Modeling Chronic Versus Acute Human Risk from Contaminants in Food

by A.L. Carriquiry, H.H. Jensen, and S.M. Nusser

Working Paper 90-WP 69
September 1990

Center for Agricultural and Rural Development
Iowa State University
Ames, Iowa 50011

A.L. Carriquiry is an Assistant Professor of Statistics, H.H. Jensen is an Associate Professor of Economics, Iowa State University, and S.M. Nusser is a Researcher, The Procter and Gamble Co., Cincinnati, Ohio.

This research was supported in part by the Human Nutrition Information Service, USDA, Research Agreement No. 58-3198-9-032, and in part by the Cooperative States Research Service, USDA Food Safety Consortium. This paper was presented at the Economics of Food Safety Conference, Alexandria, VA, June 4-6, 1990.

Abstract

Modeling Chronic Versus Acute Human Risk Contaminants in Food

A.L. Carriquiry, H.H. Jensen, and S.M. Nusser

The development of policies and regulations to address food safety concerns depends critically on appropriate assessment of health risk in foods. This paper evaluates the methods for assessing the population's exposure to a hazardous substance or contaminant in food and some aspects of the quantification of risk. We review current federal programmatic approaches to risk assessment and potential problems with these approaches. After developing procedures for estimating exposures of individuals in a population to chronic and to acute risks, we illustrate their application by using available food consumption data to estimate exposure and highlight issues related to the data requirements for risk assessment.

1. Introduction

The supply of food in the United States is abundant and varied year round. While consumers benefit from a wide range of food products available in the market, consumers are increasingly concerned about the safety and quality of food they eat. Many questions have focused recently on the safety of food additives, food production techniques, and pesticide residues in foods. These concerns have had an impact on government policymakers and government regulators responsible for maintaining a safe and adequate food supply, and on the food industry which has an interest in providing safe products and in responding to consumers' preferences for products.

The recent dialogue on the public response to food safety has made it clear that the development of policies and regulations to address food safety concerns depends critically on appropriate assessment of health risk in foods. The widely varied food supply and diverse food eating behaviors in the U.S. require a highly integrated information base to support this assessment and continual monitoring of safety in the food supply. Furthermore, the divergence in consumers' perceptions of food-borne hazards and the assessment made by food safety experts in and outside of government (Kramer 1990) highlight the need to use valid and appropriate methods for assessing and monitoring the potential for hazard in the food supply.

Food producers and processors, chemical manufacturers and other input suppliers, and government policymakers are faced with the need to address very difficult questions with respect to risk assessment. For example, those responsible for regulating pesticides must understand the extent and timing of

health risks associated with the pesticide and have reliable information on the degree of exposure of the population to residues of the pesticides in food, water, and air. Knowledge about the assessment of risk provides guidance for the development of government policies and regulation of industry, information for communicating relative risk exposures to consumers and policymakers, and indicators for monitoring changes in the quality of the food supply.

What is the appropriate method by which to assess the health risk associated with ingesting a particular pesticide residue in food? or consuming a contaminated food product? In order to best address these questions, risk assessment can be defined as a process (Barry (1987); Mauskopf (1990)). The risk assessment process involves four steps:

(i) Identification of food constituents with potentially adverse health effects.

(ii) Estimation of the exposure of the population or subpopulation to the hazard, for a certain period.

(iii) Determination of the response to different doses of the hazard (dose-response modeling).

(iv) Characterization of the risk (providing information on probable health effects of the constituent combined with exposure and dose-response estimates to produce quantitative estimates of health hazards).

In this paper, we address the methods for assessing the population's exposure to a hazardous substance or contaminant in the food. To a lesser extent, we discuss some aspects of risk quantification. This information is essential to (1) assessing and monitoring risk exposure, (2) setting

priorities for risk reduction, and (3) developing information and education programs targeted to those at greatest risk.

The paper is organized as follows. In the second section we provide background information for defining adverse health effects, current programmatic approaches to risk assessment, and potential problems with these approaches. Procedures for estimating the distributions of usual daily exposures of individuals in a population are presented in the third section. These distributions apply to a selected pesticide residue or contaminant in the food supply. In the last section we discuss some issues related to the data requirements for such assessment.

