THE VALUE OF LIFE SAVING:

LESSONS FROM THE CIGARETTE MARKET

Pauline M. Ippolito*

Federal Trade Commission

ABSTRACT

This paper reports the results of a study of the cigarette market designed to measure what consumers are willing to pay for increased life expectancy. The spread of information on the health effects of smoking has dramatically transformed the cigarette market in the last 30 years. Using survey evidence on consumer beliefs together with reductions in cigarette demand gives a direct estimate of consumers' valuation of safety. The average "value of life" is estimated to be $460,000 (1985$), but consumers are found to differ considerably in their valuations. Moreover, there appears to be a skewness to the value of life distribution; the median value of life is approximately $380,000, compared to the mean of $460,000.

KEY WORDS: Value of life, information, cigarettes, beliefs, safety policy, risk.

I. THE CIGARETTE MARKET

There is probably no market in America today that has been more affected by consumer reactions to health information than the cigarette market. Over the last 30 years, the continual flow of information to consumers on the dangers of smoking has no serious competition as an explanation for the dramatic switch to safer smoking habits. The number of people who smoke today and the types of cigarettes they smoke are very different from what would have been expected on the basis of the market's behavior before 1952, when the first Readers' Digest and Consumer Report articles began appearing on the hazards of smoking. Moreover, there is a wealth of data available on the cigarette market, including survey data on consumers' beliefs about the risks of smoking. These features make the cigarette market an ideal candidate to measure how much consumers are actually willing to give up for an increase in life expectancy.

---

* The views expressed in this paper are those of the author and do not necessarily reflect the positions of the Federal Trade Commission or of any individual Commissioner. I am indebted to Gerard Butters, Richard Ippolito and Paul Portney for helpful comments on an earlier draft.

A. The Consumer's Decision To Smoke Cigarettes

When a consumer makes the decision to smoke a particular type of cigarette, he presumably judges the enjoyment of smoking to be greater than the perceived costs. These costs are of two types: the direct money cost of cigarettes and a perceived health cost based on the individual's beliefs about the hazards of smoking. Moreover, the consumer must choose a type of cigarette. Here again the tradeoff is enjoyment versus safety. In general, higher "taste" cigarettes are also higher risk cigarettes, since the components of smoke that contribute to taste also add to the health risk. Nicotine is a convenient index of both taste and risk.[1]

Over the last 30 years, consumer perceptions of the health cost of smoking have changed substantially. This change is responsible for a large reduction in the demand for cigarettes and for a substantial switch in the type of cigarette smoked. These reactions allow us to measure how much consumers willingly gave up for the expectation of longer lives.

B. The Cigarette Market in 1980

The best recent evidence on consumer beliefs about the hazards of smoking is a nationally projectable U. S. survey conducted by the Roper Organization for the Federal Trade Commission in November 1980. Because of this evidence on consumer beliefs, 1980 was used as the base year for this study. In the survey, individuals were asked to judge the truthfulness of an assertion that 30-year-old pack-a-day smoker had a lower life expectancy than a comparable non-smoker. Those who answered in the affirmative were then asked how many years of life were lost on average. The survey results are shown in Table 1.

While these results are suggestive of consumer knowledge of smoking risks, some problems are evident. For instance, 30 percent of the population (and 40 percent of smokers) deny that smoking affects life expectancy. Yet, these responses are directly contradicted by the individuals themselves in other parts of the survey and by overall market response behavior.[2] Notwithstanding these shortcomings, the survey results suggest several qualitative features of current consumer knowledge on smoking.

First, on average, individuals do not appear to underestimate the risk of smoking. Even taking the survey at face value, and ignoring non-responses, beliefs about the life loss have an average value of approximately 3.5 years lost, which closely corresponds to epidemiological estimates.[3] More realistically, if those who responded that smoking had no life expectancy effects are placed in the 0-2 year category to adjust for the response bias, the average belief rises to 4.67 years. Second, the belief distribution is not symmetric. While a greater number of people underestimate the life loss of smoking, those who do overestimate, do so by a greater margin on average. Third, consumer beliefs show a relatively wide variance, enough to suggest that differences in beliefs are an important determinant of behavior in the cigarette market.

According to surveys conducted by the National Center for Health Statistics (HHS (1982)), approximately 33 percent of the adult population smoked cigarettes in 1980. Cigarette production statistics indicate that per capita consumption was 195 packs per year.[4] Nicotine content by brand and variety and corresponding market share data (FTC (annual) and Maxwell (1981)) indicate that the average cigarette sold in 1980 had approximately 1 milligram of nicotine. Nearly 30 percent of cigarettes sold had less

TABLE 1

Consumer Beliefs About the Life Expectancy Cost of Smoking, 1980[1]

Estimated life expectancy loss from smoking[2] (Years)	Survey Response Distributions (Percent)		
	Total	Smokers	Non-smokers
Zero	30.4	40.9	24.7
Less than 2	5.2	5.6	5.0
2-4	11.9	13.3	11.3
4-6	15.5	14.2	16.2
6-8	10.0	8.0	11.0
8-10	10.7	6.2	13.1
More than 10	4.6	2.7	5.6
Don't know how much[3]	11.7	9.1	13.1
Total	100.0	100.0	100.0

Source: Roper Survey, November 1980.

1 In 1980, the Federal Trade Commission asked the Roper Organization to include a number of smoking questions in their 1980 random survey. The survey included 1005 individuals, including 339 smokers, reflecting the national smoking rate.

2 Individuals were asked whether a 30-year-old person reduces his life expectancy if he smokes at least one pack a day for life. If answered in the affirmative, the respondent was then asked to estimate the life expectancy cost.

3 These individuals said they thought that smoking reduced life expectancy but were unable to assign a particular number of years to the loss.

than .75 milligrams of nicotine and about 13 percent had more than 1.25 milligrams.

C. The Cigarette Market That Would Have Existed in 1980 Without Health Information

If there had never been disclosures about the health effects of smoking, the 1980 cigarette market would have been very different from the one just described. While it is always somewhat precarious to attempt predictions of the world that might have been, the cigarette market is one where such predictions can be made with some confidence. In most respects, the cigarette market was on a stable path prior to the health discoveries. There was a strong growth trend in the incidence of female smoking, a more modest growth in male smoking, and virtually no change in the product itself for the twenty-five years before the first disclosures. Using standard statistical techniques and data from the HHS surveys covering 1947-1975, it has been estimated that if there had been no health information, approximately 54 percent of the adult population would have smoked and per capita consumption would have been 586 packs per year by 1980.[5]

Finally, the market share of the different nicotine-type cigarettes is projectable from simple historical evidence. Prior to 1953, a few non-filter cigarettes with very similar nicotine contents dominated sales for over twenty-five years (Maxwell (1975)). The sales-weighted nicotine content of cigarettes sold was virtually stable from 1926 to 1953. Even when new brands entered, their nicotine content was essentially the same as those already in the market. Therefore, it seems reasonable to assume that absent the health concerns, nicotine content would have remained constant and virtually all cigarettes would have had the same nicotine content. Using FTC and Maxwell data, the nicotine content of these no-information cigarettes is put conservatively at 1.49 milligrams.

This basic description of smoking behavior in 1980 with the health information and what it would have been without the health information[6] is shown in Table 2. From these estimates it is clear that consumers have reacted dramatically as their beliefs about the health risks have changed. By 1980, per capita consumption had fallen to approximately 50 percent of its projected level and the average nicotine content of cigarettes had been reduced by at least one-third. As more consumers become convinced of the risks of smoking, demand should fall still further and the distribution of cigarette types should continue its trend towards low tar/low nicotine cigarettes.

II. LESSONS FROM THE CIGARETTE EXPERIENCE

What can be learned about consumers' willingness to pay for longer lifespans from this reaction in the cigarette market? In a recent study (Ippolito and Ippolito (1984)), estimated reductions in individuals' demands for cigarettes were compared with their changed beliefs about the life expectancy effects of smoking.[7] Since the reductions in demand were the direct result of the changed beliefs, they provide a clear measurement of how much consumers were willing to pay (in reduced smoking pleasure or more technically, in reduced consumer surplus) to increase their life expectancy. More specifically, the vertical shifts in consumers' demand curves reveal their perceptions of the dollar value of the hidden health cost of smoking. Using the adjusted Roper survey as a measure of consumers' beliefs, this health cost translates directly into a "value of life" measure.

TABLE 2

Available Evidence About Smoking Behavior
With and Without Health Information, 1980.

Smoking Behavior	Without Information	With Information
Per capita cigarette consumption (packs per year; 18 years old and over)	386*	195
Percent of population smoking (18 years old and over)	54.2*	32.5
Elasticity of cigarette demand	-0.48*	-0.48*
Nicotine content per cigarette smoked (milligrams)		
Mean	1.49*	0.996
Standard Deviation	0.00	0.34

Source: The numbers not marked by an asterisk are data reported by or calculated from published sources in 1980. The numbers marked by an asterisk are estimates of what smoking behavior would have been in 1980 if cigarette-health disclosures had never been made available. All data sources are described in the text. Detailed estimations are in Ippolito and Ippolito (1984).

Three types of consumer differences were accounted for in the study: differences in beliefs (taken to be reflected by the adjusted Roper survey), differences in the underlying taste for cigarettes, and differences in the "value of life" itself. These last two factors in the cigarette reaction were estimated as part of the study by assuming particular functional forms for the underlying distributions of tastes and values of life and then finding the distribution parameters that best fit the aggregate reaction.

A. Consumers Differ in Their Willingness to Pay For Safety

The average "value of life" for the population estimated from the cigarette reaction is approximately $460,000 (1985 dollars).[8] This implies that on average individuals are willing to pay up to $460 to reduce the risk of death by 1/1000 or up to $46 to reduce the risk of death by 1/10,000.[9]

This estimate is in the lower range of those in the literature. It is based on a cleaner situation from which to measure the willingness to pay for safety -- one where there is a direct connection between the observed behavior and the risks to life. Moreover, the study corrects for possible errors in consumers' beliefs about the risks in question and for differences in consumers' valuations of safety. That these improvements in study methodology lead to lower estimates than many in the literature suggests that more serious attention must be given to these issues and to the potential bias they introduce into measures of the "value of life."

While the average willingness to pay for safety is interesting in its own right, it obscures a potentially important variation across the population. This study was specifically designed to estimate the distribution of consumers' valuations of safety. The estimated population density is shown in Figure 1. It is apparent from the figure that consumers vary greatly in their preferences. While the mean "value of life" is $460,000, a substantial portion of the population has values significantly above or below this average. The standard deviation of the estimated distribution is approximately $350,000 or about $350 to remove a 1/1000 risk of death. This variation in willingness to pay for risk reduction suggests that selection problems may indeed color estimates drawn from cross-section data. For instance, those who continued to smoke in 1980 were estimated in this study to have an average value of life that is approximately half that of non-smokers, that is, approximately $275,000 versus $550,000 (1985 dollars).

Equally significant is the decidedly skewed nature of the distribution. More than 40 percent of the population is not willing to spend even $335 to eliminate a 1/1000 risk of death, but 20 percent of the population is willing to spend more than $670 to remove the same risk. These figures reflect the relatively large portion of the population concentrated in the lower ranges of the willingness-to-pay distribution; and the smaller portion of the population in the more skewed right tail of the distribution.

