PROBLEMS IN DETERMINING IF A COMMONLY USED HERBICIDE (2,4,5-T) HAS AN EFFECT ON HUMAN HEALTH

THEODOR D. STERLING
WASHINGTON UNIVERSITY

One has to be reminded occasionally that the roots of statistics lie in problems of inference, especially in the study of efficient and useful experimental designs from which conclusions can be drawn. I start with this reminder because many of the observations and conclusions concerning the effects of pollutants derive from experiments which "happened" more than they were designed and which "presented" their data rather than analyzed them. However, there is yet one other reason for this reminder. Problems of inference raised by studies on pollution seem to show that what is needed is not so much a study of the design of experiments but a study of the strategies of acquiring and analyzing wide ranges of observations and that the end products are not simple inferences about "states of nature" but formulations of "public policies."

Keeping these reminders in mind, we shall next turn to a review of the issues and problems surrounding the question if a commonly used herbicide, 2,4,5-T, has effects that are of concern to the large community of this country.

1. 2,4,5-T is a general pollutant

The widespread (albeit inadvertent) consequences of present practices to control animal and plant pests by use of chemical agents have been recognized only recently as a general pollution problem. In view of the known toxicity of many of the agents, industrial physicians and engineers have been concerned with the manufacture and distribution of these toxic materials and with instituting proper warning procedures so as to avoid what has been commonly called "accidents." It is recognized now that while herbicides and pesticides are designed to affect only a specific target species, their indiscriminate and widespread use creates very general pollution problems. First, all herbicides and pesticides have inadvertent effects on nontarget species. The toxic agent may be extremely widely spread by wind and water to places where it was never intended to show up. For instance, 2,4,5-T applied as a spray can be transported in the atmosphere as a drop of spray, as a gaseous state of 2,4,5-T, or adsorbed on dust or other particulate matters in the air. In this way 2,4,5-T was found adsorbed on dust in a trace of rain in Cincinnati, Ohio, presumably from applications in Texas [32].

Secondly, traces of the chemical may show up in the food and water consumed by human users. Heretofore concern has been mainly for residues that may appear in foods that have been grown with the help of pesticides. In most circumstances this residue can be kept to a very minimum (although this is not always possible). Much more serious, however, is the danger that the pesticide becomes concentrated in the ecological food chain. Man, being on the apex of a pyramid in this food chain, might thus be exposed to large doses of the chemical in a wide variety of foods, especially meats. Thus, herbicides and pesticides used for specific purposes, even in rather isolated rural areas, might find their way in large doses to man wherever he resides. Finally, because the active chemical agents may be also mutagenic and teratogenic, their effect may not be limited to the present population exposed to that toxicity but to future generations through mutations and introduction of malformations. (We should note in passing that until now the mutagenic and teratogenic dangers inherent in some of the commonly seen air and water pollutants have received very little attention.)

Pesticides are of immeasurable benefit. They make it possible for man to inhabit his planet in large numbers by insuring an ample food supply and eliminating or controlling carriers of disease and discomfort. Yet, the dangers created by these same pesticides to human health and survival have to be examined very carefully and weighed against the benefits derived from their use.

2. Patterns of use of 2,4,5-T

The herbicide, 2,4,5-T (2,4,5-trichlorophenoxyacetic acid), is used primarily to control woody plants and a few herbaceous species against which it is more effective than other herbicides. 2,4,5-T is never "pure." It has not been possible so far to manufacture 2,4,5-T without one specific impurity, 2,3,7,8-tetrachlorodibenzo-p-dioxin (commonly referred to as TCDD). While even pure 2,4,5-T has a known (albeit small) toxic and teratogenic effect, the toxicity and teratogenicity of TCDD can be described only as "virulent." While most species tested (with the exception of the dog) can survive a single oral dose of 2,4,5-T in excess of 100 mg/kg, and several can survive daily treatments for a number of days at this high level [12], embryo toxicity appears in litters of females given TCDD in doses as low as .000125 mg/kg per day, and a dose of .008 mg/kg per day caused pallor and debilitation. The LD_{50} for TCDD has been found to be .022 to .045 mg/kg in the rat and .0006 mg/kg in the guinea pig [18]. Manufacturers of 2,4,5-T had paid only limited attention to the amount of impurities present. Both Dow Chemical Company and Hercules Incorporated developed procedures in which the contaminant, dioxin, was held to less than 1 ppm, but other companies manufactured a product that contained as much as 40 ppm of the dioxin, TCDD.