2. Background

2.1 A Brief Comment on Possible Adverse Health Effects

One possible classification of hazardous substances in foods is into carcinogenic and noncarcinogenic agents. This distinction is relevant both for assessing exposure and for modeling response to dose. Modeling the response to dose is different for carcinogens and noncarcinogens because the toxic endpoints vary. For carcinogens, only one endpoint, death from cancer is frequently considered (Mauskopf 1990), as we do here, even though adverse health effects may be more complex. In the case of noncarcinogens, the health effects associated with a toxic substance may be multiple. Adverse health effects from carcinogenic agents are generally recognized to be chronic, and this implies that long-term exposure is the major concern. For noncarcinogens, however, both acute and chronic health effects are likely, and hence, exposure assessment concerns both long and short-term effects. In this

paper, we consider both long-term and short-term exposure assessment, and therefore, the methodology is appropriate for both types of agents.

2.2 Some Current Risk Assessment Programs

Due to continuing regulatory activities, the Environmental Protection Agency (EPA) and the Food and Drug Administration (FDA) have been concerned with the problems of risk assessment and management. In 1987, the EPA published five technical guidelines to aid agency personnel in their risk assessment activities. The Guidelines for Estimating Exposures set forth general principles and procedures for estimating the degree of chemical contact with an affected population, including the steps to be followed for exposure assessment (EPA, 1987). These are:

(i) Source characterization

(ii) Pathways and fate analyses

(iii) Estimation of environmental concentration

(iv) Demographic analysis

(v) Integration.

The FDA has followed such an approach as well, using a representative diet to track changes in the food supply (Pennington and Gunderson, 1987).

For exposure assessment, the general approach recommended by EPA for obtaining exposure estimates for most chronic exposures is to estimate average daily lifetime exposure, in mg/kg/day (EPA, 1987). That is,

$$\text{Average daily lifetime exposure} = \frac{\text{Total dose (mg)}}{\text{Body weight (kg)} * \text{Lifetime (days)}} \qquad (1)$$

where

$$\text{Total dose} = \text{Environmental concentration} * \text{Contact rate} * \text{Exposure duration} * \text{Fraction absorbed} \quad (2)$$

In the case of exposure to pesticide residues or other contaminants in the food supply, we can express the total dose as:

$$\text{Total dose} = \text{Concentration of toxicant in food supply} * \text{Amount ingested} * \text{Days ingestion} * \text{Fraction absorbed}, \quad (3)$$

where fraction absorbed refers to the effective proportion of the contaminant crossing an exchange membrane (i.e., gastrointestinal tract). The fraction absorbed is presumably difficult to assess for each individual; it depends on a large number of individual attributes such as age, genetic makeup, health status, type of residue in the food, etc.

Expression (1) or slight modifications of expression (1) are used by exposure analysis software such as that developed by Technical Assessment Systems (TAS) (1985). TAS, under request from the EPA, developed a menu-driven program called Exposure-1, which allows for estimation of chronic exposure of the population at large, and of 22 subpopulations, to any toxicant in the food supply. Exposure-1 outputs exposure estimates in two different formats: (1) as mgs of the chemical/kg body weight/day, or (2) as percentage of acceptable daily exposure (ADE).

The TAS program draws information from two sources: food consumption files and chemical residue files. The food consumption files are given, and contain measures of the estimated daily intakes of each food and food forms by individuals (based on the USDA 1977-78 Nationwide Food Consumption Survey (NFCS)). By using weighted means for daily food intake, the TAS Exposure

system estimates usual daily intake of foods for each of the 22 subpopulations. Subpopulations are defined in terms of demographic variables such as age, gender, and race. The chemical residue files are supplied by the user and contain, for each chemical, the food or food forms in which the chemical appears, the concentration in which it appears (either tolerance level or anticipated residue) and appropriate adjustment factors for the concentration of the chemical in the food or food form after different stages of processing. Thus, the system estimates the total dose for the 22 subpopulations.