This result probably comes as no surprise to those who have attempted to market safety; some consumers are willing to pay a sizable premium for safety, but a relatively large portion of the population is not. In particular, this estimate implies that a majority of the population (nearly 65%) would not be willing to pay the estimated mean of $460 to eliminate a 1/1000 risk of death. If these results are valid, this distribution has important implications for marketers and for safety policy.[10]

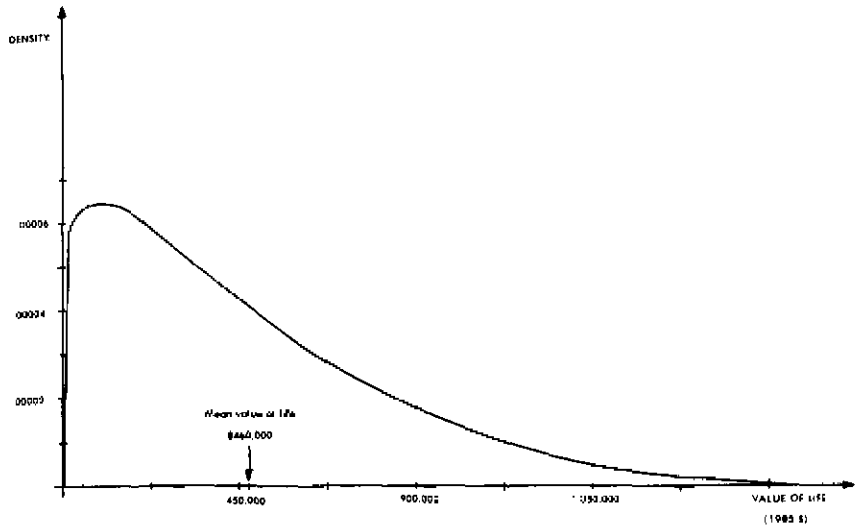

Figure 1

Estimated Value of Life Density Function from Cigarette Market Reaction

TABLE 3

Smoking Reaction to Health Information By Type of Smoker, 1980

Individual's Smoking Rate Without Information (Packs/day)	Average Reduction in Smoking Rate With Information[1] (Percent)		Percent Who Quit Smoking
	Including Those Who Quit	Those Who Continue to Smoke	
1	80%	44%	65%
2	60	38	37
3	44	33	16
4	29	27	3

Source: Estimates of cigarette consumption based on Roper survey of beliefs about smoking health hazards and aggregate cigarette consumption data from USDA and HHS sources (see text).

1 Reduction in smoking rate does not account for the related adjustment in the type of cigarette smoked. Virtually all who continued to smoke in 1980 also reduced the tar and nicotine content of the cigarettes they smoked.

In the regulatory arena, for instance, this skewed distribution implies that mandatory safety standards that are consistent with consumers' average willingness to pay for safety will seem excessive to the majority of the population, but will be decidedly too weak for those who value safety highly. This factor makes labeling and other information approaches to safety regulation more attractive, since the information approaches allow greater freedom for the market to satisfy widely varying safety preferences. Similarly, in liability matters, standards of care that are based on less than the average willingness to pay for safety might be appropriate; this is especially true if there is a defense based on a lower price relative to otherwise similar items on the market or on disclosure of relative safety features of the product.

B. Underlying Consumer Characteristics Affect Overall Reactions to Health Information

In estimating the underlying taste for smoking (before health concerns), it was found that consumers who had a "high taste" for cigarettes also had a lower sensitivity to price changes.[11] Thus, while all consumers reacted to the health information, light smokers reacted more than heavy smokers. This helps explain why on average current smokers consume nearly as many cigarettes as smokers in the past.

Table 3 shows the estimated reduction in the smoking rate of different types of smokers caused by the change in beliefs about smoking. Those who would have smoked 1 pack a day without the health information reduced their smoking by 80% on average. This reduction came from two factors: 65% of 1 pack-a-day smokers quit smoking altogether and those who continued to smoke reduced their rate by 44% on average (that is, to a little more than half a pack a day). In contrast, those who would have been 4 pack-a-day smokers reduced their consumption by only 29% (that is, to about 2.8 packs per day). This reduction is a combination of the 3% who quit smoking and a reduction of 27% for those who continued to smoke.

This result illustrates the importance of underlying demand characteristics in predicting the consumer response to health information. Other things equal, consumers with a higher demand elasticity will respond more to new health risks. The overall pattern of consumer reaction thus depends directly on the distribution of price sensitivity across the consuming group, as well as on consumers' ability to absorb the information and on their valuation of the safety itself.

In particular, the effectiveness of hazard warnings (or other information policies) cannot be judged solely by whether all groups of consumers have reacted similarly to the information One should expect instead that some groups of consumers will put a relatively greater value on the consumption of the hazardous product and will therefore reduce their consumption less in response to the warnings. For instance, even if the saccharin labels are working perfectly, consumers for whom overweight problems are a serious health issue will probably cut their saccharin consumption by less than other consumers -- for them, low calorie food products may simply be more valuable than for the average consumer.

C. Most Consumers Do Absorb Significant Health Information: Differences Remain Across Groups

One of the clear lessons from the cigarette experience is that most consumers will absorb significant health information that affects them

(and will act on this new information). Measuring consumer beliefs is always a difficult exercise, subject to many problems of survey design. Nonetheless, the pattern of responses to a wide variety of questions about the health hazards of smoking strongly supports the view that most consumers are aware of the fundamental cigarette health issues. Knowledge seems weaker on some of the particulars of the health risks and across some subgroups of the population, but the general health concern appears to have been absorbed quite well.

On the basic issue of whether smoking is "hazardous to health," for instance, a 1978 Roper survey done for the Tobacco Institute (FTC 1981) found that only 5% of the population responded "Smoking isn't hazardous" and only 4% responded "Don't know." Even smokers had a high level of acknowledgment on this issue with only 8% responding "Smoking isn't hazardous" and only 5% responding "Don't know."

The more specific question of whether a 30 year old who smokes at least a pack a day reduces his life expectancy is a good illustration of the general pattern of the more specific survey results (Roper 1980). As shown in Table 4, the population as a whole reports a somewhat lower (but still high) level of knowledge of this fact: 69.7% of the population chose "know" or "think it's true," 22.1% chose "don't know if it's true," and 7.8% chose "know" or "think it's not true." Smokers and non-smokers show a distinctly different pattern, with smokers reporting significantly less knowledge of this fact.[12]

Beliefs about smoking's effects also generally differ by income and education level. The results for the life expectancy question for different income groups are shown in Table 5 and are typical of the survey responses generally. Overall, higher income and education groups appear to have absorbed the health information about cigarettes more completely than lower income or education groups. Only 3% of those earning more than $25,000 per year responded that they "know" or "think it's not true" that life expectancy is reduced. Seventy eight percent responded that they "know" or "think it's true." In contrast, for the lowest income group (under $7000 per year), the corresponding figures were 17.7% for "know" or "think it's not true" and 53.2% for "know" or "think it's true."

This relationship between beliefs and income or education is consistent with reported smoking behavior. Surveys show that individuals in higher income and education groups are less likely to smoke cigarettes (Roper 1980 or HHS 1979) and are more likely to smoke low tar and nicotine cigarettes when they do smoke (HHS 1981).

It is plausible that the reported differences in beliefs are directly responsible for the observed differences in smoking behavior. This is the approach taken in measuring the value of life above. It is also possible, however, that higher income and education groups place greater value on safety, and for this reason, have been more attentive to the particulars of the health information; if so, different concerns for safety have led to the formation of different beliefs about the risks. Similarly, the fact that higher income and education groups have absorbed the health information more completely may reflect a relative efficiency in processing information or it may reflect information flows that have somehow have been directed towards these groups.[13] These are potentially fruitful areas of research that have not been explored to date.

III. CONCLUSION

In the last 30 years, the cigarette market has been dramatically

TABLE 4

Consumer Beliefs About Life Expectancy Effects of Smoking, 1980

Smoking Reduces Life Expectancy[1]	Survey Response Distribution (Percent)		
	All	Smokers	Non-smokers
True	69.7	59.0	75.2
Don't Know	22.1	26.8	19.7
Not True	7.8	13.8	4.5
No Answer	.5	.3	.6

Source: Roper Survey, November 1980.

TABLE 5

Consumer Beliefs About the Life Expectancy Effects of Smoking, By Income Group, 1980

Smoking Reduces Life Expectancy[1]	Income Group Response Distributions (Percent)			
	<$7,000	$7-15,000	$15-25,000	>$25,000
True	53.2%	65.8%	73.2%	78.2%
Don't Know	28.2	24.7	20.1	18.4
Not True	17.7	9.1	6.7	3.0
No Answer	.8	.4	-	.4

Source: Roper Survey, November 1980.

1 Individuals were asked whether a 30 year-old person reduces his life expectancy if he smokes at least one pack a day. Responses were "Know it's true", "Think it's true", (grouped as "true" in the table), "Know it's not true", "Think it's not true", (grouped as "not true" in the table), and "Don't know if it's true", and "No Answer."

transformed by the spread of information on the health effects of smoking. There is considerable evidence that consumers have absorbed much of this information and have acted on it. Virtually nobody in the United States today smokes the type of cigarette that had dominated the market for 25 years before the health discoveries. Compared to the world that would have been without the health information, by 1980 the portion of the population that smoked cigarettes was 40 percent lower, per capita consumption was 50 percent lower, and the average nicotine content of cigarettes was at least one third less.

Based on the evidence in the cigarette market, the need for government imposition of direct safety regulation (as opposed to the provision of safety information) appears to be less than often supposed. Information has been quite effective in changing consumers' beliefs about smoking and has led most consumers to alter their consumption patterns appropriately. Moreover, because consumers are found to differ substantially in their valuation of safety, uniform safety standards result in too much safety for many consumers and too little safety for others. In those cases where regulation may still be justified, the best estimate of the mean "value of life" from this study is approximately $460,000 (in 1985 dollars) and under a number of sensitivity tests stays under $1 million.

Finally, a common complaint about efforts to measure the value of safety and its use in safety policy can be clearly addressed with the cigarette experience. It is often argued that "value of life" estimates are too low for policy use, because low income individuals' ability to pay for safety is too strong an influence in the estimates. The willingness to pay for safety measured in the cigarette market is free of these income concerns -- once informed of the potential hazard, individuals must pay an out-of-pocket cost to expose themselves to the risks of smoking. The fundamental point of all value of life studies -- that individuals are willing to pay only limited amounts for safety -- may be more convincing to skeptics when drawn from studies where income does not limit individuals' choices.

NOTES

[1] The amount of nicotine in a cigarette is highly correlated with the other components of cigarette smoke (FTC), e.g., the amounts of tar and carbon monoxide, the components most directly linked to the health effects of cigarettes. The "taste" of cigarettes is generally tied to the amount of smoke in each puff, and thus, is also linked to the amount of tar, nicotine and carbon monoxide. Nicotine, in particular, seems to cause the pleasurable physiological effects on the brain that makes smoking attractive. The general technology for increasing the safety of cigarettes is to dilute the smoke in each puff through a variety of techniques, thus creating the link between "taste" and safety.

[2] For instance, in the same survey, only 2.9 percent of the population and 5.1 percent of smokers denied that smoking causes lung

cancer, a widely acknowledged fatal disease. Moreover, aggregate data shows that while 40 percent of smokers may say that smoking does not cause early death, only seven percent persisted in smoking non-filtered cigarettes in 1980 (Maxwell (1981)).

[3] Based on smoking surveys and subsequent follow-ups upon the death of respondents, several studies have estimated the effects of smoking on life expectancy.] The estimates are generally based on the consumption of pre-information nicotine content cigarettes. Standardizing to the same intensity levels, the results range from 2.3 years to 4.8 years of expected life lost for lifetime pack-a-day smokers (Hammond (1967), Ippolito et al. (1979), and U.S. Department of Health, Education and Welfare (1979)).

[4] The HHS surteys also contain quantity information, but these responses are known to be seriously biased. When compared to sales figures, it is clear that consumers significantly underestimate the amount they smoke when responding to surveys.

[5] A more detailed discussion of these estimates is available in the appendix to Ippolito and Ippolito (1984). When extrapolated to 1980, other studies of per capita consumption in the cigarette market are generally consistent with the predictions here (see Hamilton (1972), Ippolito et al. (1979), Klein et al. (1981) and Porter (1985)).