Because of its effectiveness, the chlorophenoxy-herbicides, 2,4-D and 2,4,5-T, have been widely used for over twenty years to control broad leafed weeds. Since a number of studies conducted in the 1950's were interpreted to indicate that the herbicide was not toxic to man at the doses at which man is exposed,

2,4,5-T has been widely used in forestry, agriculture, by the park service, and around the home. One of the major uses of 2,4,5-T has been for rights of way (approximately 50 per cent) and for nonfarm forests (approximately 10 per cent). However, roughly 20 per cent of the produced 2,4,5-T goes into hay, pasture, and rangeland and other farm uses where residues of the herbicide, in one form or the other, may show up in food. The remainder of the 2,4,5-T is used for parks, especially lawn and turf care, gardens, and other purposes where it may come into direct contact with human users in a variety of ways [30]. The pattern of use of the herbicide is especially important for its pollution potential. Its use in grazing and forestry areas may make it subject to entering the food chain and being concentrated in a large number of ways. For instance, it may be concentrated through grazing animals [15] or, by being washed into the lakes and oceans where it is picked up by plankton, it may contaminate fish, crustacea, and mollusks [3]. 2,4,5-T may come into contact with humans through residues on foodstuffs, especially on domestic rice. Finally, 2,4,5-T is a favorite herbicide around the home and garden. It is especially effective in controlling poison ivy. Its unsafe use by a home user may contaminate the family's food in a large number of ways.

One of the big problems, therefore, is whether or not 2,4,5-T can be accumulated in the food chain and stored in tissues. Unfortunately, not much is known about the fate of 2,4,5-T after it is applied, and even less is known about the fate of its active teratogenic and toxic impurity, TCDD. 2,4,5-T is immediately subject to physical and chemical actions that continually reduce the amount remaining at a site of application (by degradation by soil microorganisms, leaching and surface movement in water, volatilization, movement by wind, and photochemical decomposition). Chemically detectable amounts of two pounds per acre have been found in soil after three to seven months after application, but no detectable amount was found in the same soil after one year [1]. The half-life of 2,4,5-T has been estimated as 40 days, at least in forests [22]. Thus, although 2,4,5-T disappears from the soil relatively quickly, it remains long enough to appear as a possible residue in foodstuffs, and certainly its half-life is long enough to enable it to enter the ecological food chain. TCDD is not as easily eliminated from the environment as is 2,4,5-T. Because of its low water solubility (only .2 ppb), TCDD does not move through the soil nor is it easily leached out [30]. Using radiolabeled TCDD, it was found that the radioactive material (probably TCDD) in the soil decreased only 15 to 20 per cent in 160 days, indicating that this compound was very slowly degraded and would persist for more than a year [19]. Thus, there is ample evidence that TCDD remains in the environment for a relatively protracted period.

3. The recent 2,4,5-T controversy

Although a number of studies were done some twenty years ago to evaluate the toxicity of 2,4,5-T, and although the possible teratogenicity of this agent has been under careful investigation during the last two years, there is disagree-

ment on whether or not 2,4,5-T is toxic and teratogenic to humans exposed to low doses. Scientists who are familiar with existing studies are divided in how they interpret the data. For statisticians it is important to notice such divisions, especially if the disagreement is not at what level of a probability (computed or subjective) a hypothesis ought to be accepted or abandoned. What makes the study of the controversy surrounding 2,4,5-T so valuable is that it gives us the opportunity to see the scientific process in operation and, perhaps, enables us to evaluate the usefulness of present methods of designing and analyzing experiments. After all, 2,4,5-T originally was evaluated by a number of conventionally designed toxicity studies, and the recent public controversy surrounding its possible teratogenicity has motivated a number of experiments by both government and industry that were designed specifically to clarify, once and for all, if 2,4,5-T was teratogenic—at least in the experimental animals.

The review of 2,4,5-T came about quite inadvertently through a screening study which the National Cancer Institute contracted with Bionetics Research Laboratories. In this study Bionetics Research Laboratories performed screening studies for carcinogenicity and teratogenicity on a number of pesticides and industrial chemicals. The results released in October, 1969, indicated that 2,4,5-T showed embryo toxicity in two stocks of mice at doses of 113 mg/kg per day when given for several days during organogenesis. As it turned out, the sample of 2,4,5-T used in the Bionetics study was known to have been contaminated with 27 ± 8 ppm of TCDD, and the results were no longer considered valid indication of the teratogenicity of the herbicide. (The experiment has been summarized in a number of forms [8], [23], [31].)

The findings of the Bionetics Laboratories, together with resports by South Vietnamese newspapers of an increased occurrence of birth defects during June and July of 1969 (2,4,5-T had been the major defoliant used in Vietnam), elicited far reaching reactions from governmental agencies and segments of the scientific and concerned lay communities. A number of animal experiments performed early in 1970 confirmed that even the pure samples of 2,4,5-T did, indeed, result in delivery of malformed offspring. In April, 1970, the Secretaries of Agriculture, of Health, Education, and Welfare, and of the Interior jointly announced the suspension of the registration of 2,4,5-T.