2.3 Some Problems with Current Exposure Assessment Programs

Exposure analysis methods, such as those currently in use by EPA and TAS, do not rely on estimates of the intake distributions but summarize the whole distribution in a point estimate of an average individual's usual intake. In an attempt to take into account the interindividual variation, methods like TAS' Exposure-1, separate the general population into several subpopulations according to factors such as age and race, and estimate exposure to a food constituent using the mean consumption in the subpopulation. A similar approach is outlined in the EPA's Technical Guidelines for Exposure Assessment (1987).

There are several problems with such an approach. First, the estimates of average daily lifetime exposure assume that there is no interindividual variation regarding total dose or body weight. Even within subpopulations of individuals grouped by age, there exists variation in contact rate, exposure duration and fraction absorbed. Therefore, a better method for assessing exposure would rely on estimating expression (1) for each individual in the

sample and using this information to estimate the distribution of "usual daily exposure" of individuals to a certain toxicant. By using a point estimate to quantify exposure to a food constituent for the whole subpopulation, individual information is lost.

Furthermore, using a single average daily lifetime exposure assumes that the total dose is constant throughout the individual's lifetime. This is clearly not true in the case of pesticides ingested with the food, since the type and amount of food consumed varies with age. A simple correction for this consists of considering average daily exposure in a certain age range, and then computing lifetime exposure as the sum of the exposures in each period. In this case, lifetime in expression (1) would be changed to the number of days in each period considered. Furthermore, contact rate, exposure duration, and absorption rate would also be changed to their appropriate values for each period.

Estimates of usual daily intakes of individuals often are based on intake data sets, such as the 1977-78 NFCS, which contain intake data for a sample of individuals for a few days. However, because the observed daily intakes measure usual daily intake with error, it is important to account for the intraindividual variation in estimating usual daily intake in order to avoid attributing higher reliability to the estimates than is justified by the data. In the next section we present an approach for estimation of usual intake distributions which takes into account inter- and intraindividual variation and incorporates both into the analysis. It should be noted that assessing chronic exposure to a food constituent can be viewed as the same problem as assessing nutrient adequacy. Therefore, we adopt methodology developed by Nusser et al. (1990).

In the analysis which follows, we present a procedure for estimating distributions of usual daily exposures of individuals in a population to pesticide residues or contaminants in the food supply. Total dose (mg) is taken to be the concentration of the pollutant, contaminant (or pesticide residue) in each food or food form times the amount of food or food forms containing the chemical that is ingested by individuals in the population. Exposure duration is taken to be one day, and fraction absorbed is assumed to be equal to 1. Clearly, the fraction absorbed could be changed to show an appropriate alternative value. In that case, we would be estimating the distribution of usual daily absorption rather than that of usual daily intake. The distributions we estimate are those of usual daily intake of a food constituent per individual. It would be a simple matter to obtain estimates of usual daily intake per kg of body weight by including information on individuals' weights.

3. Statistical Methodology for Assessing Exposure

3.1 Overview of Issues

Chronic and acute exposure to pesticide residues or other agents in the food supply can be estimated from dietary intake data and information on residues in foods and food forms. In this paper, we refer to chronic exposure as the low-intensity, daily intake of a pesticide residue, which accumulates for a long period of time before any adverse health effects are evident. By acute exposure, we mean a one-time intake of a toxic agent, in quantities enough to produce an adverse health effect.

To assess chronic or long-term exposure to a toxic agent in the food, it is necessary to estimate the average or usual daily intake of foods containing

the agent by individuals in the population. This is appropriate, for example, for carcinogenic agents. Usual daily intake of a pesticide residue (or of a nutrient) is defined as the normal or long-run average intake. It is explicitly recognized that intake of the toxicant on one day is not an indicator of chronic exposure; rather, it is intake of the residue over a long period of time that places an individual at risk of adverse health effects.

Usual intake of chemicals can be obtained from dietary data. Ideally, the data should include information on dietary intake for a large number of individuals, on a large number of days. Unfortunately, it is usually possible to obtain just a few days of intake data for individuals in the sample. An individual's usual daily intake is often estimated by the individual's mean daily intake of the residue under consideration. While the individual's mean intake is a reasonable estimator of the individual's usual daily intake, the distribution of mean intakes is not a good estimator of the distribution of usual intakes. The distribution of means has always a larger variance than the usual intake distribution. Therefore, exposure estimates obtained from the distribution of mean intakes could be inflated. The degree by which exposure is overestimated depends on the shape of the distribution of usual intakes. This is illustrated in Figure 1.

