EPA/SAB/81/002

Review of "A General Method for Assessing Health Risks Associated With Primary National Ambient Air Quality Standards" (Draft dated April 1981)

A Report of the Subcommittee on Health Risk Assessment

September 1981

Science Advisory Board
U.S. Environmental Protection Agency
Washington, D.C. 20460

EPA NOTICE

Table of Contents

Background

This is the third report of the Subcommittee on Health Risk Assessment of the Science Advisory Board critiquing approaches to health risk assessment for alternative national ambient air quality standards developed under the auspices of EPA's Office of Air Quality Planning and Standards (OAQPS).

In 1979, the Subcommittee reviewed a methodology for assessing health risks associated with alternative ambient air quality standards for ozone. The Subcommittee stated its comments, conclusions, and recommendations in a report dated September 1979. */

Among other comments and suggestions, the Subcommittee

- found that the methodology developed by OAQPS was not ready for application in the process of establishing national ambient air quality standards; and

- recommended that OAQPS should formulate a plan outlining how the Agency will (a) develop the proposed methodology, including standards and protocols for application, (b) consider alternative approaches, and (c) select and establish the credibility of the best methodology.

Consistent with the Subcommittee's recommendations, the Office of Air Quality Planning and Standards developed a "Risk Assessment Program Plan" and contracted with six research teams to develop integrated conceptual risk assessment procedures. This effort resulted in six documents which the Subcommittee reviewed at a public meeting in September 1980. Also at that meeting, the Subcommittee heard update reports on the F/B Risk Assessment Method (named after its authors, Thomas B. Feagans and William F. Biller) and on contemplated next steps in the OAQPS "Risk Assessment Program Plan." The Subcommittee's comments, findings, and conclusions are contained in a report dated December 1980.**/

*/ Review of "A Method of Assessing the Health Risks Associated With Alternative Air Quality Standards for Ozone" (Draft dated July 1978), A Report of the Subcommittee on Health Risk Assessment, EPA/SAB/79/001, September 1979. The method is referred to as the F/B Risk Assessment Method or the F/B Method after its authors, Thomas B. Feagans, OAQPS Analyst, and Dr. William F. Biller, Consultant.

**/ Approaches to Health Risk Assessment for Alternative National Ambient Air Quality Standards, A Report of the Subcommittee on Health Risk Assessment, EPA/SAB/80/003, December 1980.

Briefly, the Subcommittee concluded that

- the program has produced two and possibly three promising approaches to health risk assessment for use in developing alternative national ambient air quality standards; and

- information presented on the F/B Risk Assessment Method did not reflect significant improvements since the Subcommittee first reviewed it and discussed its deficiencies in public session in April 1979.

Following the Subcommittee's recommendations, OAQPS selected two approaches as alternatives for further development: the approach presented by M.W. Merkhofer and the one by Robert L. Winkler and Rakesh K. Sarin. As regards the F/B Method, OAQPS advised the Subcommittee that the method will be presented in a comprehensive report and will be subject to extensive in-house and external reviews including another review by this Subcommittee. OAQPS then planned to proceed as follows. If the reviews indicate that further development of the F/B Method is warranted, OAQPS will hold it in abeyance until alternative methodologies can be brought into a parallel state of development. If the reviews indicate that further OAQPS support of the F/B approach is not warranted, all work on that method will be terminated except for the generally-applicable exposure analysis module. Resources currently designated for further development of the F/B Method would be applied to the alternative approach(es). */

A comprehensive report describing the F/B approach was issued as a draft and became available for Subcommittee review in mid-April. **/

---

*/ Letter from Joseph Padgett, Director, Strategies and Air Standards Division, OAQPS, to Dr. Anita S. Curran, Subcommittee Chairman, dated December 17, 1980.