[6] In characterizing these projections as pertaining to the market without information, it is assumed that prior to 1952 consumers were generally not aware of the life-threatening risks to smoking. To the extent that these risks were known, the "value of life" estimates below are biased low.

[7] To the extent that consumers were reacting, in part, to other newly discovered health costs of smoking, the estimated "value of life" is biased high, since it includes these other health costs. The estimate also does not treat smoking as an "addiction." Studies show that there are physical withdrawal effects to smoking, but that these effects are significantly reduced within a week of quitting (Krasnegor 1979). Empirical estimates to test whether the addictive characteristics of smoking are significant in explaining aggregate smoking behavior suggest that they are not. If the addictive characteristics of smoking were important in the aggregate, it would follow that after 1964 (the date of the first Surgeon General's Report), the reduction in start rates would have been proportionally larger than the corresponding reduction in overall participation rates; that adjusting for other factors, pre-1964 starters would smoke either more cigarettes or higher nicotine content cigarettes than post-1964 starters; and that post-1964 quit rates would be lower for older smokers than for younger smokers. Available empirical evidence rejects these hypotheses (Ippolito et al. (1979)). If there is a one-time cost of changing smoking behavior, our estimated value of life is biased low by that amount; however, it is unlikely that this cost is large enough to affect the order of magnitude of the estimates.

[8] A more detailed description of the theory and estimating procedure is outlined in Ippolito and Ippolito (1984) where the estimates are expressed in 1980 dollars. The Consumer Price Index was used to inflate these 1980 estimates (conversion factor is 1.28).

This estimate assumes that consumers discount the future only to account for the risks of survival. If there is an additional subjective discounting of the future above and beyond the risk of survival, these estimates would increase somewhat. For instance, if there is an

additional (real) discount rate of 1.25%, then the estimated mean "value of life" would increase by about 10 percent to $505,000. For added discount rates of 2.5 percent and 5 percent, the estimates increase to approximately $560,000 and $765,000 respectively.

Similarly, the estimate is based on regression coefficients that are subject to statistical uncertainty. Using the 95% confidence bounds on the regression coefficients that underlie the estimates in Table 2 to determine the sensitivity of the estimates, the mean value of life is bounded between $276,000 and $844,000.

Finally, the estimate is based on the adjusted Roper survey where those who denied any life expectancy effect were placed in the lowest response category of 0-2 years of life lost. Alternative adjustments where these individuals were placed in a 0-1 year category or in a 0-1/2 year category were decidedly inferior in fitting the aggregate data.

[9] A variety of studies have used labor markets to estimate the wage premiums attached to risky occupations as a means of estimating the value of safety. These studies, which generally assume that workers are correctly informed about the risks, use highly aggregated industry-level data to attempt to isolate the safety effect from the variety of other factors that determine wages. The studies generate "value of life" estimates ranging from approximately $505,000 (Thaler and Rosen) to about $3,500,000 (Brown) in 1985 dollars. See Bailey (1980) and Blomquist in Jones-Lee (1982) for reviews of this literature. Moreover, because these studies estimate the market price of risk, it is impossible to determine whether these estimates represent "value of life" figures for individuals with above or below average valuations of safety.

[10] A recent British survey which asked individuals directly about their willingness to pay for safety (rather than estimate it from actual market behavior) found the same type of skewed distribution, though with substantially higher values of life (Jones-Lee et al. (1985)).

[11] In particular, the individual's demand for cigarettes was specified as $Q=r-cP/r$ where P is price per pack, Q is quantity of packs purchased per year, c is an estimated constant and r is a "taste" parameter that varies across individuals and reflects income and other individual-specific factors that affect consumption. Simpler demand specifications where individual specific factors did not affect price sensitivity (such as $Q=r-cP$) were decidedly inferior in fitting the aggregate data.

[12] Certainly these results are colored by some degree of response bias where both smokers and non-smokers rationalize their choices, but the differences seem large enough to suggest a real difference in beliefs.

[13] For instance, the cigarette ad ban on electronic media has the effect of reducing the amount of advertising for low tar and nicotine brands that lower income groups are exposed to, since they are more intensive users of these media. Most of the information on the health effects of smoking has appeared in print media.

REFERENCES

Bailey, Martin J. 1980. Reducing risks to life. Washington, D.C.: American Enterprise Institute for Public Policy Research.

Federal Trade Commission. Annual. Report of tar and nicotine content. Washington, D.C.: Federal Trade Commission.

Federal Trade Commission. 1981. Staff report on the cigarette advertising investigation. Washington, D.C.: FTC.
Hamilton, James L. 1972. The demand for cigarettes, advertising, the health scare and the cigarette advertising ban. Review of Economics and Statistics. 39:401-411.
Ippolito, Pauline M. and Richard A. Ippolito. 1984. Measuring the value of life saving from consumers reactions to new information. Journal of Public Economics. 25:53-81.
Ippolito, Richard A., Denis Murphy and Donald Sant. 1979. Consumer responses to cigarette health information. Staff Report to the Federal Trade Commission. Washington, D.C.
Jones-Lee, M. W. Editor. 1982. The value of life and safety. Amsterdam: North-Holland.
Jones-Lee, M. W. et al. 1985. The value of safety: results of a national sample survey. The Economic Journal. 95:49-72.
Klein, Benjamin, Kevin Murphy and Lynne Schneider. 1981. Government regulation of cigarette health information. Journal of Law and Economics. 24:575-612.
Krasnegor, Norman A. Editor. 1979. Cigarette smoking as a dependent process. National Institute on Drug Abuse Research Monograph 23.
Maxwell, John P. 1975. Historical sales trends in the cigarette industry. Richmond, Virginia: Maxwell Associates.
Maxwell, John P. 1981. Historical sales trends in the cigarette industry. New York: Lehman Brothers Kuhn Loeb Research.
Porter, Robert H. 1985. The impact of government policy on the U.S. cigarette industry. In Empirical Research in Consumer Protection Economics. Pauline M. Ippolito and David S. Scheffman, Editors. Washington, D.C.: Federal Trade Commission.
U.S. Department of Health and Human Services. Annual. The health consequences of smoking. Washington, D.C.: GPO.

VALUING FOOD SAFETY*

Tanya Roberts

United States Department of Agriculture
Washington, D.C. 20005-4788

ABSTRACT

This paper reviews methods for assessing the economic costs of foodborne disease and estimates the annual costs of salmonellosis and campylobacteriosis. These foodborne diseases cause intestinal disturbances in approximately 4.1 million Americans annually. Annual medical costs and lost wages from salmonellosis and campylobacteriosis are estimated at $1.4 to $2.6 billion. Inclusion of the economic value of leisure time lost and other factors would increase these cost estimates.

Chicken was associated with 9.5 percent of the outbreaks of salmonellosis reported in 1981, and fresh chicken may cause half of the cases of campylobacteriosis. Irradiation is one method proposed to reduce the incidence of these diseases caused by chicken. Irradiation is estimated to have a favorable benefit/cost ratio of between 2.2 to 4.2. Estimated net benefits range between $186 to $498 million annually.

KEY WORDS: food safety, foodborne disease, salmonellosis, campylobacteriosis, food irradiation, benefit/cost analysis.

Two major outbreaks of foodborne disease occurred this year in the United States. One was an outbreak of salmonellosis from contaminated milk that infected approximately 16,000 persons. The other was an outbreak of listeriosis caused by contaminated soft cheeses that resulted in approximately 100 deaths. These outbreaks are noteworthy because of the unusually high incidence and potential economic implications. These outbreaks illustrate that modern sanitation and processing have not eliminated foodborne disease.

Although some diseases such as typhoid fever have decreased markedly, other foodborne diseases are thought to be increasing (Kamplemacher, 1985). The reasons are diverse and range from practices on the farm to the kitchen: Greater concentration of animals in larger production units permits easier transmission of disease from one animal to the other. The considerable geographic movement of animals and birds can spread disease

* The views expressed here are those of the author and not necessarily those of the U.S. Department of Agriculture.

across the countryside. Today the use of improperly processed animal byproducts and wastes in animal feeds can introduce and perpetuate disease cycles. Concentration of animal slaughter in fewer and larger plants increases the possibilities for cross-contamination between carcasses (Schwabe, 1985, pp. 552-3; Snoyenbos, 1985). An increased number of distribution stages means more mass production of food and the greater inherent possibilities of improper heating and refrigeration -- two of the most common contributors to foodborne disease in meat and poultry (Bryan, 1980). People are traveling more and eating more exotic foods and being exposed to a greater variety of foodborne hazards. Finally, the organisms themselves have been evolving. They are adapting to modern food processing and are more able to survive (Archer, 1985). Also, they are developing resistance to human drug therapies (Holmberg, et. al., 1984).

Another factor is our changing life styles--eating more of our meals away from home where additional health risks from further handling of food and improper preparation may occur. The 20 percent of the meals consumed outside the home (Consumers, 1985) caused 68 percent of the foodborne disease outbreaks reported in 1981 (CDC, 1983). Conversely, the 80 percent of the meals consumed at home caused only 32 percent of the reported foodborne outbreaks. But perhaps this is an unfair inference to make because of the severe underreporting of foodborne disease and perhaps the greater likelihood that food poisoning at home will go unreported, undiagnosed, or misdiagnosed.

The severity of recent foodborne disease outbreaks and the accompanying publicity have renewed interest in developing estimates of the economic costs of these outbreaks and identifying and quantifying food safety management techniques. This paper discusses several of the practical difficulties with making such estimates: (1) valuing the growing number of sectors associated with a foodborne disease outbreak, and (2) estimating the costs of control The specific foodborne diseases evaluated here are salmonellosis and campylobacteriosis, both intestinal diseases of mild, but occasionally life threatening, severity.

Methodology

Foodborne disease costs can be classified into 3 categories: individual, industry and public (Table 1). The individual's costs associated with illness and death include medical resources used, loss of wages or productivity during sickness, reduction of leisure time choices during the illness and recovery, and pain and suffering. The costs to the industry or firm found responsible for the outbreak may include the value of product recalled, reduction in future demand for the product due to reputation damage, plant cleanup, and liability awards[2]. Public costs include investigation, surveillance and possibly part of the cleanup expense.

Traditionally, only the easily monetizable, direct costs have been estimated, namely the medical costs and wages (or productivity) lost during an illness. However, averting behavior costs (behavior designed to avoid or reduce the risk of illness) can be a significant cost item and, in fact, may swamp the traditional medical and productivity costs. A recent Resources for the Future study of the contamination of a water supply addresses the willingness of the public to pursue a variety of measures to avoid illness--boiling water, travelling to another community to obtain water, and purchasing bottled water (Harrington, et. al., 1985).

Table 1

Social Costs of Foodborne Illness

Costs to individuals
Medical costs
Income or productivity loss
Pain and suffering
Leisure time cost
Averting behavior costs
Risk aversion costs
Travel cost
Child care cost
Industry costs
Product recall
Plant closings and cleanup cost
Product liability costs
Reduced product demand
Public health surveillance costs
Costs of investigating outbreak
Costs of maintaining disease surveillance
Cleanup costs

Averting behavior by the public can result in diet and consumption expenditure changes that affect sales and revenues of the involved industry. An opinion survey by the National Pork Producers Council found that 40 percent of the surveyed consumers claimed that they had reduced their consumption of pork because of health concerns about salt and 17 percent claimed decreased poultry consumption because of disease concerns (Weise Research Associates, 1984). Of course, opinion surveys by themselves don't provide empirical evidence of actual reductions in consumption and impacts upon industry revenues.

Finally, the costs to firms in the industry and the public are typically excluded from cost estimates, although Ewen Todd has found they are often significant components of foodborne outbreak costs (Todd, February 1985 and July 1985).