Next, the Dow Chemical Company and Hercules Incorporated exercised their right under Section 4.c. of the Federal Insecticide, Fungicide, and Rodenticide Act (*U.S.C.* Vol. 7, p. 135 ff.) to petition for referral of the matter to an advisory committee. Such a committee was then formed from a list supplied by the National Academy of Sciences and met during the early five months of 1971. The committee submitted its report in May, 1971. Although the committee was in substantial agreement on the facts of the case (as far as the evidence was palpable enough to supply the facts), its recommendations on the restoration of registration were divided [6], [14].

At the same time, the United States Military in South Vietnam undertook a survey to evaluate the human evidence of birth defects possibly due to defolia-

tion practices. An Army report was issued in December, 1970 [10]. The Army studies surveyed obstetrical records for the years 1960 to 1969 of 22 provincial, district, and maternity hospitals in 18 cities and other areas in various geographical localities. The findings of the Army studies, in the main, were that no differences in stillbirth rates were observed geographically that could not be attributed to better maternal and neonatal care or to more competent or thorough examination for congenital malformations in the capitol area, and that the rates of stillbirth declined and of congenital malformations remained unchanged during this ten year period despite the heavy spraying between 1966 and 1969. At the same time, the American Association for the Advancement of Science appointed a commission led by Dr. Meselson to investigate the Vietnam charges. This report (HAC) noted that the Army report had been heavily influenced by data from the capitol area which generally experienced little or no exposure to 2,4,5-T. By deducting the capitol area data and considering only data from other parts of the country, the declining trend was reversed, giving considerably increased stillbirth and malformation rates in 1966–69 (heavy spraying period) over 1960–65 (no or light spraying period). Also there was an increase in the incidence of cleft palates and spina bifida (in some regions) that was in need of an explanation. In addition, the HAC report pointed out that in the Tay Ninh Provincial Hospital, which serviced a population that was directly exposed to heavy defoliation or lived near rivers draining areas of defoliation and serving as the source for fish, the stillbirth rate was much higher than in any of the other hospitals surveyed by the Army [21]. (There are two other reported incidents where spraying was thought to be followed by malformed human offspring. One came after a spraying project near Globe, Arizona, and the other from a Swedish defoliation project in Lapland. Claims for both areas were investigated by teams of experts and found to be, most likely, not associated with the spraying incidents.)

4. Sources of uncertainty

It would be a mistake to view the present attitude toward the use of 2,4,5-T as indicating a deep division in the scientific or concerned lay communities. There is general agreement that 2,4,5-T (with or without its impurities) is toxic. The disagreements reflect the burden of responsibilities or of special interests. Governmental agencies tend toward the restriction and more stringent regulation of 2,4,5-T, and agricultural and manufacturing industry and representatives of the forest service tend toward the restoration of the use of 2,4,5-T to its prior status. However, the division does reflect the uncertainty about the dangers to human health. Despite a large number of experiments, some of them directly designed to assess the toxicity and teratogenicity of 2,4,5-T, a substantial uncertainty remains about almost every important question that needs to be answered about the toxic and teratogenic dangers entailed in human exposure.

The problem which causes uncertainty consists of two parts.

(1) The determination of toxicity and teratogenicity of a substance is not a simple matter. Sometimes the experimental solution to questions of fact concerning toxicity cannot be answered by any known experimental procedure, and sometimes they can be answered only partially.

(2) The types and kinds of experimental designs required and the sort of analyses necessary in order to extrapolate from present findings to questions of human health require a high degree of sophistication in experimental design, mathematical analysis (especially in extrapolation techniques), and in information processing. Unfortunately, these skills appear to be very sparsely distributed among toxicological or teratogenic researchers, and proper help from statistics and data processing sources appear to be either unavailable or unused.

As a consequence, a difficult scientific problem is infinitely compounded because proper tools are not brought to bear on its solution. These two facets create and compound the uncertainty about the danger inherent in the use of 2,4,5-T to such an extent that we shall treat them separately and refer to them, respectively, as first and second order uncertainty.

4.1. *First order uncertainty: the experimental model.* All toxicological and teratogenic studies suffer from three major shortcomings.

(1) Animal experiments are, by and large, inadequate;

(2) the dose and response relationship is difficult to determine; and

(3) information on the effects on humans is almost impossible to obtain.

4.1.1. *The proper animal "model."* It is naively assumed that the antecedents of disease can be easily studied on animals and that a demonstration that certain conditions lead to disease in animals may invariably teach us how disease is caused in humans. Unfortunately, this is not always true. Although we can learn a great deal from animal experiments, toxicological and teratological information from animal experiments turns out to be much less useful than is commonly thought. Animals react with a wide range of physical and behavioral responses to the presence or absence of chemical stimuli. There is not only a wide difference in the reaction of different animal species, but the difference becomes even more pronounced when we compare animal to man. Difficulties in inferring human reactions from animal studies are further compounded by the use of rodents as a favorite experimental animal because of their convenience and low cost as laboratory animals. However, rodents are much further removed phylogenetically from the human animal than are dogs or monkeys. On the other hand, experiments with dogs and monkeys are inordinately expensive. The problem raised by the use of rodents is clearly seen in 2,4,5-T toxicology studies, where it was found that rats may be able to maintain on doses of 100 mg/kg per day for a number of days while dogs may succumb from a single dose of 20 mg/kg [12]. Yet, most inferences about toxicity of 2,4,5-T to man were based on work with rodents. It should be noted that the problems of choosing a proper model (that is, an animal system that reacts in the same way as a human subject) have had marked effects on such recent affairs as thalidomide, riboflavin deficiencies, vitaminoses, and others.