**/ "A General Method for Assessing Health Risks Associated with Primary National Ambient Air Quality Standards" (Draft dated April 1981) by Thomas B. Feagans and William F. Biller

OAQPS also arranged for six experts to provide in-depth technical reviews of the F/B report which became available shortly before or at the meeting. */

*/ The six experts selected by OAQPS to provide technical reviews were:

Dr. David E. Bell, Harvard University Business School

Dr. C.S. Burton, Systems Applications, San Rafael, California

Dr. Richard C. Jeffrey, Department of Linguistics and Philosophy, Massachusetts Institute of Technology; after June 15, 1981: Department of Philosophy, Princeton University

Dr. Isaac Levi, Department of Philosophy, Columbia University

Dr. John H. Seinfeld, Department of Chemical Engineering, California Institute of Technology

Dr. Andrew J. Van Horn, Energy and Environmental Analysis Division, Teknetron Research, Orinda, California

Comments

We again wish to commend and compliment the Office of Air Quality Planning and Standards (OAQPS) on their efforts in developing health risk assessment and evaluation methodologies and recommend that they continue these efforts.

The contracts commissioned over this past year and a half have produced valuable research on the application of decision analysis principles and techniques in the development of health risk assessment and evaluation methodologies to aid in setting ambient air quality standards. But we can see little or no impact from this research or from other sources in remedying the deficiencies this Subcommittee has perceived with the F/B approach. It remains unnecessarily complex, continues to be poorly communicated despite repeated attempts at clarification, contains concepts that have not been fully worked out, and in some areas contains unclear mathematics. (See Appendix B, Dr. Watson's critique of Appendix G of the F/B Report.)

Our perception is that EPA's goal is to have an implementable methodology within the next several years for assessing health risks to aid the Administrator in setting ambient air quality standards. While the F/B approach may have commendable aspects as a research effort, it is not, in its present form, an implementable tool for public policy decision making, and, we believe, will not be accepted by a large part of the scientific community or the public. As we have repeatedly suggested, Feagans and Biller should publish their work in peer-reviewed journals, and the professional community should judge the merits of their viewpoint. It is our belief that even if the F/B viewpoint were to become widely accepted, this would not occur in the foreseeable future.

The scientific data and reasoning should be emphasized, not the mathematical complexities of processing probabilities. There is a lack of emphasis in the F/B approach on what constitutes an adverse health effect, what constitutes the most sensitive group, and which health effects should be most important in establishing ambient standards. The discussion of responses of concern, sensitive populations, seriousness of effect, uncertainty about causality (pages 7-2 to 7-4 of the F/B Report) is unsatisfactory and superficial. Agreement on basic definitions and measurement criteria is essential if comparisons between experts' probabilities are to have any meaning.

Another area of difficulty is that of working through the mathematics of a complex model using upper and lower probabilities. In this situation a probability density function does not exist, and integration over the range of an uncertain quantity becomes much more difficult. In complex parts of their analysis, such as the dose-response discussion in Appendix E, Feagans and Biller go back to using single-valued probabilities and density functions. Feagans and Biller state (p. E-2), "With care, it is relatively straightforward to recast the results and the supporting treatment in the form of upper and lower probabilities." We are not convinced that the extensive development of pages E-2 to E-29 and Appendix F can be carried through with upper and lower probabilities without encountering some intractable or severely restrictive mathematical difficulties. Only the most diligent and determined probability expert could understand the mathematical development of Appendices E and F, given that it is carried out using upper and lower probabilities.

A third area of difficulty was pointed out by Professor Levi. It is not clear how the upper and lower probabilities are to be used to arrive at decisions. The risk ribbon approach advocated by Feagans and Biller is ad hoc and unsatisfactory. Professor Levi notes that approaches to decision theory using upper and lower probabilities have been developed by himself and others and published in the literature, but these approaches have not been incorporated into the F/B approach.