Evaluation of Costs of Illness

The three components of the cost of illness estimate are the number of persons affected, the severity of the illness, and the costs associated with that severity:

- o Estimated Disease Incidence. All diseases are typically underreported. By looking at outbreaks of foodborne illness, epidemiologists have found that only 1 in 75 or 1 in 100 salmonellosis cases are reported (Smith and Blaser, 1985). Rather than just rely on reported cases, estimates of the total U.S. incidence are used.

- o Severity of Illness. The range of illnesses caused by the diseases considered here vary from essentially unnoticeable to life-threatening. Among other things, the impact depends on the number of microbes ingested and on the efficiency of the individual's immune system to fight off the diseases. However, fatalities can occur in relatively normal human adult hosts (Smith and Blaser, 1985). Costs have been estimated for three disease severity levels--mild, moderate, and deadly.

- o Cost of Illness. Secondary data sources (updated to 1985 prices) are used for the cost estimates. Often these costs are derived from surveys of people involved in an outbreak of foodborne disease. Generally, they are confined to medical costs and wage losses and are underestimates.

Evaluation of Death

Traditionally, deaths have been evaluated traditionally by the human capital method which measures the individual's contribution to productive output. The income stream that would have been produced by the individual is collapsed into a present value for that production at today's prices.

$$\text{Human Capital Method} = \sum_{t}^{T} \frac{L_t + H_t}{(1 + i)^t}, \text{ where}$$

- T = remaining lifetime
- t = a particular year
- L_t = labor income in year t
- H_t = value of nonmarket time spent on homemaking services
- i = social discount rate; opportunity cost of society investing in life-saving programs

The human capital method only places a value on what the individual produces for society.

From the perspective of the individual and consumer demand theory, a life ought to be valued by what the individual is willing to pay to avoid a particular risk of death. The individual's non-labor sources of income are included as resources to pay for risk reduction along with wages. Even more important are the nonmarket activities that may be of more value to the individual than his/her income loss. These include pain and suffering, loss of leisure time, and aversion to risk.

$$\text{Revealed Preference Willingness to Pay Method} = \sum_{t}^{T} \frac{B_t}{(1+\rho)^t} - a, \text{ where}$$

T = remaining lifetime
t = a particular year
B_t = benefits of living = $L_t + NL_t + NM_t + P_t$, where L_t = labor income,
NL_t = nonlabor income, NM_t = nonmarket activities and leisure,
P_t = premium for pain and suffering
ρ = individual rate of time preference
a = risk aversion factor

Historically, the range of value of life of estimates resulting from the willingness to pay method has been large (Landefeld and Seskin, 1982).

A hybrid approach attempts to bridge the gap between the two methodologies (Landefeld and Seskin, 1982). The adjusted willingness to pay/human capital approach includes only measurable economic losses associated with death. It is based on after-tax income from labor and nonlabor sources, discounts at the individual's rate of return after taxes, and includes risk aversion shown by investment in life insurance, security systems, etc.

$$\text{Adjusted willingness to Pay/Human Capital Method} = \sum_{t}^{T} \frac{Y_t}{(1+r)^t} - a, \text{ where}$$

T = remaining lifetime
t = a particular year
Y_t = after-tax income = $L_t + NL_t$, where L_t = labor income, NL_t = nonlabor income
r = individual's opportunity cost of investing in risk-reducing activities
a = risk aversion factor

Perhaps most important, data exist for estimating with this methodology.

Costs of Salmonellosis and Campylobacteriosis

Two of the most prevalent foodborne diseases in the United States are salmonellosis and campylobacteriosis which annually cause intestinal disorders in an estimated 2 million and 2.1 million persons, respectively (Holmberg, 1985). The most extensive data on costs of illness come from the Centers for Disease Control (CDC) survey of a 1976 salmonellosis outbreak in Colorado (Cohen, et. al., 1978). These data are used to estimate the medical costs, productivity losses, and miscellaneous costs for mild cases ($230) and moderate cases ($1,290, (Table 2.. The incidence of moderate salmonellosis cases is conservatively assumed to be those 40,000 salmonellosis cases reported annually to CDC (Holmberg, 1985). The mild cases are the remainder, or 2.0 million estimated salmonellosis cases minus the 40,000 reported cases which equals 1,960,000 mild cases.

Table 2

Annual Cost Estimates for Salmoneliosis, United States (1985 prices).

Item	Number of cases	Cost per case	Total cost	
			low	high
	thousand	dollars	million $	
Moderate severity	40	1,290	51.6	
Mild severity	1,960	230	450.8	
			502.4	
Loss of life	2	85,800[a]	171.6	
		351,500[b]		703
Total cost			673	1,205

[a] Based on human capital value method.
[b] Based on the adjusted willingness-to-pay/human capital method.

For campylobacteriosis, the same costs are used because of the similarity of the course of the diseases (Seattle, 1985)(Table 3). The U.S. incidence for moderate cases is based on a study in Denver for the low estimate (Smith and Blaser, 1985) and the high estimate is based on a Seattle study (Seattle, 1984). Mild cases are the residual of the 2.1 million cases.

Death's are evaluated by using the CDC estimated death rate of one in a thousand for these diseases (Holmberg, 1985) and the actual age distribution of reported deaths due to salmoneilosis (Table 4). The low estimate, $85,800 is based on the present value of life with the human capital method. The high estimate, $351,500, is based on the adjusted willingness to pay/human capital method.

For the medical and productivity categories, U.S. costs for salmonellosis and campylobacterosis are estimated to range from $1.4 billion to $2.6 billion annually (Table 5). Note that while deaths are the largest component of the high estimate, productivity losses are the largest component of the low estimate. This changing position highlights the importance of what assumptions about incidence and severity are made in deriving the estimates and the methodology used to evaluate deaths. Also note the fewer number of cost categories estimated in Table 5 when compared to Table 1 implies that this estimate is quite conservative.

Table 3

Annual Cost Estimates for Campylobacteriosis, United States, (1985 prices).

Item	Number of cases	Cost per case	Total cost low	Total cost high
	#	$	million $	
Moderate[a] severity	57,340 or 168,025	1,290	74	217
Mild[b] severity	2,042,660 or 1,931,975	230	470	444
Loss of life	2,100	85,800[c]	180	
		351,500[d]		738
Total cost			724	1,399

[a] The low estimate is based on the incidence reported in the Denver area (Smith and Blaser, 1985) while the high estimate is based on the Seattle study.

[b] Total cases of campylobacteriosis are estimated at 2.1 million (Holmberg, 1985). The moderate cases are subtracted from this number to get the estimated mild cases.

[c] Based on human capital method.

[d] Based on the adjusted willingness-to-pay/human capital method.

Table 4

Value of Life for Salmonellosis Fatalities, 1985

Method	Age	Male			Female			Average Value[c]
		Deaths	Present Value	Total Value	Deaths	Present Value	Total Value	
		#	thousand $		#	thousand dollars		$
Human capital[a]	0-4	3	88	264	4	77	307	
	5-14	0	159	0	2	139	278	
	15-24	2	300	600	1	241	241	
	25-44	0	371	0	4	238	954	
	45-64	10	189	1,890	6	144	867	
	65+	20	14	271	27	41	1,107	
		35		3,025	44		3,754	85,800
Adjusted willingness to pay/human capital[b]	0-4	3	1208	3,624	4	836	3,102	
	5-14	0	1408	0	2	961	1,922	
	15-24	2	1655	3,309	1	1086	1,086	
	25-44	0	1432	0	4	866	3,462	
	45-64	10	548	5,480	6	410	2,459	
	65+	20	34	680	27	90	2,443	
		35		13,091	44		14,675	351,500

a Data from Dolan, et. al., 1980; Vital; Updated to March 1985 dollars.
b Data from Landefeld and Seskin, 1982; Vital; Updated to March 1985 dollars.
c Value calculated by dividing the total values for male and female by total number of deaths.

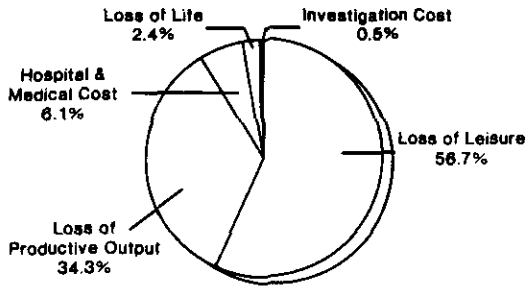

Figure 1

Economic Costs of Human Salmonellosis in Canada, 1982. Source: Curtin

Table 5

U.S. Annual Costs for Salmonellosis and Campylobacteriosis, 1985

Cost Categories	All Cases		
	low		high
	million dollars		
Medical Costs			
Mild Cases	0		0
Moderate Cases & Deaths	91	or	195
Lost Productivity			
Mild Cases a/	921	or	895
Moderate Cases a/	27	or	59
Deaths-Human Capital or Adjusted Willingness to Pay/Human Capital	352	or	1,441
Miscellaneous costs b/	6	or	13
TOTAL	$1.4 to 2.6 billion		

a This is a low estimate because it leaves out the value of homemaking.
b Miscellaneous costs include transportation, child care, other laboratory tests, finding new jobs.

Source: Data from Morrison and Roberts, 1985.

A Canadian study has also estimated costs of human salmonellosis and included two additional categories: loss of leisure and investigation (Curtin, 1984). The leisure cost is even more important than the loss of productive output, 56.7 percent of the total vs. 34.3 percent (Fig. 1). Some of their key assumptions on valuing leisure time are: there are 5 hours of leisure time available on a working day, 12.5 hours of leisure on a non-working day, and the value of leisure time is the average hourly wage rate (wage rates were calculated for five groups of people: working men 14-65, working women 14-65, working men and women 65+, non-working men and women 14-65, non-working men and women 65+). Productivity and leisure time losses are more important than hospitalization and other medical costs for salmonellosis and campylobacterosis in the Canadian study because the vast majority of the cases are mild and do not require hospitalization or even a visit to the doctor.

However, there are chronic medical syndromes that can follow both diseases that are difficult to quantify in their frequency and severity. For salmonellosis, rheumatoid arthritis is thought to be a common complication (Archer, 1985). Less common complications can affect the heart, the thyroid, the spleen, the pancreas, or even cause blood poisoning (Mossel, 1984). For campylobacteriosis, complications can include arthritis, blood poisoning, or inflammation of the heart, the colon, or the brain (ibid). If effects of these complications were accounted for, the estimates of the medical costs would be higher.

Food Safety Management Options

Pathogens in food are dependent upon three variables according to Mossel's and Stegeman's formula: $N_R = (N_O)(A^{-1})(\Delta)$

pathogens in food	=	pathogen contamination in raw product	pathogen reduction during processing	↑ or ↓ in pathogens after processing

Control options for reducing pathogens in food occur at all three points in this formula. Raw product contamination can be decreased by reducing or eradicating animal disease on the farm. One key factor in reducing salmonellosis on the farm is reducing or eliminating Salmonella in the feed. Improved sanitation practices in chicken slaughter and packaging plants could reduce the transfer of fecal matter and pathogens onto chicken meat. At the second stage, a processing technique such as canning or freezing can reduce pathogens. Irradiation has been proposed as a new method of reducing pathogens in raw pork, chicken and beef. Finally, there are procedures which could reduce salomonellosis and other pathogens in commercial and private kitchens. These include cooking to higher or more uniform internal temperatures and adopting better and more effective sanitation practices to keep raw chicken from contaminating cooked chicken or fruits and vegetables which are eaten raw.

Irradiation as a Control Option

Irradiation kills pathogens by interrupting their DNA. Work in the Netherlands by Mulder suggests that a 93 percent reduction in Salmonella-contaminated chicken carcasses can be achieved by an absorbed irradiation dose of 250 krads. A colleague, Rosanna Mentzer Morrison, estimated that irradiating fresh chicken packed at large and medium sized plants in the U.S. would cost $155 million a year. In comparison, the public health prevention benefits of irradiating U.S. fresh chicken to reduce Salmonella are estimated at $48 to 86 million. Chicken was identified as the source of 9.5 percent of the salmonellosis outbreaks reported in 1981 to CDC and 81 percent of the fresh chicken in the United States is packaged at these large and medium sized plants. ($48-86 million = $673-1,205 million x .93 x .095 x .81). In addition, irradiation kills 100% of the Campylobacter (Maxcy, 1983). Chicken was the cause of half of the campylobacteriosis cases in the Seattle study. The benefits for reducing campylobacteriosis by irradiating fresh chicken packaged at large and medium sized plants are $293 to $567 million annually ($724-1,399 million x .50 x .81).