4.1.2. *Relationship between dose and response.* As a corollary to the problem of finding the right model of animal stands the observation that every known substance (including water and oxygen) is harmful and may have toxic, teratogenic, mutagenic, and carcinogenic effects if it is given in a large enough dose [4], [7], [29]. As a consequence, the meaning of a toxic response in an animal system to a very high dose of an agent may be unclear. True, the kind of physiological or anatomical reaction occurring offers an important clue to how human tissue may respond to the same agent. But the mere fact that an agent is toxic at some high dose is not considered surprising, since a toxic reaction is to be expected for some dose. What is important is to see the rate at which a reaction disappears as a function of lower and lower doses.

Unfortunately, the investigation of the dose response curves in animals is beset by two extremely difficult problems.

(1) A low exposure does not necessarily result in a smaller reaction in an animal but, rather, in fewer animals in which this reaction may be observed. Thus, with a large enough exposure to kill all animals, the question of whether or not a particular dose is toxic creates no great experimental problems. However, an immensely large number of animals would be needed to determine doses with the toxicity to affect 1 per cent, 0.5 per cent, or 0.01 per cent of the animal population.

(2) The response of biological systems to any agent is not uniformly good or bad. It is true that biological systems do respond to a large variety of chemical agents in different ways at different doses. For instance, copper, which is a very toxic substance at high doses, is not toxic at all in low doses and is, in fact, necessary to sustain life, so that its complete absence is a definite hazard. The same is, of course, also true with many other substances, including water and oxygen. Thus, the major utility of an animal experiment is to see whether or not the reactions to relatively high doses can serve for extrapolating a dose response function that will tell if a zero or nontoxic response to some dose will occur or whether or not the agent may be toxic at any dose.

4.1.3. *Clinical (human) studies.* The best measure of toxic and teratogenic effects is man. Unfortunately, man is also the most unusable experimental subject. Three major difficulties meet the attempts to study effects of pollution on man (including that of herbicides and pesticides).

(1) Toxic agents that have dramatic effects are easily spotted and eliminated. Pollutants that affect the occasional individual, individuals who are ill from other causes, or that cause small reactions in individuals tend to go unnoticed.

(2) Even when noticeable reactions occur, the information that they do occur may be almost impossible to obtain and evaluate. Abnormal human reactions usually are recorded only when they motivate the diseased individual to seek medical attention. Thus, processes of self selection emanate which make it almost impossible to find proper control groups for prospective or retrospective studies, even if these records become available. However, the fact that records exist somewhere about a reaction to a pollutant is, in most instances, immaterial

since these recorded instances are difficult to find or, once located, to concentrate for review and analysis.

(3) There are restraints to either withdrawing or exerting a specific treatment on humans, even if an agent or treatment is deemed to be harmful (or its opposite).

4.2. *Second order uncertainty: the experimenter.* Experiments that may yield answers to such questions as those asked here require sophistication in the design, the analysis of multivariate data, the mathematical techniques for extrapolation, and data processing facilities.

Yet, what do we find? The literature contains less than two dozen key reports of studies on the toxicity of 2,4,5-T and on its teratogenicity. While most of these experiments were performed using some variation in combination of different dose levels, different concentrations at which dose levels were applied, different vehicles which carried the doses, different amounts of impurities (especially TCDD), and some were performed on different species or used different products coming from different manufacturers (and so ideally suited to factorial designs), the reports of some 22 experiments analyzed by the 2,4,5-T Advisory Committee did not contain a *single* experiment that was designed to tell something about the effect of 2,4,5-T at very low doses and attempted, by statistical analysis or mathematical techniques, to milk the available data (although inadequate) for whatever information could have been obtained about reactions of animals to very small doses.

The relevant design features and approaches to the analysis of data of ten toxicological and ten teratological experiments are summarized in Tables I and II. In view of the lack of statistical sophistication practiced in this important field, much is to be learned from their study.

All experiments were designed to test the effect of 2,4,5-T at high doses on relatively small numbers of animals. No provisions were built into the experimental design to permit picking up the effect of 2,4,5-T at a low dose. In fact, some of the conclusions reached by the scientists who reviewed these data were that 2,4,5-T has a toxic and teratogenic effect only at high doses. This impression may have been created by the fact that no adequate experimental design existed to assess the effect of 2,4,5-T at low doses.