When the objective is to assess acute or short-term exposure to a food constituent different procedures must be employed. Is it not the usual or average consumption of the food component which is relevant, but the amount ingested on any given day. Consider, as an example, exposure to salmonella from contaminated eggs. Even if, on the average, an individual consumes small amounts of eggs, it is the number of contaminated eggs consumed in one day which will determine whether the individual gets sick or not. In general

terms, the probability that a randomly chosen individual from the population suffers acute exposure to a food constituent can be estimated from the population's probability distribution of consumption of the food on a given day, times the probability that the food is contaminated.

Two types of assumptions are made in this assessment of acute exposure: (1) it is assumed that unless the food looks or smells differently than usual, the food intake distribution is independent of contamination, and (2) if an individual consumes more than one portion or unit of food, it is assumed that the portions are either (a) independent, or (b) not independent. The choice of 2 (a) or 2 (b) will depend, for example, on the type of contaminant under consideration, the foods or food forms likely to contain it, and individual eating patterns. If portions are assumed to be independent, then whether a portion is contaminated will not affect the status of the other portions. Survey-based dietary data provide information from which to estimate the food intake distributions. In some cases, we may be interested in a frequency distribution, as in the hypothetical case of eggs mentioned earlier. However, the distribution of the presence of the contaminant in each food or food form in which it may appear is usually not known. Most often, the information available includes only the probability of finding the contaminant in each food or food-form at a level deemed hazardous.

Sections 3.2, 3.3, and 3.4 are organized in the following manner. Section 3.2 contains a description of available dietary intake data. The methodology for assessing chronic exposure to a food constituent is presented in Section 3.3. The proposed procedure is described in some detail, since obtaining reliable estimates of chronic exposure is very relevant from a public policy viewpoint. In Section 3.4 we state the problem of estimation of

acute exposure more precisely, and suggest estimation procedures appropriate for different scenarios.

3.2 Dietary Intake Data

When collecting dietary intake data for the estimation of exposure to food contaminants, it must be recognized that different data attributes are important for chronic or acute exposure assessment. In the case of chronic exposure assessment, it is necessary to obtain reliable estimates of the usual intake distribution for the food constituents of interest. Data suitable for estimating usual intake distributions of dietary components should allow for the estimation of between- and within-individual variances. One-day intake data on individuals allow for the estimation of between-individual variance. However, estimation of the within-individual variance requires that the data set include more than one day of intake on each individual. Thus, using only one day intake data is not appropriate for estimation of usual intake distributions of food constituents, since one-day intake data sets do not provide a means for estimating the within and the between-individual variances.

Assessment of acute exposure, however, does not rely on the estimation of usual daily intake distributions. It suffices to be able to estimate between-individual variances, but it is not necessary to obtain estimates for within-individual variation. Therefore, inferences about acute exposure can be based on one-day dietary intake data.

The U.S. Department of Agriculture's intake data sets provide multi-day data from which to estimate the usual intake distribution of food constituents. For this study, data from the Continuing Survey of Food

Intakes by Individuals (CSFII) were used to help develop the methodology described in Sections 3.3 and 3.4. The CSFII data were collected by the Human Nutrition and Information Service (HNIS) of the USDA in 1985-86. Women between 19 and 50 years of age provided data on their own daily dietary intakes and those of their pre-school children, in addition to information on household composition, sociodemographic information and eating behaviors (e.g., meal patterns) (see USDA, 1987). The sample was a nationwide, multi-stage stratified area probability sample drawn from the 48 coterminous states. The primary sampling units were area segments, and the probabilities of selection of area segments were proportional to the numbers of housing units in the segments as estimated by the Bureau of the Census. USDA constructed a data set on four days of data available for analysis. The days' data, collected throughout the year, were assumed to be independent. The analysis described below was based on a subset of the CSFII 4-day data set corresponding to 23-50 year old women who were not pregnant or lactating. The dietary intake data were matched to the extensive nutrient data banks at USDA to obtain data on nutrient intakes. These data on nutrient food components were used to develop the methodologies for estimating usual intake distributions.