The combined benefits of reducing salmonellosis and campylobacteriosis are $341 to $653 million. The estimated benefit/cost ratio of irradiating fresh chicken then ranges from 2.2 to 4.2, or $341/$155 to $653/$155.[3] Estimated net benefits range between $186 million ($341 minus $155) to $498 million ($653 minus $155).

CONCLUSION

The categories and methods which economists use to estimate the costs involved in foodborne illness have been briefly outlined. The example of estimating benefits and costs of irradiating fresh chicken to reduce salmonellosis illustrates many of the important considerations that must go into making these estimates. For example, the fact that irradiation also eliminates Campylobacter and other pathogens from chicken must be accounted for in benefit estimates.

ACKNOWLEDGEMENTS

I wish to thank Edna Loehman, Clark Burbee, Carol Kramer, and Doug McNiel for thoughtful comments on the paper.

NOTES

[1]Tanya Roberts is an economist with the National Economics Division, Economic Research Service, U.S. Department of Agriculture. The views expressed in this paper do not necessarily reflect those of the U.S. Department of Agriculture.

[2]In adding up costs, care must be taken to assure that product liability costs to firms are not already counted in the estimated pain and suffering costs to individuals.

[3]A number of other issues are overlooked in a benefit/cost analysis. The ratio depends upon what is included in the analysis. Are consumers concerned enough about salmonellosis and campylobacteriosis to be willing to cover the expense of irradiation? Are consumers convinced that irradiation is safe? Also, the estimates assume that these diseases are uniformly distributed throughout plants and not associated primarily with large or small packing plants.

The 250 krad absorbed dose level used in this analysis has not been approved by the Food and Drug Administration. However, Codex Alimentarius, an international cooperative body affiliated with the United Nations has approved doses of up to 700 krads for chicken.

REFERENCES

Archer, Douglas L., "Enteric Microorganisms in Rheumatoid Diseases: Causative Agents and Possible Mechanisms", J. of Food Protection, vol. 48, no. 6 (June 1985)538-45.

Bryan, Frank L., "Foodborne Diseases in United States Associated with Meat and Poultry", J. of Food Protection, vol. 43, no. 2 (1980) 140-50.

Centers for Disease Control. Foodborne Disease Outbreaks, Annual Summary 1981. Issued June 1983. HHS publication No. (CDC) 83-8185.

Cohen, M.L., R.E. Fontaine, R.A. Pollard, S.D. Von Allmen, T.M. Vern and E.J. Gangarosa, "An Assessment of Patient-Related Economic Costs in an Outbreak of Salmonellosis," The New England J. of Medicine, vol. 299 (1978)459-60.

"Consumers Eat Out More in 1985," NRA News Vol. 5, No. 7(Aug. 1985) 36-39.

Curtin, Leo. "Economic Study of Salmonella Poisoning and Control Measures in Canada." Agriculture Canada Working Paper 11/84, August 1984.

Dolan, T.J., T.A. Hodgson, and L.M. Wun, "Present Values of Expected Lifetime Earnings and Housekeeping Services, 1977", Division of Analysis, National Center for Health Statistics, HHS (1980).

Field, R. A., F. C. Smith, D. D. Dean, G. M. Thomas and W. A. Kotula, "Sources of Variation in Bacteriological Condition of Ground Beef," J. of Food Protection, vol.40, no. 6 (June, 1977) 385-388.

Harrington, Winston, Alan J. Krupnick, and Walter O. Spofford, Jr. "The Benefits of Preventing an Outbreak of Giardiasis Due to Drinking Water Contamination," ELI/EPA Seminar, Washington, D.C. 18 July 1985.

Holmberg, Scott D., Enteric Diseases Branch, Divisions of Bacterial Disease, Centers for Disease Control, HHS, letter to Tanya Roberts, April 24, 1985.

Holmberg, Scott D., Michael T. Osterholdm Kenneth A. Senger, and Mitchell L. Cohen, "Drug Resistant Salmonella from Animals Fed Antimicrobials", New England J. of Med. vol. 311, no. 10 (Sept. 6, 1984) 617-622.

Kamplemacher, E.H., "Benefits of Radiation Processing to Public Health", paper presented at IAEA/FAO/WHO Symposium, San Diego, Fall, 1984.

Landefeld, J. Steven, and Eugene P. Seskin, "The Economic Value of Life: Linking Theory to Practice", Amer. J. of Public Health, vol. 72, no. 6 (June 1982) 555-565.

Kraus, Nancy Nighswonger, "Taxonomic Analysis of Perceived Risk: Modeling the Perceptions of Individual and Representing Local Hazard Sets," Ph.D. Dissertation, University of Pittsburgh, 1985.

Maxcy, R. B. "Significance of Residual Organisms in Foods After Sub-sterilizing Doses of Gamma Radiation: A Review," J. of Food Safety, 5 (1983) 203-211.

Morrison, Rosanna Mentzer and Tanya Roberts, "Food Irradiation: Technical, Regulatory, and Economic Considerations", Draft for the Office of Technology Assessment, May 1985.

Mossel, D.A.A., "Intervention as the Rational Approach to Control Diseases of Microbial Etiology Transmitted by Foods," J. of Food Safety, vol. 6 (1984) 89-104.

Mossel, D.A.A. and E.F. Drion, "Risk Analysis: Its Application to the Protection of the Consumer Against Food-transmitted Disease of Microbial Aetiology", Antonie van Leeuwenhoek, 45(1979) 321-3.

Mossel, D.A.A. and H. Stegeman "Irradiation: An Effective Mode of Processing Food for Safety," Paper presented at the IAEA/FAO International Symposium on Food Irradiation Processing, Washington, D.C. (March 4-8, 1985).

Mulder, R. W. A. W. Salmonella Radicidation of Poultry Carcasses, PhD. Dissertation, Agricultural University, Wageningen, The Netherlands, 1982.

Schwabe, Calvin W. Veterinary Medicine and Human Health. 3rd ed. Williams Wilkins: Baltimore, 1985.

Seattle - King County Dept. of Public Health, Communicable Disease Control Section. "Surveillance of the Flow of Salmonella and Campylobacter in a Community", USHHS Contract No. 223-81-7041. August, 1984.

Smith, Gordon S. and Martin J. Blaser, "Fatalities Associated with Campylobacter jejuni Infections", J. of the Amer. Med. Assoc., vol. 253, no. 19 (May 17, 1985) 2873-5.

Snoyenbos, H.G. (ed), Proceedings of the International Symposium on Salmonella (Amer. Assoc. of Avian Pathologists: Kenneth Square, PA) 1985.

Todd, Ewen C. D., "Economic Loss from Foodborne Disease and Non-illness Related Recalls Because Of Mishandling by Food Processors", J. of Food Protection, 48(July 1985)621-633.

Todd, Ewen C. D., "Economic Loss from Foodborne Disease Outbreaks Associated with Foodservice Establishments", J. of Food Protection, 48 (Feb. 1985) 169-180.

Vital Statistics of the United States, 1979: Volume II - Mortality, Part A. U.S. Dept. of Health and Human Services, Public Health Service, National Center for Health Statistics (Hyatsville, Maryland), 1984.

Wiese Research Associates, Inc., "Consumer Reaction to the Irradiation Concept", prepared for the U.S. Department of Energy and the National Pork Producers Council, March 1984.

THE COCHRAN-ARMITAGE TEST

FOR TRENDS OR THRESHOLDS IN PROPORTIONS

S. Stanley Young

Statistical and Mathematical Services
Lilly Research Laboratories

ABSTRACT

The Cochran-Armitage test for trends or thresholds in proportions. Young, S. S. (1985) Society for Risk Analysis, 1985 Annual Meeting. the Cochran-Armitage (C-A) test (1954, 1955) is widely used as a test for linear trends in proportions in the analysis of long term rodent studies. Implicit in the use of this test is the assumption that the dose response pattern is known. Although in practice the dose response pattern is often assumed linear on a log dose scale, this test can be used to test for nonlinear dose response patterns. The effect of the choice of different dose response patterns is examined using hypothetical and actual examples of tumor data in rodents. The choice of a particular dose response pattern can greatly influence the p-value from the C-A test. As an alternative to the usual C-A test, a sequential testing procedure, similar to the Williams-t (Williams, 1971, 1972), is suggested. Under the assumption of a threshold model, this procedure gives improved testing and estimation and leads to better inferences.

KEY WORDS: Cochran-Armitage test, Bioassay, Thresholds, Trends, DDT, TDE, DDE

INTRODUCTION

It is common practice in toxicology experiments to give increasing doses of a compound to groups of animals and to compare their response to the response of animals in an untreated group. Experimenters are interested in the progressive nature of any induced toxic phenomenon. They are also interested in does that are without effect (or with effects that are so small as to be considered unimportant). Typically, many responses are measured in each animal. Progressivity of a dose response is often used to help assess whether an observed group response is treatment related or random. For example, a statistically significant response at a low dose, unconfirmed at higher doses, would usually be considered indicative of a false positive result.

It is often stated, see for example the Report of the NTP Ad Hoc Panel on Chemical Carcinogenesis Testing and Evaluation (1984), that a trend test is more powerful than multiple tests of each treated group versus the control; trend tests gain this advantage since numerous tests

are replaced with a single test. However, there are some disadvantages to trend tests. They usually do not address the question of a "no-effect" dose. Also there is an implicit assumption that the shape or the form of the dose response curve is known.

The purpose of this paper is to examine the Cochran-Armitage (C-A) test (1954, 1955) for trends in proportions with regard to its implicit assumption of a known dose response pattern. A threshold test alternative to the usual C-A test is given. Numerous examples are included for three reasons. First, p-values for the same data vary greatly depending on the form of the test. Second, a number of data sets offer empirical evidence for thresholds. Third, computations are exemplified.

METHODS

The Cochran-Armitage trend test can be computed as the square of Z, given below:

$$Z = \frac{\Sigma\, di(xi - p*ni)}{(p*q*(\Sigma ni(di - \bar{d})**2))\ **0.5}$$

where

i = index for treatment group,
0 control, 1 low dose, etc.
xi = the number of tumor bearing animals in group i
ni = the number of animals in group i,
di = selected to match the expected dose response,
N = Σni
p = Σni / N
q = 1 - p
d = Σni*di / N.

Z is approximately normally distributed with a mean of 0 and a variance of 1. Negative Z-values indicate a decreasing trend in the proportions; positive Z-values indicate an increasing trend. The di are selected by the experimenter to match the expected dose response curve. They are often chosen as 0,1,2 etc to match a dose response that is presumed linear on a log scale.

EXAMPLES: HYPOTHETICAL DATA

To examine the influence of the selection of di on the C-A test, several examples were constructed. Three examples are given in Table 1. These three examples were constructed to have doses placed in three different ways on a sigmoid dose response curve. See Figure 1. Example (a) has the low dose on the initial flat portion of the sigmoid curve. Example (b) has the low dose in the "linear phase" of the dose response and Example (c) has both the two treated groups on the upper flat portion of the dose response curve. For each example, the C-A z-values and corresponding p-values are given for three different sets of di. The first set of di, (0,0,1), is constructed to match doses that are located at control, lower plateau, and higher plateau. The second set of di, (0,1,2), assumes a linear response. The third set of di, (0,1,1), models control and both treated groups on the upper plateau of a sigmoid curve. It is obvious from Table 1 that the z-value and corresponding p-value are greatly influenced by the choice of di. The more closely the choice of di corresponds to the observed dose response pattern, the larger the z-value and smaller the p-value.