Very serious, in view of the wealth of available statistical and mathematical techniques, is the relative naiveté with which these studies were analyzed. Many of these reports did not subject their data to any statistical analysis whatsoever. Simple summary figures, such as the arithmetic average, were in most instances not accompanied by measures of dispersion. The most sophisticated statistical analyses were multiple application of Student's t tests comparing sets of measurements from each individual group (resulting from combinations of doses and other factors) with a control group. The numbers so analyzed often were discrete rather than continuous and came from obvious nonnormal distributions. Those values of "t" that met the famous criteria of statistical significance at 0.05 were duly starred with an asterisk, adding perhaps insult to injury [28]. Thus, not a

TABLE I

SUMMARY OF RELEVANT DESIGN FEATURES AND ANALYSES OF DATA OBTAINED FROM KEY EXPERIMENTS INVESTIGATING THE TOXICITY OF 2,4,5-T

Reference	Kind of animal	Number of animals	Dose levels (and other conditions)	(Statistical) analyses
Rowe [24]	Rats Mice Guinea pigs Chicks	10 10 10 3	Different doses but amounts not given	LD_{50} (by method of Lichtfield). No statistical analyses.
Drill [12]	Dogs	48	4 dose levels each for 2 separate experiments (acute and chronic), for 2,4-D and 2,4,5-T and for males and females	Tabular and descriptive summaries (given for each animal) of body weight, symptoms, autopsy findings, blood counts, organ weights, microscopic changes. LD_{50} is estimated but no indication given on how it was done. No statistical analyses.
Rowe [25]	Rats Guinea pigs Rabbits Mice Chicks Dogs	Unclear how many	15 different herbicidal materials and 12 different herbicidal formulations	LD_{50} (by method of Lichtfield). Various qualitative observations listed. No statistical analyses.
Dow [11]	Rats	100	5 dose levels each for males and females	Tabular summary of mortality, autopsy findings, averages for food consumption, body and organ weights, blood and urine analyses. Statistical analyses limited to comparing some groups to controls using Student's t tests.
Hazelton [16]	Rats Rabbits Rabbits Rabbits	30 3 16 12	For eye irritation phase—6 dose levels in 3 concentrations. For first dermal irritation phase—4 dose levels. For second dermal irritation phase—3 dose levels.	Tabular summaries and qualitative descriptions of weekly body weights, signs of dermal irritation, blood and urine analyses. No statistical analyses.
Butler [3]	Crustacea Mollusks Fish Phyloplankton	Not given either by volume or number	Different pesticides at different water conditions as salinity and temperature and time intervals	Tables of LD_{50} values. Qualitative description of a variety of observations. No statistical analyses.
Hegyi [17]	Rabbits	32	10 different solutions	Pictures and qualitative description of a variety of observations. No statistical analyses.
McCollister [20]	Rats	50	5 dose levels each for males and females	Tables giving blood and urine counts (for each individual animal), individual body and organ weights and averages, listing of gross histological results. No statistical analyses.

TABLE II

SUMMARY OF RELEVANT DESIGN FEATURES AND ANALYSES OF DATA OBTAINED FROM KEY EXPERIMENTS INVESTIGATING THE TERATOGENICITY OF 2,4,5-T

Reference	Kind of animal	Number of animals	Dose levels (and other conditions)	(Statistical) analyses
Emerson [13]	Rats	175	6	Qualitative description of observations. Tabular summaries of organ and body weights, number of pregnancies, implantations, corpora lutea, viable pups, resorptions, average and range of pup weights, skeletal and visceral abnormalities. No statistical analyses.
Courtney [8]	Mice (two different strains) Rats	Not given. However, number of litters is.	4, 2, and 3 for different strains or time periods.	Tabular summaries of per cent (of live litters) of live fetuses, fetal mortality, abnormal litters, abnormal fetuses, cleft palate, cystic kidney, fetal and maternal average body and liver weights. Used "2 × 2 Chi Square tests" to compare porportion of abnormal fetuses with controls. (It is worthy of note that there were no concurrent controls but controls used were the cumulated experience over preceding three years.) No further statistical analyses.
Sparschu [26]	Rats	75	6 dose levels and concentrations	Qualitative description and tabular summaries of average maternal body weights, means of number of pregnancies, corpora lutea, viable fetuses, resorption sites, sexes, numbers of gross, visceral, and skeletal abnormalities. Student's t test used for comparing test groups to controls. No further statistical analyses.
Emerson [13]	Rabbits	80	4	Qualitative description of observations and tabular summaries of average maternal body weights, total and average number of implantations, corpora lutea, resorptions, viable kits, average and range of kit weights, sex and number of skeletal and visceral abnormalities. No statistical analyses.