3.3 Assessing Chronic Exposure

3.3.1 Overview

Preliminary analyses of the CSFII intake data (Jensen et al., 1989) indicate that intake data for nutrient and other food components are not normally distributed. Intake distributions are sometimes severely skewed, which makes the assumption of normality untenable. In order to estimate the

distribution of usual intakes of a food component, therefore, it is necessary to adopt one of two possible approaches: 1) assume an appropriate parametric model for the intake distribution, (such as a Gamma or a Weibull distribution) and derive the estimators within that parametric framework, or 2) transform the data to normality. The first approach was adopted by Battese et al. (1988) in the context of the estimation of usual nutrient intake distributions. This parametric approach is computationally involved. The transformation to normality approach, on the other hand, can be applied to any food component without modifications, and estimators and predictors of usual intakes can be derived by invoking results from normal theory. In what follows, we describe the transformation approach to estimating usual intake distributions.

The estimation of the distribution of usual intakes of a food component is based on a non-parametric approach to transforming the data to normality. The objective of this approach is to produce transformed observations that are normally distributed and have homogeneous variances. The methodology is developed in Nusser et al. (1990).

The approach we suggest for estimating the distribution of usual intakes of a food constituent involves the following steps: 1) observed intakes are transformed to normality, 2) the normal data are assumed to follow a measurement error model that decomposes the observed daily intake of an individual into the usual intake for that individual plus a measurement error associated with the individual on the day the intake was observed, 3) normal theory is then used to obtain predictors of usual intakes in normal space for each individual, 4) application of an inverse transformation to the predicted normal usual intakes produces a set of pseudo usual intakes in the original

scale, which can then be used to estimate the distribution of usual intakes. The measurement error model approach requires an estimate of the within-individual variation, which can be obtained only if data for each individual are available for more than one day.

3.3.2 Transforming the Observed Data to Normality

The transformation of the observed data to normality consists of the following steps: first, a smoothed empirical cumulative distribution function (c.d.f.) of the observed daily intakes is evaluated at each of these values to produce a set of uniform random variables. The inverse normal c.d.f. is then used to transform the uniform variates into a set of standard normal random variables. Let Y_{kij} denote the observed intake of a dietary component k for individual i on day j, where k=1,...,p components, i=1,...,n individuals, and j=1,...r days. Assume that individuals, as well as daily intakes within individuals, are independent. The empirical c.d.f. constructed from the nr Y_{kij} values is a step function. By connecting the midpoints of the rises between the steps defined by the empirical c.d.f., a continuous piecewise linear estimate of the true c.d.f. F_{Y_k} is constructed. For this choice of midpoints, the continuous c.d.f. yields approximately the same mean value of the data as the empirical c.d.f.

The estimated continuous c.d.f. provides a means of generating a set of uniform (0,1) variates, p_{kij}, from the observed intakes. Therefore, given the standard normal cumulative distribution function $\Phi(\cdot)$,

$$X_{kij} = \Phi^{-1}(p_{kij})$$

are N(0,1) variates (e.g. Lindgren, 1976). The X_{kij} represent the transformed observed values. It may be the case that the transformed values do not have homogeneous within-individual variances. If so, a further transformation is required to homogenize the within-individual variances. The methodology presented later in this paper relies on the assumptions of normality and of homogeneous within-individual variances.

3.3.3 Predicting Usual Intakes in Normal Space

Normal theory and a measurement error model can be used to generate predicted usual intakes from the transformed observed intakes. The prediction methodology is well suited for application to a vector of dietary components. The multivariate approach permits incorporation of information contained in the relationships among intake patterns of dietary components into the prediction of normal usual intakes.

Assume that data are available for p dietary components on each individual. Suppose that for each dietary component k, the nr values of Y_{kij} are transformed, using the methodology in Section 3.3.2, to generate nr X_{kij} normally distributed values. Denote the p x 1 vector of transformed observations for individual i on day j by $\mathbf{X}_{ij}$.