The Subcommittee suggests that all material relating to upper and lower probability, the theoretical foundations of probability, and risk ribbons should be considered basic research and separated from the approach. For an implementable tool, we suggest the analysis should be based on standard decision analysis or Bayesian methods, using single-valued probability distributions. Each such distribution would represent the judgment of one expert or a consensus among a group of experts for each uncertain input to the risk analysis, such as the percentage of a well-defined sensitive group suffering a health effect from a given pollutant concentration. These distributions would be combined with meteorology and human activity models to produce probability distributions on threshold exceedances and the number of adverse health effects for alternative standards. Where experts disagree on a probability distribution that is used as input, EPA should carry out a sensitivity analysis to determine the impact of the disagreement. Where the disagreement is significant to the decision on choice of standard, each of the disagreeing experts should be asked to communicate the rationale for his or her probability assessments in such a form that this rationale can be reviewed by scientific peers and presented to EPA's senior management.

Any analysis which is to be successfully used and defended by EPA in the standard-setting process must be as simple as is consistent with a satisfactory solution. Its developers and users must be able to describe their work to semi-technical and non-technical people in clear, unambiguous, conventional English. Over the past two years the F/B approach appears to have become more ponderous, complex and inflexible. Moreover, the doctrinaire presentation of the approach may greatly intensify the antagonism of those scientists who are generally skeptical of quantitative risk assessment methods.

The major departure in the F/B approach from widely accepted decision analysis practice is the use of upper and lower probabilities instead of a single probability number for an event and a "sharp" or single-valued function for the probability distribution over the range of an uncertain quantity. The concept of upper and lower probabilities is an interesting one that has received some attention in the literature. However, Feagans and Biller have not worked out a complete approach based on this concept, and the use of upper and lower probabilities greatly complicates an already complex analysis. The basis advanced for using upper and lower probabilities is that the authors believe that the concept is not only theoretically sound, but indeed preferable to the standard "behaviorist" approach based on the work of Frank Ramsey, Bruno di Finetti, and L. J. Savage. However, the latter approach has been widely taught and practiced over the past twenty years, and few members of the relatively large decision analyis/Bayesian statistics community have as yet been persuaded of the merits of the forty-year old B.O. Koopman viewpoint on probability that has been embraced by Feagans and Biller.

The upper and lower probability approach involves a number of difficulties. One is that of probability encoding. Feagans and Biller cite the techniques used by Stanford Research Institute (SRI), but neither SRI's work nor the extensive literature on which it is based employes upper and lower probabilities. The work on encoding health experts that Tom Wallsten presented at the May 13, 1981 meeting of the Subcommittee was interesting but unique; there is essentially no existing literature or reservoir of practical experience in probability encoding using the upper and lower probability approach. The professional community can consider the implications of Dr. Wallsten's work in conjunction with the already existing literature if and when Dr. Wallsten's work is published.

Rereading the comments of the six reviewers, as regards the merits of the F/B approach, it seems to us that they were mainly addressing the question: Is the F/B approach sufficiently interesting research that it deserves continued support? This, in our opinion, is very different from the question of whether the F/B approach could or should be implemented for an ambient air quality standard in the next several years. We read most of the reviews to state or imply that the approach and the communication of it had significant problems yet to be resolved. We find Bell's comments on efficiency and the caveat in Burton's review particularly important in relation to the feasibility of successfully implementing the F/B approach.

Bell states:

> "It is well known that there are decreasing returns to scale with the complexity of a model to the point where you can end up worse off than no model at all. I think the models here are overly complex for the current state of applied art. I agree with the approach but I believe it is a little too much all at once. I would be happier seeing more modest goals at this point. If the report is only intended to be a look at the future or as a research document as opposed to a draft of an EPA manual then I'm content. I don't believe it's realistic to expect a methodology such as this to be performed with much creedence [sic] given to it in the next 10 years." (p.3)

Burton adds:

> "The concepts and methodology embodied in the approach appear sound and merit (immediate) further development towards the goal of incorporating the techniques into the standard setting process. However, while I support completely the approach, it is apparent that neither the style nor the tone of the report will facilitate acceptance of the approach by the technical, regulating, or regulated community. To take, as the authors do, such an unyielding position regarding the approach of eliciting probability estimates, and to use language that to many, if not most, will be perceived as offensive, appears to be both unnecessary and counterproductive. This statement should not be construed to mean that the authors should not be bold and clear in stating their views; it (the statement) is meant to suggest, however, that there

exist less polarizing means of communicating their views. The authors should be satisfied that both their peers and time will be the final judge. In short why make the introduction and use of this new methodology into the decision-making process more difficult than necessary." (p.1)

The comments of the reviewers reflect in general many of the sentiments expressed by this Subcommittee 2½ years ago in that they endorse the use of judgmental probabilities as a useful additional tool in the standard-setting process, while raising a variety of specific concerns with the methodology. Though the overall comments appear to be favorable, it must be recognized that this is a first view of the F/B approach by these reviewers. We have had the advantage of watching the development of the process over time, and our opinion is that there has been a lack of significant progress toward developing an implementable tool. We see no reason to believe this situation will change.

Feagans and Biller deserve to be commended for their pioneering efforts, and we would expect other analysts to learn from their work. But, as we stated in our September 1979 report, the goal of the methodology should be effective communication between the experts and the decision maker. We believe this goal can be accomplished with a minimally complex methodology employing judgmental probabilities to describe the occurrence of adverse health effects in sensitive groups at low doses.

We urge that EPA not consider the development of these techniques as a competition where a "winning" methodology is selected without provision for flexibility and modification. What has been learned from F/B and the other approaches, and the specifics of the air pollutant on which the approach is to be implemented may require substantial changes within any approach that EPA selects.

Despite our concerns about this specific piece of analysis, we believe that the overall program of health risk assessment in which OAQPS is now engaged is well conceived and making excellent progress. We believe that the OAQPS is the appropriate unit within EPA to pursue this work and we recommend that it be continued.

Advisory Statement */

The Subcommittee again wishes to commend and compliment the Office of Air Quality Planning and Standards (OAQPS) on their efforts in developing health risk assessment and evaluation methodologies and recommends that these efforts be continued. But we see no significant progress in the F/B Report. The comments of the reviewers reflect, in general, many of the sentiments expressed by this Subcommittee 2 1/2 years ago in that they endorse the use of judgmental probabilities as a useful additional tool in the standard-setting process while raising a variety of specific concerns.

We recognize that this is a first view of the F/B approach by these commentators, but we have had the advantage of watching the development of the process over time, and our opinion is that there has been a lack of significant progress toward developing an implementable tool. We see no reason to believe this situation will change.

Despite our concerns about this specific piece of analysis, we believe that the overall program of health risk assessment in which OAQPS is now engaged is well conceived and making excellent progress. We believe that the OAQPS is the appropriate unit within EPA to pursue this work, and we recommend that they continue to develop the alternative approaches now under consideration.

Conclusions

In the view of the Subcommittee, the F/B method is not a promising approach for developing an implementable tool in the foreseeable future.

The Subcommittee recommends that OAQPS continue to develop the two alternative approaches.

*/ This is a slightly edited version of a statement, reflecting a consensus of Subcommittee members, read into the record by the Subcommittee Chairman, Dr. Curran, at the May 13, 1981 meeting of the Subcommittee.

SCIENCE ADVISORY BOARD

SUBCOMMITTEE ON HEALTH RISK ASSESSMENT

Chairman:

Dr. Anita S. Curran, Commissioner, Department of Health, Westchester County, County Office Building 2, 150 Grand Street, White Plains, New York 10601

Members:

Dr. Bernard Altshuler, Director of Biomathematics Laboratory, Institute of Environmental Medicine, New York University Medical Center, 550 First Avenue, New York, New York 10016

Dr. Judy A. Bean, Associate Professor, College of Medicine, Department of Preventive Medicine and Environmental Health, University of Iowa, Iowa City, Iowa 52242

Dr. Granger Morgan, Head, Department of Engineering and Public Policy, Carnegie-Mellon University, 5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213