TABLE II, (continued)

Reference	Kind of animal	Number of animals	Dose levels (and other conditions)	(Statistical) analyses
Bionetics [2]	Mice	Not given	1 dose of each of 2 compounds given via 2 vehicles and controls in each vehicle	Tabular description of observations on each individual animal and litter and tabular summary. No statistical analyses.
Collins [5]	Hamsters	6 per group	Up to 4 dose levels each for 7 different sources of 2,4,5-T and 3 sources of 2,4-D and one control group.	Tabular summaries of number and averages and per cents. No statistical analyses.
Sparschu [27]	Rats	75	2 doses and control	Tabular summaries of numbers, averages, and per cents. Student's t and Chi Square tests—individual test groups vs. control.
Bionetics [2]	Mice	Not given directly	1 dose of 2 compounds in 2 vehicles and controls for each vehicle	Tabular description of observations on each individual litter and summaries of totals, per cents, and averages. No statistical analyses.
Courtney [9]	Mice (3 strains) Rats	Not given, but number of litters is.	Up to 4 dose levels (varying) of 2 compounds each	Tabular summaries of number, per cent, averages. No statistical analyses.
Wilson [33]	Rats	Not given, but number of litters is.	4 dose levels, 3 of these given on day 9, 1 given on various days (7 to 13)	Tabular summary of total implants, per cent dead or resorbed, mean weight of survivors, per cent of survivors malformed. It is noteworthy that there are no concurrent controls. No statistical analyses.

single one of these experiments subjected its data to the robust and generally available factorial techniques (not to say anything about taking out the effect of confounding variables). Not a single one of these experiments made any attempt to extrapolate their data toward low doses.

5. Conclusion

It is not easy to understand why so much inadequate data had been generated. After all, these toxicological problems are not new and solutions exist for most of them. One simple answer might be that these practices have grown up around the need to *license* or *register* possibly toxic products and not around their widespread *use*. Another explanation might be that statisticians and mathematicians have failed to *educate* the community of scientists, especially those in medical areas. Undoubtedly, immediate interests in licensing and lack of proper preparation of scientists are influential factors. But it would be negligent not to look further. After all, a simple explanation would hardly fit a case involving so many investigators from so many different sources. Surely, lack of education or surrender to special interests are not *that* widespread in our scientific community.

There is an alternative answer to this question. Perhaps proper statistical and experimental techniques do not exist that would enable us to deal with the issues created by pollution and community health! Perhaps to derive a statement of public policy is quite a different problem from making an inference about a state of nature. For it is true that the end product of the process of inference in instances of pollution and community health is *administrative policy decisions* and not judgments about possible states of nature. These policy decisions may be to build or not to build nuclear power plants, to build them at particular sites, or to choose other strategies to deal with a power shortage. Similarly, the policy decision might be to register or license the use of a herbicide, restrict its use to certain instances, restrict its use with certain limitations on impurities, or examine alternative ways of pest control. Fortunately, the field of statistics is prepared to deal with a broad experimental strategy of the kind required to serve as a foundation for public health policy. Statisticians have brought the science of making inferences in the face of uncertainty to a high plateau. What may be needed as a next step is to assess problems created by pollutants for community health and to make decisions about community action and to develop the kinds of experiment to decision strategies that are best suited to derive rational and sensible public policies.

REFERENCES

[1] W. L. Bamesberger and D. R. Adams, "Organic pesticides in the environment," *Adv. Chem.* Ser. 60 (1966), Washington, D.C., ACS Publication.

[2] Bionetics Research Laboratories, "Teratogenic effects of 2,4,5-T in mice, report submitted to Hercules Incorporated, December, 1970," Report submitted to Hercules Incorporated, March, 1971.

[3] P. A. BUTLER, "Pesticide-wildlife studies; a review of Fish and Wildlife Service investigations; commercial fisheries investigation," Circular 167 (1963); Circular 199 (1964); Circular 266 (1965).

[4] W. CARRUTHERS, "Carcinogens related to the aetiology of bronchial carcinoma," *Physiotherapy*, Vol. 44 (1958), pp. 307–312.

[5] T. F. X. COLLINS and C. H. WILLIAMS, "Teratogenic studies with 2,4,5-T and 2,4-D in the hamster" (prepublication copy to 2,4,5-T Advisory Committee, AE-9), Department of Health, Education, and Welfare, 1971.

[6] "Conflicting philosophies over 2,4,5-T," Editorial in *Nature*, Vol. 231 (1971), pp. 483–485.

[7] J. CORNFIELD, W. HAENSZEL, E. C. HAMMOND, A. M. LILIENFELD, M. B. SHIMKIN, and E. L. WYNDER, "Smoking and lung cancer: recent evidence and a discussion of some questions," *J. Nat. Cancer Inst.*, Vol. 22 (1959), pp. 173–203.

[8] K. D. COURTNEY, D. W. GAYLOR, M. D. HOGAN, H. L. FALK, R. R. BATES, and I. MITCHELL, "Teratogenic evaluation of 2,4,5-T," *Science*, Vol. 168 (1970), pp. 864–866.