Table 1

Results of Cochran-Armitage Test for Constructed Examples of Tumor Data for Different Sets of Di, See Figure 1

Example	Dose Group 0	1	2	C-A di	Z Value	P Value
(a)	0/50	0/50	10/50	001	4.629	0.000002
				012	4.009	0.00003
				011	2.315	0.010
(b)	0/50	5/50	10/50	001	2.887	0.002
				012	3.333	0.0004
				011	2.887	0.002
(c)	0/50	10/50	10/50	001	1.698	0.045
				012	2.942	0.002
				011	3.397	0.0003

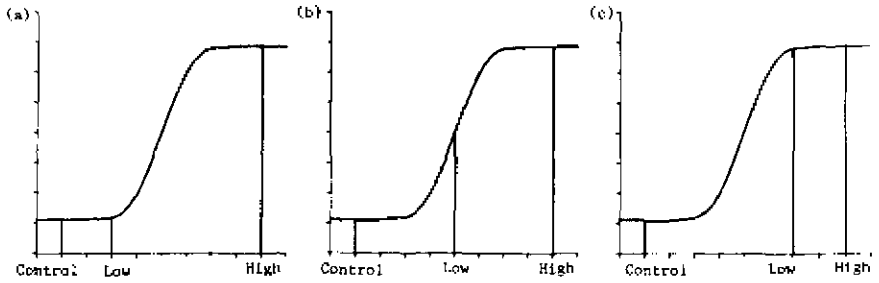

Figure 1

Dose Placement on a Signoid Response Curve

EXAMPLES: REAL DATA

Two examples are now given of mouse experiments conducted within Lilly Research Laboratories. For Study 1, Table 2(a) gives the survival at one and two years and the incidence of animals with proliferative hepatic lesions: hyperplasia, adenoma, or carcinoma. Survival is given to indicate that early deaths in the treated groups did not produce the decline in hepatic lesions. Table 2(b) gives the results of C-A tests for several sets of di. For both males and females a simple test of control versus treated (0,1,1,1) is unimpressive. A test assuming the doses are linearly spaced on the dose response curve (0,1,2,3) is significant for both males and females and indicate a decrease with increasing dose. A set of di that matches the apparent threshold nature of the dose response

for each sex gives an even smaller p-value. The latter di were selected by looking at the observed dose response and so the choice of di is subject to post hoc criticism.

Table 2

Lilly Study 1 in Male and Female Mice

(a) Survival at one and two years and the incidence of proliferative hepatic lesions.

Dose Group	Males			Females		
	Survival 1 Year	Survival 2 Years	Liver Lesions	Survival 1 Year	Survival 2 Years	Liver Lesions
0	59/60	53/60	19/60	59/60	47/60	5/60
1	60/61#	52/61#	22/61#	59/59#	48/59#	5/59#
2	60/60	48/60	20/60	60/60	51/60	2/60
3	59/60	51/60	9/59*	58/60	47/60	0/60
4	60/60	54/60	10/60	60/60	57/60	1/60

\# Animal missexed.
* One animal lost for evaluation.

(b) Cochran-Armitage trend test using different sets of di

Sex	Dose Group 0	1	2	3	4	C-A di	Z Value	P Value
Males	19/60	22/61	20/60	9/59	10/60			
(%)	31.7	36.1	33.3	15.3	16.7	01234	-2.81	0.997
						01111	-0.98	0.836
						00011	-3.40	0.9997
Females	5/60	5/59	2/60	0/60	1/60			
(%)	8.3	8.5	3.3	0.0	1.7	01234	-2.61	0.995
						01111	-1.69	0.955
						00111	-2.80	0.997

For Study 2, Table 3(a) gives the survival at 12 and 22 months and the incidence of animals with proliferative hepatic lesions. Again, survival is given to indicate that animals in the different groups were equally at risk so that increased survival did not produce the increase in hepatic lesions. Also note that the increase in liver lesions was not associated with a decrease in survival. Table 3(b) gives several C-A tests for males and females. For both males and females a simple test of control versus treated is not significant. Tests assuming the doses are linearly spaced on the dose response curve are significant for both sexes and indicate an increase with increasing dose. A set of di that matches the apparent threshold nature of the dose response gives an even smaller p-value. This set of di was selected by looking at the observed dose response and is again subject to post hoc criticism.

Table 3

Lilly Study 2 in Male and Female Mice

(a) Survival at 12 and 22 months and the incidence of proliferative hepatic lesions.

	Males			Females		
Dose Group	Survival 12 Months	Survival 22 Months	Hepatic Lesions	Survival 12 Months	Survival 22 Months	Hepatic Lesions
0	108/120	44/120	10/120	107/120	38/120	0/120
1	69/ 80	31/ 80	6/ 80	72/ 80	21/ 80	0/ 80
2	64/ 80	17/ 80	7/ 80	71/ 80	17/ 80	0/ 80
3	67/ 80	35/ 80	18/ 80	68/ 80	27/ 80	2/ 80

(b) Cochran-Armitage trend test using different sets of di

Sex	Dose Group 0	1	2	3	C-A di	Z Value	P Value
Males	10/120	6/80	7/80	18/80			
(%)	8.3	7.5	8.8	22.5	0123	2.78	0.003
					0111	1.29	0.099
					0001	3.54	0.0002
Females	0/120	0/80	0/80	2/80			
(%)	0.0	0.0	0.0	2.5	0123	2.05	0.020
					0111	1.00	0.159
					0001	2.65	0.004

Table 4(a) gives the incidence of urinary bladder neoplasia of female mice treated with 2AAF in animals that survived 18 or 24 months, Littlefield, et al. (1979). The incidence of neoplasia in the three lowest doses is essentially equal to that of control animals. A positive dose response appears to begin at 60 ppm. A formal test of threshold for this data set is given in Table 4(b) and will be described later.

A final example is a composite experiment on three structurally related compounds, DDT, DDE and TDE, see Figure 2. DDE and TDE are metabolites of DDT. These compounds, each at two dose levels, were run together in an experiment by the National Cancer Institute, Report 131 (1978). Table 5(a) gives the incidence of liver carcinoma for this experiment. Table 5(b) gives the results of several C-A tests. A test of control versus treated is significant. A test of linear trend has a substantially smaller p-value. A C-A test for a threshold followed by a linear trend followed by a plateau has an even smaller p-value. The decrease in p-values can be taken to indicate progressively better models of the dose response. A formal test of threshold for this data set is given in Table 5(c) and will be described next.

Table 4

NTP Study in Female BALB/c Mice Treated with 2AAF, Killed at 18 or 24 Months, Littlefield et al (1979).

(a) Urinary bladder neoplasia

Group	Dose (ppm)	Bladder Incidence	Neoplasia Percent
0	0	2/ 784	0.26
1	30	4/2473	0.16
2	35	3/1434	0.21
3	45	2/ 828	0.24
4	60	6/ 684	0.88
5	75	4/ 578	0.69
6	100	30/ 291	10.31
7	150	162/ 251	64.54

(b) Results of threshold test, see Figure 3.

Test	"Control" Incidence	"Control" Percent	"Treated" Incidence	"Treated" Percent	z Value	p Value
0 vs 1	2/ 784	0.26	4/2473	0.16	-0.531	.798
0,1 vs 2	6/3257	0.18	3/1434	0.21	0.180	.429
0-2 vs 3	9/4691	0.19	2/ 828	0.24	0.296	.384
0-3 vs 4	11/5519	0.20	6/ 684	0.88	3.199	.001
0-3 vs 4,5	11/5519	0.20	10/1262	0.79	3.421	.00031

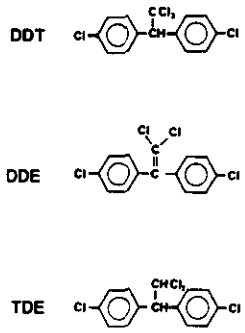

Figure 2

Structure of DDT, DDE and TDE

Table 5

NCI Studies on DDT, DDE and TDE in Male B6C3F1 Mice.

(a) Liver carcinoma

Group	Treatment	Liver Carcinoma Incidence	Liver Carcinoma Percent
0	Control	4/56	7.1
1	22 ppm DDT	1/49	2.0
2	44 ppm DDT	1/48	2.1
3	148 ppm DDE	7/41	17.1
4	261 ppm DDE	17/47	36.2
5	411 ppm TDE	12/44	27.3
6	822 ppm TDE	15/50	30.0

(b) Cochran-Armitage trend test using different sets of *di*

C-A di	Z Value	P Value
0111111	2.22	1.3×10^{-2}
0123456	5.25	7.6×10^{-8}
0001222	6.06	9.3×10^{-10}

(c) Results of threshold test, see Figure 3.

Test	"Control" Incidence	"Control" Percent	"Treated" Incidence	"Treated" Percent	Z Value	P Value
0 vs 1	4/ 56	7.1	1/49	2.0	-1.22	.889
0,1 vs 2	5/105	4.8	1/48	2.1	-0.79	.786
0-2 vs 3	8/153	5.9	7/41	17.1	2.99	.002

THRESHOLD TESTING METHOD

The C-A test can be modified to give a statistical testing procedure similar to that of Williams (1971, 1972). The pattern of the test is given in Figure 3. The approach is to first test the control vs the low dose. If the result is significant then there is evidence of a problem: the low dose is an effect level (EL). If the low dose does not differ significantly from the control, then the low dose and the control are combined and used to test the next dose. This procedure is continued until an effect dose is found or all the dose groups have been considered. Dose levels below the effect level are considered to be no-effect levels (NOEL).

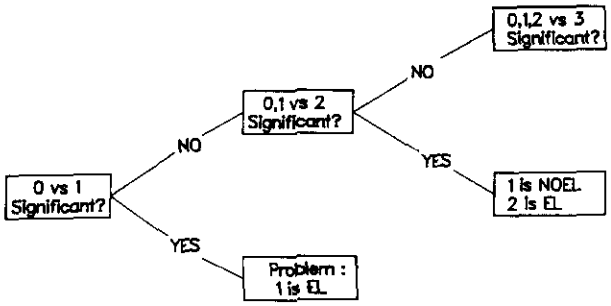

Figure 3

Schematic for the Application of a Williams Type Test to Binomial Data

Threshold testing is applied to the 2AAF and DDT composite data sets and the results are given in Tables 4(b) and 5(c). In both cases no effect levels are determined.

There are two practical aspects of the threshold method that deserve comment: control of experiment-wise error rate and allocation of animals to groups. First, it is common practice in any multiple comparison procedure to adjust the α-level of the statistical test so that a specified experiment-wise error rate is maintained. The greater the number of statistical tests, the more extreme the test statistic must be before significance is declared. The determination of critical values can be difficult. Two aspects of the threshold test can be used to simplify computation of critical values. The first aspect is a question of experimental intent. Often it is the intent to establish a no-effect level. Dose levels are set so that if the low dose is a no-effect level safety is considered assured. In this case, the control versus low dose comparison can be made at the nominal error rate as the experimental intent is satisfied if that test is not significant. The second aspect is that the construction of the test assures statistical independence of the step-wise statistical comparisons. Each sequential test can be formulated as a polynomial and the polynomials are orthogonal. For example, for a control and three treatment groups

	Group			
Comparison	0	1	2	3
0 vs 1	-1	1	0	0
0,1 vs 2	-1	-1	2	0
0-2 vs 3	-1	-1	-1	3.

The independence of the comparisons allows a simple determination of a per-comparison critical value that assures a specified experiment-wise error rate. Critical z-values for the threshold test are given in Table 6. The first row, only one treated group, can be used if establishment of one no-effect level is the only issue. The other rows give critical z-values if a specified experiment-wise error rate is to be maintained. It should be pointed out that if the premise of the test is correct -- threshold exist -- then adjustment for multiple comparisons does not make sense. Once a no-effect level is declared, safety is assured. Correction for multiple comparisons only makes sense if, even while using a threshold model, safety is not considered assured unless all tests are not significant. Anyone considering this strategy would be better advised to simply test control versus high (one test) at the experiment-wise error rate.