Dr. D. Warner North, Principal, Decision Focus Inc., 5 Palo Alto Square, Suite 410, Palo Alto, California 94304

Dr. Geoffrey Watson, Professor and Director of Graduate Studies, Department of Statistics, Fine Hall, Princeton University, Princeton, New Jersey 08544

SAB Staff Officer:

Mr. Ernst Linde, Scientist Administrator, Science Advisory Board, A-101M, U.S. Environmental Protection Agency, Washington, D.C. 20460 Phone: (202) 472-9444

Comments by Geoffrey S. Watson on Appendix G, Mathematical Treatment of Uncertainty in Pollutant Concentrations of "A General Method for Assessing Health Risks Associated With Primary National Ambient Air Quality Standards" (Draft dated April 1981)

Appendix G seeks the probability distribution of the n largest in a series of m values. Three years ago I pointed out the need to consider dependent values. The authors now recognize this but make totally incorrect statements. The 19 lines giving the theory for the independence case are marred by five typos, one leading to nonsense. The discussion of this topic is at best opaque.

If the m values ordered from largest to smallest are $x_{(1)}>x_{(2)}>\ldots>x_{(n)}>\ldots>x_{(m)}$, we seek $\mathrm{Prob}(x_{(n)}<C)$. If $G(x) = P(x_i>x)$, for any i, (as in (G-1)), the chance that exactly v of the m values will be greater than C is, from the binomial distribution,

$$\frac{m!}{v!(m-v)!} \quad G(C)^v \ (1-G(C))^{m-v}$$

Now $x_{(n)}<C$ if and only if $v=0,1,\ldots,n-1$.

Hence

$$\mathrm{Prob}(x_{(n)}<C) = \sum_{v=0}^{n-1} \frac{m!}{v!(m-v)!} G(C)^v \ (1-G(C))^{m-v}$$

which is (G-5) corrected.

The particular case of the Weibull distribution (first paragraph on p. G-3) correctly sets

$$G(C) = \exp - (C/\delta)^k \qquad \text{(G-6)}$$

The expected number of exceedances E of a standard C_{std} in n_E is

$$E = n_E\, G(C_{std}) \qquad \text{(i)}$$

To proceed one must know what is meant by the P_C function. This is apparently defined in Section 8.2, but this section is absolutely baffling.

From Section 8.1, it seems clear that if the standard requires E in (i) to be less than unity, then C_{std} could be determined from (i) if δ and k are known.

Returning to Section 8.2 the notation and definition of P_C seems to change from line to line. Hence I cannot say whether (G-7) and the paragraph below it are right or wrong.

The remainder of Appendix G deals with the problems caused by hour-by-hour correlations and non-stationarity. The latter means daily and seasonal variations. For daily maximum the authors refer to a paper by Horowitz and Barakatz (which I have not been able to see) who use the representation

$$\log C(t_i) = U(t_i) + a(t_i) \ ,$$

$$a(t_i) = b_1 a(t_{i-1}) + \ldots + b_q a(t_{i-q}) + a_0(t_i) \ ,$$

where $a_0(t_i)$ are "normal with mean zero" and

$$1 \geqslant b_1 \geqslant \ldots \geqslant b_q \qquad \text{(ii)}$$

I put a finite number of b's because that is presumably intended. I <u>presume</u> that the $a_0(t_i)$ are supposed to be uncorrelated and that the t_i are equally spaced. The condition (ii) no doubt is supposed to be necessary and sufficient to guarantee stationarity It is NEITHER as the trivial example

$$ε(t) = a(t-1) + a(t-2) + a_0(t)$$

should make clear to anyone knowing this theory.

Further, Feagans and Biller assert that Horowitz and Barakatz have derived the distribution of the maximum (presumably over a long period like a year) from their model. This is a major achievement because mathematicians have been trying fruitlessly to solve this problem for at least 30 years (see, e.g., Watson, Ann. Math. Stat. 1954).

Finally no one can guess these results. I have done many simulations and yet am totally unable to guess the results when I vary the parameters.