[9] D. K. COURTNEY and J. A. MOORE, "Teratology studies with 2,4,5-T and tetrachlorodioxin" (manuscript for 2,4,5-T Advisory Committee), National Institute of Environmental Health Sciences, 1971.

[10] R. T. CUTTING, T. H. PHUOC, J. M. BLLO, M. W. BENENSON, and C. H. EVANS, *Congenital Malformations, Hydatidiform Moles, and Stillbirth in the Republic of Vietnam, 1960–1969*, Washington, D.C., U.S. Government Printing Office, 1970.

[11] DOW CHEMICAL COMPANY, "Results of 90-day dietary feeding studies of Dowanol, 97B Ester 2,4,5-T in rats," Biochemical Research Laboratory Report, November 27, 1961.

[12] V. A. DRILL and T. HIRATZKA, "Toxicity of 2,4-dichlorophenoxy acetic acid and 2,4,5-trichlorophenoxy acetic acid, A report on the acute and chronic toxicity in dogs," *AMA Arch. Indust. Hygiene Occupat. Med.*, Vol. 7 (1953), pp. 61–67.

[13] J. L. EMERSON, D. J. THOMPSON, C. G. GERBIE, and V. B. ROBINSON, "Results of teratogenic studies of 2,4,5-trichlorophenoxy acetic acid in rats," Dow Chemical Company, Human Health Research and Development Laboratory Report, March 30, 1970.

[14] ENVIRONMENTAL PROTECTION AGENCY, Report of the Advisory Committee on 2,4,5-T to the Administrator, May 7, 1971, to appear.

[15] B. H. GRISBY and E. D. FARVELL, "Some effects of herbicides on pasture and grazing livestock," *Quart. Bull.*, Vol. 32 (1950), pp. 378–385.

[16] HAZELTON LABORATORIES INCORPORATED, Report to the Diamond Alkali Company, February, 1962, mimeograph.

[17] E. HEGYI, Z. ST'OTA, and A. LUPTAKOVA, "Die Rolle des Alkali im Hinblick auf die keratogene Wirkung des Technischen 2,4,5-Trichlorophenol," *Derufs-dermatosen*, Vol. 6 (1969), pp. 327–337.

[18] J. E. JOHNSON, "Symposium on possible public health implications of widespread use of herbicides," AIBS Meeting, August 26, 1970.

[19] T. C. KERNEY, "Chlorinated dioxin research," presented before a Joint Meeting on Pesticides, United Kingdom, Canada, United States, November 5, 1970.

[20] S. B. MCCOLLISTER and R. J. KOCIBA, "Results of 90-day dietary feeding study on 2,4,5-trichlorophenoxy acetic acid (2,4,5-T) in rats," Biomedical Research Laboratory, The Dow Chemical Company, September 18, 1970.

[21] M. S. MESELSON, A. H. WESTING, and J. D. CONSTABLE, "Background material relevant to presentation of the 1970 annual meeting of the AAAS," Herbicide Assessment Commission of the American Association for the Advancement of Science (revised January 14, 1971).

[22] L. A. NORRIS, "Degradation of herbicides in the forest floor," *Tree Growth and Forest Soils* (by C. T. Youngberg and C. B. Davey), Corvallis, Oregon State University Press, 1970, pp. 397–411.

[23] OFFICE OF SCIENCE AND TECHNOLOGY, *Report on 2,4,5-T of the Panel on Herbicides*, April, 1971.

[24] V. K. ROWE, D. D. MCCOLLISTER, and H. C. SPENCER, "The acute oral toxicity of 2,4,5-trichlorophenoxy acetic acid to rats, mice, guinea pigs, and chicks," Report of the Biochemical Research Laboratory, Dow Chemical Company, 1950.
[25] V. K. ROWE and T. A. HYMAS, "Summary of toxicological information on 2,4-D and 2,4,5-T type herbicides and an evaluation of the hazards to livestock associated with their use," *Amer. J. Veterinary Res.*, Vol. 15 (1954), pp. 622–629.
[26] G. L. SPARSCHU, F. L. DUNN, and V. K. ROWE, "Teratogenic study of 2,3,7,8-tetrachlorodiobenzo-p-dioxin in the rat," Biochemical Research Laboratory Report, Dow Chemical Company, 1970.
[27] G. L. SPARSCHU, F. L. DUNN, R. W. LISOWE, and V. K. ROWE, "Study of the effect of high levels of 2,4,5-trichlorophenoxy acetic acids (2,4,5-T) on rat fetal development," Chemical Biological Research Report, Dow Chemical Company, 1971.
[28] T. D. STERLING, "Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa," *J. Amer. Statist. Assoc.*, Vol. 54 (1959), pp. 30–34.
[29] T. D. STERLING, "Epidemiology of disease associated with lead," *Arch. Environ. Health*, Vol. 8 (1964), pp. 333–348.
[30] UNITED STATES DEPARTMENT OF AGRICULTURE, "Progress report on dioxin research IV," March 25, 1970, Unpublished report.
[31] U.S. DEPARTMENT OF HEALTH, EDUCATION, and WELFARE, "Report of the Secretary's Commission on Pesticides and their Relationship to Environmental Health," Washington, D.C., 1969.
[32] S. R. WEIBEL, R. B. WEIDNER, J. M. COHEN, and A. G. CHRISTIANSON, "Pesticides and other contaminants in rainfall and runoff," *J. Amer. Waterworks Assoc.*, Vol. 58 (1966), pp. 1075–1084.
[33] J. G. WILSON, experiment reported to the 2,4,5-T Advisory Committee, 1971; this experiment is part of the Final Report of the 2,4,5-T Advisory Committee [23].