A measurement error model is used as a basis for predicting the usual intakes given the observed intakes. This model recognizes that the observed daily intake for an individual on a given day is equal to the sum of the usual daily intake of the individual and a measurement error associated to the individual on that day. Let

$$\begin{aligned} \mathbf{X}_{ij} &= \mathbf{x}_i + \mathbf{u}_{ij} \\ \mathbf{x}_i &\sim N\ (\mu_x,\ \Sigma_{xx}) \\ \mathbf{u}_{ij} &\sim N\ (0,\ \Sigma_{uu}), \end{aligned} \tag{4}$$

where $\mathbf{x}_i$ is the vector of unobservable usual intakes for individual i; $\mathbf{u}_{ij}$ is the unobservable measurement error for individual i on day j; the $\mathbf{x}_i$ are independently distributed; the $\mathbf{u}_{ij}$ are independent across days; and $\mathbf{x}_i$ and $\mathbf{u}_{ij}$ are uncorrelated. Assume that Σ_{xx} and Σ_{uu} are positive definite. This model implies that the $\mathbf{X}_{ij}$ are $N(\mu_x, \Sigma_{xx} + \Sigma_{uu})$ variates, and that the sample individual means

$$\bar{\mathbf{X}}_i = r^{-1} \sum_{j=1}^{r} \mathbf{X}_{ij} \tag{5}$$

are independent random variables from a $N(\mu_x,\ \Sigma_{\overline{XX}})$ distribution, with

$$\Sigma_{\overline{XX}} = \Sigma_{xx} + r^{-1}\ \Sigma_{uu}\ . \tag{6}$$

It should be noted that if the normal observed intakes from the initial transformation described in Section 3.3.2 are used in this model, $\mu_x = 0$. However, μ_x may be non-zero if further transformations are required to obtain homogeneous error variances for the transformed intakes.

Our objective is to produce a set of pseudo usual intakes whose distribution is close to that of true usual intakes. That is, we want to predict a set of pseudo usual intakes $\mathbf{x}_i$ whose covariance matrix is Σ_{xx}. The Best Linear Unbiased Predictor of $\mathbf{x}_i$ (BLUP) has smallest prediction error variance among all unbiased linear predictors, and so would be appropriate if the objective was to predict individual $\mathbf{x}_i$. However, if the BLUP is used to

predict a set of $\mathbf{x}_i$, the variance of the predicted $\mathbf{x}_i$ is smaller than Σ_{xx} , therefore, the distribution of the BLUPs of $\mathbf{x}_i$ is not close to that of the true $\mathbf{x}_i$. Predictors of $\mathbf{x}_i$ with variance Σ_{xx} can be obtained by using

$$\mathbf{x}_i^* = \mu_x + \Sigma_{xx}^{1/2} \underset{xx}{\Sigma_{--}^{1/2}} (\bar{\mathbf{x}}_i - \mu_x) \quad . \tag{7}$$

The values of μ_x, Σ_{xx}, and $\underset{xx}{\Sigma_{--}}$ are unknown. Therefore, to implement the procedure of equation (7), estimates of μ_x, Σ_{xx}, and $\underset{xx}{\Sigma_{--}}$ can be substituted into (7) in the appropriate places.

Usually, inferences are made about the exposure of the target population regarding a single food constituent (in which case, p=1). It may also be of interest, however, to assess exposure with respect to a vector of constituents. The methodology can be used to make simultaneous inferences about exposure of the target population to more than one food constituent. For example, suppose that we want to know the proportion of the population exposed to all p constituents, where exposure is indicated by usual intakes above a vector $\mathbf{k}$ of Acceptable Daily Intakes (ADI's). In normal space, this proportion is given by $Pr(\mathbf{x}_i^* > \mathbf{k}^*)$, where the $\mathbf{x}_i^*$ are obtained from the X_{ij}, and $\mathbf{k}^*$ is the transformed vector $\mathbf{k}$. Alternatively, predicted normal usual intakes can be transformed back to the original scale using the transformation described in Section 3.3.4, and inferences can then be made from usual intake distributions estimated in the original scale. Note, the transformation procedure outlined in Section 3.3.2 produces a set of N(0,1) variables, but the transformed intakes of the p food constituents are not necessarily multivariate normal. If it is desired to estimate simultaneous exposure to a

set of food constituents, then further transformations can be used to approximate multivariate normality (Nusser et al., 1990).