Table 6

Critical Per-Comparison Z-Values that Maintain a Specified Experiment-Wise Error Rate, One-Tail; for Two-Tailed Testing use 1/2 Experiment-Wise Error Rate.

Number of Treated Groups	Experiment-wise error rate			
	.050	.025	.010	.005
1	1.6449	1.9600	2.3263	2.5758
2	1.9545	2.2390	2.5749	2.8066
3	2.1212	2.3909	2.7119	2.9348
4	2.2340	2.4944	2.8059	3.0229
5	2.3187	2.5723	2.8769	3.0896
6	2.3862	2.6347	2.9338	3.1434

The second practical aspect of the threshold method is how to allocate animals to the various groups. There is much confusion in the literature about "optimal" allocation of animals, see for example Portier and Hoel (1983, 1983, 1984) and Krewski (1983). Much of this confusion can be attributed to an unclear statement of the objective of the experiment.

Once the goal of the experiment is determined, the allocation of animals is straightforward. With a threshold experiment there are two reasonable ways to allocate animals depending upon the goal of the experimenter. First, if the goal is to establish a no-effect level at a particular dose, then the animals should be divided equally between the control group and the dose level, see Table 7. Second, if the goal is to determine where the threshold begins, then there are three questions.

Table 7

Allocation of Animals to a Dose-Response Study as a Function of the Goal of the Experiment. N. is the Total Sample Size.

Goal	Allocation	Comments
Low dose effect	N/2, N/2, 0,0,etc	Only issue is safety of low dose.
High dose effect	N/2, 0,...,0, N/2	Only issue is toxicity of high dose
Where is threshold	3 Groups: N/4,N/4,N/2 4 Groups: N/8,N/8,N/4,N/2	Maximum power at top dose level
Determine shape	3 Groups: N/4,N/2,N/4 4 Groups: N/8,N/4,N/4,N/8	Maximum power for quadratic response
Control vs several groups	For k treated groups, control group replicated k times the trt group size.	Decreases false positives from aberrant controls

1. How many dose groups?
2. How to space the dose groups?
3. How many animals at each group?

The answer to the third question offers some guidance to the first two questions; unfortunately answers to questions 1 and 2 depend upon knowing the shape of the true dose response. If a control and three dose levels have been chosen, then the optimum allocation of animals is 1/8 at control, 1/8 at low dose, 1/4 at mid dose, and 1/2 at high dose. The reason for this allocation is that at each stage of the threshold test it is optimal to have an equal number of animals in the two groups being compared. It is well known that the optimal allocation in a two group comparison is equal allocation of animals. This result is somewhat disturbing: there is less power to detect a threshold at a particular level than at the next. The answer to this dilemma is that maximum power to detect an effect at a particular level is a different goal and as was mentioned before the best allocation of animals in that situation is 1/2 at control and 1/2 at that level. Note that if the threshold occurs at the highest dose, then final test of the threshold method will have 1/2 of the animals combined and tested against the highest dose. This test is identical to optimum allocation for the question "Is there any effect?" That test would use two groups: a control and a group at the highest possible dose. Table 7 gives suggested optimum allocations of animals for various experimental situations.

The above discussion offers some guidance to the questions of how many dose groups and how to space them. Fewer dose groups will have more animals in each dose group (assuming a fixed total number of animals) so each sequential test will be a powerful as possible. The top dose should be as high as is reasonable so that maximum response can be observed. The

low dose should be as low as possible commensurate with human exposure, safety factors etc. Intermediate doses should be spaced as far apart as possible. It is often argued that the dose scale be in log units. It is doubtful that increasing the number of dose groups beyond three adds much beyond insurance against misplacement of the dose levels.

DISCUSSION

The Cochran-Armitage trend test requires that the dose response pattern, i.e. the di, be specified by the experimenter. The true dose response pattern must be known for the test to be optimal. It is clear that the choice of di can have a large effect on the test statistic. Unfortunately, the optimal set of di is seldom known until the results are scrutinized. If the dose response is sigmoid, the actual spacing of doses might fall in any of a number of ways. If the actual doses do not fall in a linear portion of the response curve, then the C-A test with linear di will not give a z-value that is as large as that given by a test with a set of di that matches the real dose response.

Several data sets are given where the dose response is nonlinear. In Lilly Study 1, the control and low doses are similar in response and the higher doses give a decrease in response. In Lilly Study 2, the control and two low doses are similar and in the high dose there appears to be an increase in response. Both studies along with the 2AAF and DDT-DDE-TDE experiments can be taken as empirical evidence for the existence of thresholds. The usual application of the C-A test which assumes a linear dose response can be misleading when there is a threshold or nonlinear dose response present.

Although it has been taken on faith by some that thresholds do not exist, this faith is far from universal and a number of recent papers support the existence of thresholds; see for example the book CANCER AND THE ENVIRONMENT, Possible Mechanisms of Thresholds for Carcinogens and other Toxic Substances, Cimino, ed. (1983). Also in an extensive review of the potential carcinogenic risk from formaldehyde, Squire et al. (1984) question the presumption that thresholds can not exist. Matters of faith are not testable, hence are outside the usual scope of science; the proposed threshold test does bring thresholds into the realm of testability.

The DDT-DDE-TDE experiment is included for several reasons. First it is illustrative of how the p-values of the C-A test can change as a function of the di. The p-value can change by eight orders of magnitude over rather reasonable sets of di, Table 5(b). If it is presumed that the compounds are so similar structurally and metabolically that they will be roughly identical in the induction of hepatic tumors, then this data set is illustrative of a sigmoid dose response pattern with a threshold and a plateau. The premise that all three compounds are roughly equal in their ability to induce hepatic tumors is not unreasonable as the nonresponding compound, DDT, has been shown to induce hepatic tumors in other studies, see for example Wahrendorf, 1983. There are few long term studies with enough dose groups to span the entire range of a sigmoid dose response curve. The low doses of DDT, 22 and 44 ppm, appear somewhat protective. In studies in trout, Hendricks et al. (1977), and Sheldon et al. (1984), show that PCB, another enzyme inducer, is protective against induced hepatic tumors.

There are several methods in the literature for addressing the testing of thresholds and the fitting of dose response functions with

plateaus. Williams (1971, 1972) gives a method of testing for a threshold that we have modified here for proportions. Anderson and Nelson (1975) give a method of piece-wise linear fitting of a response function so that thresholds and plateaus can be modeled. Both methods are for continuous data, but are easily extended to binomial data. The methods of Anderson and Nelson are aimed at curve fitting whereas the methods of Williams are aimed at hypothesis testing. Daly (1962) gives a method for testing for trends in a contingency table which is the mechanical equivalent of the method given here, but he combines his sequential tests into one composite test. Tukey et al. (1985) have proposed a sequential testing scheme to find a "no-statistical-significance-of-trend" level. Their scheme is for quantitative variables and although it could be extended to binomial data, they do not recommend extension, Ciminera (personal communication). Poon (1980) compares several methods for binomial data.

There are a number of reasonable objectives in the assessment of a dose response. The particular objectives and the methods to be employed should be stated before the conduct of the study. Care is needed in the application of the Cochran-Armitage test; the results are dependent on the particular dose response pattern assumed, i.e. the set of di. If a crude indicator of treatment effect is all that is desired, then the Cochran-Armitage test with a linear set of di, i.e. 0,1,2, etc, could suffice although a control versus high appears preferable. This test has essentially equal power and makes no assumption on the shape of the dose response. If thresholds are considered possible and no effect levels are of interest, then the threshold test given here appears reasonable. It allows for the determination of "no effect" levels and still has good power for the detection of treatment effects. If curve fitting is of interest, then a method similar to that of Anderson and Nelson appears reasonable. Neither method requires that the complete shape of the dose be known.

There is a final important caveat: All of the methods discussed herein assume that the single response of interest has been identified before the conduct of the study. In the case of long-term rodent carcinogenic studies, assessment of the data often constitutes a survey of several hundred possible tumors rather than an explicit testing scheme. It is not known how curve fitting and hypothesis testing that were designed for a single predetermined response behave when they are applied to several post hoc selected response variables. It is probable that hypothesis testing, by any usual method, is greatly upset and that false positives will result, Muller et al. (1984).

REFERENCES

Anderson, R. L. and L. A. Nelson. "A family of models involving intersecting straight lines and concomitant experimental designs useful in evaluating response to fertilizer nutrients" Biometrics 31:303-318 (1975).

Anolymous. "Bioassay of DDT, TDE, and p,p'-DDE for possible carcinogenicity" NCI Technical Report, No. 131. (1978).

Armitage, P. "Tests for linear trends in proportions and frequencies" Biometrics 11:375-386 (1955).

Board of Scientific Counselors, NTP. Report of the ad hoc panel on chemical carcinogenesis testing and evaluation. U.S. Dept. of Health and Human Services. (1984).

Cimino, J.A. Cancer and the Environment, Possible Mechanisms of Thresholds for Carcinogens and other Toxic Substances. Mary Ann Liebert. New York. (1983).

Cochran, W.G. "Some methods of strengthening the common Chi-Square test" Biometrics 10:417-451 (1954).

Daly, C. "A simple test for trends in a contingency table" Biometrics 18:114-119 (1962).

Hendricks, J.D., T.P. Putman, D.D. Bills and R.O. Sinnhuber. "Inhibitory effect of a polychlorinated biphenyl (Aroclor 1254) on aflatoxin B1 carcinogenesis in rainbow trout (Salmo gairdneri)" JNCI 59:1545-1551 (1977).

Krewski, D., J. Kovar, and M. Bickis. "Optimal experimental designs for low dose extrapolation" In Topics in Applies Statistics edited by T.W. Dwived: Marcell Dekker, New York (1983).

Littlefield, N.A., J.H. Farmer, D.W. Gaylor and W.G. Sheldon. "Effects of dose and time in a long-term low-dose carcinogenic study" Journal of Environmental Pathology and Toxicology 3:17-34 (1979).

Muller, K.E., C.N. Barton, and V.A. Benignus. "Recommendations for appropriate statistical practice in toxicology experiments" Neuro. Toxicology 5:113-126 (1984).

Poon, A.H. "A Monte Carlo study of the power of some k-sample tests for ordered binomial alternatives" J. Stat. Comp. and Sim. 11:251-259 (1980).

Portier, C. and D. Hoel. "Low-dose-rate extrapolation using the multistage model" Biometrics 39:897-906 (1983).

Portier, C. and D. Hoel. "Optimal design of the chronic animal bioassay" J. Tox. Env. Health 12:1-12 (1983).

Portier, C. and D. Hoel. "Design of animal carcinogenicity studies for goodness-of-fit of multistage models" Fundam. Appl. Toxic 4:949-959 (1984).

Sheldon D.W., J.D. Hendricks, R.A. Coulombe, and G.S. Bailey. "Effect of dose on the inhibition of carcinogenesis/mutagenesis by Aroclor 1254 in rainbow trout fed aflatoxin B1" J. Tox. and Envir. Health 13:649-657 (1984).

Shirley, Eryl A. C. "The comparison of treatment with control group means in toxicology studies" Applied Statistics 28:144-151 (1979).

Squire, R.A. and Cameron, L.L. "An analysis of potential carcinogenic risk from formaldehyde" Regulatory Toxicology and Pharmacology 4:107-129 (1984).

Tukey, J.W., Ciminera, J.L. and Heyse, J.F. "Testing the statistical certainty of a response to increasing doses of a response to increasing doses of a drug" Biometrics 41:295-301 (1985).

Wahrendorf, J. "Simulatneous analysis of different tumor types in a long term carcinogenicity study with scheduled sacrifices" JNCI 70:915-921 (1983).