Discussion

Question: E. B. Hook, Birth Defects Institute, Albany Medical College

Would you specify the nature of the increase in congenital malformations in the AAAS Vietnamese study (that is, which specific defects were increased), and what order of magnitude of increase there was, and what was the "denominator" of population studied?

Also, is there evidence that the other civilian ravages of war in Vietnam were associated with exposure to herbicide spraying?

Reply: T. Sterling

The HAC study found in some geographic areas an increase in spina bifida and cleft palates. These are the same malformations also noted in the teratogenicity animal studies. The increase was noticeable rather than impressive. However, this does not mean much because these data are full of rather large errors. The denominator of the population studied, incidentally, was, of course, all the children recorded in a given hospital.

Question: Dominick Mendola, Ecology, San Diego State University

In your paper you referred to a report utilizing radioactive tracer techniques to trace the fate of dioxin in soils. You said that there was a 15 per cent to 20 per cent decrease of dioxin applied to soils after 160 days. Do you know how

long this experiment was run and how much, if any, dioxin was found after the entire length of the experiment? Also, what were the soil conditions—wet or dry?

Reply: T. Sterling

I expect the experiment was run for 160 days. What was measured was the radioactivity remaining after that period of time. That this radioactivity was labeled dioxin which had been put into the soil to begin with is, of course, assumed. The soil condition was sandy but not dry.

Question: J. B. Neilands, Biochemistry, University of California, Berkeley

Is it not a fact that the increased incidence of birth defects in South Vietnam has more or less paralleled the deployment of herbicides, that is, starting in about 1961 and tapering off in the later years of the decade when the defoliation operation (but not the bombing and intensity of the war) was cut back? And is it not a fact that the highest levels of birth defects reported in South Vietnam were in Tay Ninh, the province most heavily defoliated?

Can you give us an estimate as to how many agencies of the U.S. government had access to the report of the Bionetics Laboratories and also the time when the report was available to the agencies? My motivation in asking this is the following. Some years ago the UN voted 80 to 3 to class herbicides as prohibited weapons under the 1925 Geneva Protocol—only Australia and Portugal joined the U.S. in voting "no"—and 95 nations have ratified this instrument. Even though the U.S. has not yet ratified the Protocol, we have admitted we are bound to it by customary international law and under the Constitution the latter is considered part of the body of U.S. law. Statistics aside, it is clearly a *war crime* to employ herbicidal sprays in military operations. Domestically, 2,4,5-T has also been used on a massive scale, again with assurances that it is innocuous to animal life. Yet we now know that a manufacturer was forced to close a 2,4,5-T plant because of chloracne, a dermatitis, in the workers. Thus if the Department of Agriculture knew of the toxicity of preparations of 2,4,5-T and delayed action against the herbicide, officials of that Department would seem to be irresponsible and derelict in their duty to protect the health of the American public. Can you comment on the legal status of 2,4,5-T?

Reply: T. Sterling

As both the Army and the HAC reports show, the questions on the correlation between birth defects and defoliation are not easy to answer. Certainly the analyses of the data conducted by Dr. Meselson (the HAC report) find some support for the increase in birth defects paralleling use of the defoliant. His work also uncovered that the Tay Ninh Province reported the highest level of birth defects.

Everything I have seen in my contact with 2,4,5-T would indicate that the actions of the government officials involved were extremely responsible. The Department of Agriculture, the Department of the Interior, and the Department of Health, Education, and Welfare took action immediately as soon as evidence emerged that 2,4,5-T had teratogenic potentials. While there are some scientists

and industrialists who might even accuse government officials of acting in haste, I surely do not think we would find many individuals who would accuse officials of the Department of Agriculture (or of other departments) of having been derelict in their duty to protect the health of the American public. If I may add, a certain amount of fault rests on all of our shoulders. We have now been willing to tolerate sloppy experiments and data in the area of toxicology and public health to such an extent that rational and reasonable action becomes more and more difficult to take. Perhaps we, as scientists, ought to re-examine our procedures first before examining the actions of government officials who are much less well equipped to deal with the states of uncertainty created by inadequate scientific procedures.