3.3.4. The Mean Transformation

The predicted usual intakes in normal space can be transformed to obtain a set of pseudo usual intakes in the original space. To generate a set of pseudo usual intakes in the original data scale from the normal usual intakes, a transformation from the normal space to the original scale is required. This transformation, called the mean transformation, should have the property that the usual intake in the original scale is equal to the mean transformation of the normal usual intake. Note that since the transformation from observed intakes to normal observed intakes is nonlinear, the inverse of this transformation cannot be used to transform normal predicted usual intakes (which are like means) back to the original scale. Preliminary analyses indicate that the mean transformation can be accomplished via the use of cubic splines (e.g., Ahlberg et al., 1967). The methodology consists of fitting a grafted polynomial function with linear end segments and cubic interior segments, to (X_{ijk}, Y_{ijk}) pairs. The estimated function can then be used to transform the predicted usual intakes in normal space (x_i^*) to pseudo usual intake in the original scale. A detailed discussion of the mean transformation is presented in a separate publication (Nusser et al., 1990.)

3.3.5. Assessing the Proportion of Individuals with Usual Daily Exposures Above a Critical Level.

An example for a selected food constituent illustrates the method for chronic exposure assessment based on the USDA dietary data. Following the procedures described in Sections 3.3.3 and 3.3.4, the predicted normal usual

intakes were transformed back to the original scale to obtain usual intake distributions in the original scale. The percentage of the population with usual daily exposure above a critical level can be estimated as the area under the curve to the right of the critical value. This is illustrated in Figure 1. The estimated proportion is calculated using the estimated usual intake distribution, the 4-day mean intake distribution and the one-day intake distribution. First, note that the estimated distributions differ; the distribution estimated from one-day dietary intake data has the largest variance. Second, when using an example critical value of 23, the one-day intake distribution overestimates the percentage of the population with intakes above the critical level: the usual intake distribution shows 1.1 percent of the population to have levels in excess of 23, compared to 5.2 percent estimated using the one-day intake distribution.

3.4. Assessing Acute Exposure

3.4.1. Overview

When considering short-term exposure to a toxic agent in the food, different questions may be of interest. For example, given a known No Observed Effects Level (NOEL), of a certain food constituent, it may be important to determine the probability that a randomly chosen individual from some population has an intake of the constituent on any given day which exceeds the NOEL. It may also be interesting to determine what is the probability that an individual's intake of the constituent on any given day exceeds the NOEL, given that the individual consumes a certain amount of the food containing the toxic agent. To answer the first question, information must be drawn from two sources: (1) the dietary intake data can be used to

determine proportions of the population with different levels of intake of the constituent on any given day. (2) The NOEL, as well as the probability of a unit of food or food-form of carrying the constituent at a level above the NOEL, are determined from the toxicological parameters of the constituent as well as from extensive testing of the foods. The second question is, from a statistical viewpoint, contained in the first. When the consumption of an individual is given, it is enough to know the probability of the food consumed being contaminated.

3.4.2. Estimating the Distribution of Daily Intake of the Population

Consider, for example, a food A which may be contaminated with a constituent. It is known that at levels above N_θ, intake of θ causes an adverse health effect. Further, suppose one portion is a unit of consumption for food A. The frequency distribution for the consumption of food A in the population can be determined from the data set described in Section 3.2. The proportion of individuals who, on any given day consume 0,1,2,... portions of A can be obtained in a straightforward manner.

Let $Pr\{x \text{ portions}\} = p_x$, $x = 0,1,2,\ldots$ denote the probability that a randomly chosen individual from the population consumes x portions of A on any given day. Probabilities p_x can be estimated as frequencies, that is, $p_x = n_x/N$, where n_x is the number of individuals in the sample consuming portions of A on any given day, and N is the total number of intake observations in the sample. Note that N will be larger than the number of individuals in the sample when more than one day of intake data for each individual is considered.