Williams, D.A. "A test for differences between treatment means when several dose levels are compared with a zero dose control" Biometrics 27:103-117 (1971).

Williams, D.A. "The comparison of several dose levels with a zero dose control" Biometrics 28:519-531 (1972).

DEVELOPMENT OF THE NON-DIMENSIONAL METHOD OF RANKING RISKS

John C. Chicken* and Michael R. Hayns**

J. C. Consultancy Ltd and*
John C. Chicken Consultancy GmbH
United Kingdom Atomic Energy Authority**

ABSTRACT

This paper presents a critical review of the progress made in the development of the risk ranking technique. The aim of the development of the technique has been to produce a method of making a comprehensive assessment that takes into account the technical, economic and socio-political factors involved in determining the acceptability of a risk.

The paper examines: the data available for ranking acceptability, the ranking of the Moss Morran and Eemshaven projects, the efficacy of the technique in its present state of development and the future uses for the technique.

KEY WORDS: Risk Ranking, Overall Acceptability Assessment

1. INTRODUCTION

This paper examines the practical problems associated with applying the method of ranking the acceptability of risks which was described in a paper presented at the 1984 Annual Meeting.[1] The incentive for developing the technique is to produce a unified, non-emotive, non-dimensional, easily understood way of describing the acceptability of risks. In the following the application of the technique to ranking two projects is examined.

The factors that have to be considered are grouped under the headings of technical, economic and socio-political, the rank of each project being determined by integrating the acceptability scores of each group of factors. The scores allocated are an assessment of the acceptability and uncertainty of each factor. Table 1 shows how the scores are related to the ranking and the type of control action likely to be associated with each level of ranking.

The problems involved in ranking acceptability are examined in three steps which are: the data required, ranking the acceptability of the British Moss Morran and the Dutch Rijnmond projects and assessing the efficacy of the technique. Finally, the general uses and future development of the technique are examined.

2. DATA FOR RANKING

Each group of data has quite different characteristics.

Table 1

Definition of Rank Acceptability and Control Action

RISK RANK	ACCEPTABLITY OF PROPOSAL	TOTAL SCORE ALLOWABLE FOR EACH RANK	MAXIMUM ALLOWABLE SCORE FOR EACH RANKING FACTOR (TECHNICAL, ECONOMIC, AND SOCIO-POLITICAL	CONTROL ACTION REQUIRED
1	UNLIKELY TO BE ACCEPTABLE	6 - 12	4	UNLIKELY ANY POSSIBLE
2	ONLY ACCEPTABLE IF RISK CAN BE REDUCED	4 - 6	3	ADMINIS-TRATIVE AND ENGINEERING
3	YES SUBJECT TO CERTAIN ACTION	>2 - 4	2	ENGINEERING
4	YES WITHOUT RESTRICTION	0 - 2	1	NONE

The technical group of data is essential to the evaluation of the acceptability of a risk. In one definitive technical study of risk, (NUREG 1050), a clear indication of the uncertainty associated with such data is given by statements that: estimates of the frequency of nuclear reactor core-melt may differ by two orders of magnitude and that estimates of the likelihood of operator error may deviate by an order of magnitude.[2] The magnitude of these uncertainties led to the following interesting conclusions about the usefulness of quantitative risk assessment.[2]

"Probabilistic Risk Assessment results are useful, provided that more weight is given to the qualitative and relative insights regarding design and operations, rather than the precise absolute magnitude of the numbers generated.

It must be remembered that most of the uncertainties associated with an issue are inherent to the issue itself rather than artifacts of the Probabilistic Risk Assessment Analysis. The Probabilistic Risk Assessment does tend to identify and highlight these uncertainties, however.

Probabilistic Risk Assessment results have useful application in the prioritization of regulatory activities, development of generic regulatory positions on potential safety issues and the assessment of plant-specific issues. The degree of usefulness depends on the regulatory application as well as the nature of the specific issue.

The basic attributes of a Probabilistic Risk Assessment are not highly compatible with a safety-goal structure that would require strict numerical compliance on the basis of the quantitative best estimates of Probabilistic Risk Assessment."

These conclusions draw attention to two very important findings, which are the significance of a risk will not be known accurately and the allowance for uncertainty is not generally adequate so there is a need to develop a comprehensive way of allowing for uncertainty.[2]

Further indication of the range of uncertainty that can be associated with estimates of risk is given by the results of a study of decision making and risk analysis in relation to the siting of liquefied energy gas facilities.[3] The study argued that for most risk estimates the range of uncertainty is at least 10^2. To emphasise that 10^2 may be the minimum range of uncertainty attention is drawn to two figures from the study, which are estimates for the probability of an internal system failure, one is 3.2×10^{-3} and the other is 1.0×10^{-11}. Such a wide difference shows dramatically how great the uncertainty in risk estimates can be.

The conclusion about the technical data that is justified is that the data likely to be available will contain an element of uncertainty. It is in dealing with uncertainty in risk assessment that the ranking technique can be particularly helpful.

We are all conscious of the economic significance of variations in our expenditure. With major projects the variation in cost can be many millions of pounds. In NUREG 1050 it is shown that the total financial risk for a pressurized water reactor ice-condenser plant can be between 5×10^5 and 8×10^7 dollars per plant lifetime.[2] The problem was examined further in a study which the US Nuclear Regulatory Commission had made of the socio-economic consequences of nuclear reactor accidents.[4] The study stressed the uncertainties in predicting the economic impact of an accident.[4] Such evaluation of economic losses generally includes an allowance for pain and suffering, which in a way quantifies the emotive factor associated with risk. The variability of economic factors is given by the statement: "We can reduce the risk in any sector provided we are prepared to pay the cost."[5] This exposes the concept of opportunity cost which underlies most economic decisions.

There are clear indications from ref 3 that the authorities responsible for deciding about the acceptability of sites attempt to take into account the need for economic development in the area surrounding a proposed site,[3] but there are no universally agreed ways in which such factors are taken into account and this fuzziness in the economic argument shows the need to evaluate the uncertainty or ambiguity in the data when determining the ranking.

In a study by Chicken of the correlation between expenditure on life saving and the public's perception of risk the following conclusions were drawn, which are relevant to the ranking.[6]

1) The level of expenditure on risk reduction is more related to people's perception of risk than to estimates of the probability of the risk.

2) Policy makers are willing to contemplate higher levels of expenditure to reduce involuntarily accepted risk than for voluntarily accepted risks.

3) The value of life for compensation purposes, including an allowance for pain and suffering, often seems to be put at about £200,000, but there is a considerable range in such valuations.

4) The range of cost of saving an extra statistical life (often referred to as the CSX value) used is from £0 to 2×10^9. For many decision

situations the values used are in the range £10^4 to 10^5.

5) The cost of action to save a life is sometimes higher than the compensation paid for loss of life.

A more recent study of the way human life is valued for various purposes draws attention to the view that life insurance does not really value life but just amounts to saving to provide for dependants or the future.[7] The values of human life reported, in ref 7, range from £1 x 10^3 for a child-proof drug container to £20 x 10^6 for a change in British building regulations.

The conclusion that seems to be justified about the economic data is very similar to the conclusion about the technical data, but there is even greater uncertainty about the data that has to be used, and this must be taken into account in determining the ranking score.

At the heart of assessing socio-political factors is determination of public opinion. Public opinion is fickle and opinions do change and this variability has to be allowed for. Public opinion can be assessed either from the opinions of those active in the field or by polls. Dr. Keyes, a Director of the Westinghouse Electric Corporation, has summarized the roll of polls in the following incisive way:[8] "Certainly, the techniques are not perfect. Sometimes the pollster errs; sometimes, his client. Nevertheless, polling is one of the most important tools available in measuring public attitude on certain issues in order to ascertain the public will which, in the long run, will find expression in the actions of government in a democratic society." One important weakness of polls is that unless they are carefully designed they can give too much emphasis to the views of people who are not concerned with the issues involved.

If there is opposition and it is based on a lack of information or on a misunderstanding of information given, it is possible by judicious publicity and education processes to reduce it, but such processes can take a considerable time.[9]

The conclusion about socio-political data that seems to be warranted is that although a great diversity of views are involved there are survey techniques available that enable opinions to be assessed in a way that indicates the ranking that may be justified.

3. DEMONSTRATION OF THE APPLICATION OF THE RANKING TECHNIQUE

To demonstrate how the ranking technique may be applied in practice two well known cases, Moss Morran and Rijnmond, that have caused a certain amount of controversy were assessed.

3.1 The Moss Morran Facilities and Pipeline

The Moss Morran liquefied energy gas terminal facilities and pipeline was planned as part of the development required to exploit the Brent oil and gas field in the North Sea. There were three main stages in the process leading to outline permission being granted.[3] Stage 1 was Shell and Esso formally lodging planning applications to develop a processing facility at Moss Morran. Stage 2 included the Secretary of State for Scotland calling for the decision to be made at central rather than local government level, a public inquiry into the acceptability of the proposal being held, local government authorities publicising the fact that the planning applications had been lodged and describing the general nature of the proposals, the hazards involved and the environmental impact of the

proposals being assessed, and the directors of planning of the local authorities concerned preparing a report on the socio-economic impact of the proposal. The conclusion of Stage 2 was marked by the Secretary of State receiving the report of the public inquiry. Stage 3 started when it was found that the public inquiry had exposed differences in views and concern about the possibility of a vapour cloud being ignited by radio frequency transmissions. The Secretary of State invited written comments on the subject. It was finally concluded from tests that radio frequency transmissions were unlikely to produce sufficient power to reach the minimum required for ignition.[10] In August 1979 the Secretary of State announced he would grant outline planning permission.

The risks associated with the pipeline from St. Fergus to Moss Morran were assessed by the Health and Safety Executive in 1978 and again in 1980 when it was proposed to increase the pipeline size from 16-inch to 24-inch diameter.[11] The assessment showed that the chance of leakage from the pipeline reaching people in the area of the pipeline fell in the range 1 to 4 x 10^{-6} per year. The report advised: "...the level of risk would not be such as to lead to a recommendation that a Construction Authorization should be withheld on health and safety grounds".

Although consequence analysis calculations were made the results were expressed in qualitative terms like low, very low or extremely low and gave no estimate of the possible number of fatalities.[3] The Action Group estimated[3] that the probability of an individual fatality was 7 x 10^{-4} per year.[3] The one hazard figure that seems to be very high is the shipping hazard figure, the same figure of 10^{-3} per year appears to have been used in the assessment of the acceptability of Eemshaven, Braefoot Bay and Wilhelmshaven. This apparent statistical anomaly has been adversely commented on in reference 3.

Concern about the safety and risk justification of the Moss Morran site is indicated by the conditions that were attached to the outline planning permissions granted to Shell and Esso.[3] The most important condition from the safety point of view being the requirement that a full hazard and operability audit should be conducted before the facilities are allowed to be commissioned. The importance attached to the audit is indicated by the fact that the Secretary of State decided the audit must be to his satisfaction and not just to the satisfaction of the Health and Safety Executive.[3]

From the information above and assuming that all the planning conditions are satisfied it is considered that the risk ranking that can be justified for Moss Morran is 3. The construction of the ranking is shown in Table 2.

3.2 The Rijnmond Decision

The history of the Rijnmond Decision is complicated and has its origins in the early 1970's when plans were made to import large quantities of liquefied natural gas (LNG) from Algeria.[3] Eight possible sites for the LNG terminal were considered.[3] The two main contenders were Rotterdam and Eemshaven.

The discussion that took place about which was the most acceptable site also took place in three stages.[3] Stage 1 was the period up to the final signing of the contract for the supply of LNG and included the preliminary search for a terminal site. Stage 2 involved the cabinet and several government departments and at this stage it was recognized that siting involved several issues such as energy policy, the environment, safety, land use and regional planning. At the beginning of this round,