditions of Federal grants, contracts, or financial assistance call for a different form of regulatory analysis than do other types of regulation. In some instances, a full-blown benefit-cost analysis may be appropriate to inform Congress and the President more fully about the desirability of the program, but this would not ordinarily be required in a Regulatory Impact Analysis. The primary function of the RIA for this type of regulation should be to verify that the terms or conditions are the minimum necessary to achieve the purposes for which the funds were appropriated. They should not contain conditions in pursuit of goals that are not germane to the purpose for which the funds were authorized and appropriated. Beyond controls to prevent abuse and to ensure that funds appropriated to achieve a specific purpose are channeled efficiently toward that end, maximum discretion should be allowed in the use of Federal funds, particularly when the recipient is a State or local government.

IV. RATIONALE FOR CHOOSING THE PROPOSED REGULATORY ACTION

The RIA should include an explanation of the reasons for choosing the selected regulation. Ordinarily, the regulatory alternative selected should be the one that achieves the greatest net benefits. If legal constraints prevent this choice, they should be identified and explained, and their net cost should be estimated.

Where uncertainties are substantial or a large proportion of benefits cannot be monetized, other methods of summarizing the benefit-cost analysis may sometimes be appropriate. When alternative forms of presentation are used, the objective must continue to be the maximization of net benefits (except where prohibited by law). Alternative criteria must be used with care because of the potential for errors or misinterpretation.

Agencies need not calculate the internal rate of return for a regulation. The internal rate of return is often difficult to compute and is problematical when multiple rates exist. It must not be used as a criterion for choosing between mutually exclusive alternatives. As a criterion for choosing between alternatives that are not mutually exclusive, it has no advantages over the criterion of maximizing the present value of net benefits.

Benefit-cost ratios, if used at all, must be used with care to avoid a common pitfall. It is a mistake to choose among mutually exclusive alternatives by selecting the alternative with the highest ratio of benefits to costs. An alternative with a lower benefit-cost ratio than another may have the higher net benefits. Whether a regulation's benefits are greater (or less) than its costs can be determined by whether its benefit-cost ratio is greater (or less) than one. The benefit-cost ratio may be used as a very simplified indicator of the likely sensitivity of the result: If the benefit-cost ratio is much greater than one, the conclusion that the regulation's benefits exceed its costs probably is not sensitive to likely alternative parameter values. If the ratio is only slightly greater than one, the conclusion probably is sensitive. The benefit-cost ratio may sometimes be acceptable as a rough substitute for genuine sensitivity analysis where it is not feasible to carry out a full sensitivity analysis (e.g., if the number of regulatory parameters to be tested by sensitivity analysis is large). When so used, the benefit-cost ratio should be recognized as only a crude approximation to a genuine sensitivity analysis and the analyst should be aware of its limitations (e.g., the benefit-cost ratio is sensitive to the arbitrary classification of an item as a benefit or an averted cost).

Where the benefits of proposed regulatory alternatives include reductions in fatality risks, an acceptable alternative to direct calculation of net benefits is the indirect approach of calculating incremental costs per life saved between adjacent alternatives. This is done by ranking all the alternatives according to the number of lives they save and then calculating the change in costs and the change in lives saved between each alternative and the one with the next highest number of lives saved. If the alternative selected is the one whose incremental cost per life saved is closest to the willingness-to-pay value of life, this decision criterion is analytically equivalent to that of maximizing net benefit.

In cases where important benefits cannot be assigned monetary values, cost-effectiveness analysis should be used where possible to evaluate alternatives that generate equivalent nonmonetizable benefits. Costs should be calculated net of monetized benefits. Between two alternatives with equivalent nonmonetizable benefits, the alternative with the lower net costs should be selected. Cost-effectiveness analysis should also be used to compare regulatory alternatives in cases where the level of benefits is specified by statute.

V. STATUTORY AUTHORITY

The RIA should include a statement of determination and explanation that the proposed regulatory action is within the agency's statutory authority.

Further Reading

Edith Stokey and Richard Zeckhauser, *A Primer for Policy Analysis*. Chapters 9 and 10 provide a good introduction to basic concepts.

E. J. Mishan, *Economics for Social Decisions: Elements of Cost-Benefit Analysis*. Assumes some knowledge of economics. Chapters 5–8 should be helpful on the important subjects of producers' and consumers' surpluses (not discussed extensively in this guidance document).

W. Kip Viscusi, *Risk By Choice*. Chapter 6 is a good starting point for the topic of valuing health and safety benefits. Other more technical sources are 'ven in the bibliography.

Robert Cameron Mitchell and Richard C. Carson, *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Provides a valuable discussion on the potential pitfalls associated with the use of contingent-valuation methods.

V. Kerry Smith, Ed., *Advances in Applied Microeconomics: Risk, Uncertainty, and the Valuation of Benefits and Costs*.

Judith D. Bentkover, Vincent T. Covello, and Jeryl Mumpower, Eds., *Benefits Assessment: The State of the Art*.