3.4.3 Estimating the Probability of Contamination Given a Certain Food Consumption: Independence Assumption.

When it is assumed that food units are independent with regard to the presence of a toxic constituent, it is recognized that whenever one portion of the food is contaminated, this does not affect the contamination of any other portion. In this case, we are concerned with the probability that each individual portion of the food be contaminated. This is a realistic assumption when it can be assumed that different portions came from different sources.

Let c denote the probability that one portion of food A carry amounts of θ above N_θ. Then, it is clear that given an individual who consumes on any given day only one portion of A, the probability that this individual suffer adverse health effects is given by c, where

$$\Pr\{\text{adverse health effects/one portion of A}\} = c.$$

For an individual consuming two portions of A this probability is computed as:

$$\begin{aligned}&\Pr\{\text{adverse health effects / 2 portions of A}\}\\ &= 1 - \Pr\{\text{none of the portions is contaminated}\}\\ &= 1 - (1 - c)^2.\end{aligned}$$

Similarly, given an individual who consumes on any given day x portions of A, his or her probability of suffering adverse health effects due to the toxic agent is given by:

$$\Pr\{\text{adverse health effects / x portions of A}\} = 1-(1-c)^x .$$

The risk of suffering adverse health effects increases as the number of portions of A consumed increases.

3.4.4 Estimating the Probability of Contamination Given a Certain Food Consumption: Complete Dependence Assumption

In many cases, the independence assumption on food portions is untenable. Often, it is more appropriate to assume that whenever an individual consumes more that one portion of a food, either none is contaminated or all of them are. Consumption of any one portion leads to adverse health effects. Consider, for example, a hypothetical toxicant which may show up in chicken and eggs. If an individual consumes more than one portion of chicken on a given day, most likely both portions came from the same chicken. Likewise, eggs consumed on the same day came from the same carton.

Let c^* denote the probability that a unit of food A contains levels of θ higher than N_θ. The probability that an individual who consumes one portion of A on any given day exceeds the NOEL intake of θ is given by:

$$\Pr(\text{adverse health effects} \mid \text{one portion of A}) = c^*,$$

as in the independence case. However, the probability that an individual experience an adverse health effect when consuming more than 1 portion A is also equal to c^*. It does not matter how many portions the individual consumes; if it is assumed that all portions come from the same unit, then the risk of adverse health effects is given by the probability of the unit being contaminated. The risk of suffering adverse health effects does not depend on the amount of food consumed.

It should be noted that the complete dependence assumption makes sense only when considering portions of the same food. In the hypothetical case of an individual consuming 2 portions of chicken and 3 of eggs, chicken and eggs should be considered independent regarding the chance of contamination.

3.4.5. Estimating Risks for Randomly Chosen Subjects

It is now quite a simple task to answer the following question: what is the probability that a randomly chosen individual from some population suffer an acute adverse health effect due to intake of a toxic agent in food A? Recalling that p_x denotes the probability that a randomly chosen individual consumes x portions of A, and that c and c^* represent probabilities of contamination of independent and independent units of food, respectively, we can write

$$\text{(independence)} \quad Pr\{H\} = p_0\,(x)(0) + p_1(x)(c) + p_2(x)[1-(1-c)^2] + \ldots$$

or

$$\text{(complete dependence)} \quad Pr\{H\} = p_0(x)(0) + (c^*)(x)[p_1 + p_2 + p_3 + \ldots].$$

where H = "a randomly chosen individual from the population suffers adverse health effects due to acute exposure to θ".

3.4.6. Partial Independence Assumption.

Perhaps a more realistic assumption regarding contamination of different portions of a food or food-form is that of partial independence. In the

example of the carton of eggs, it may be more appropriate to assume that there exists a dependence among the eggs, but that this dependence is not complete. This assumption, however, cannot be entertained in a practical sense, since the required information is not available. In the carton of eggs example, it would be necessary to know the joint probability distribution of contamination for all 12 eggs. Further, we would need to be able to derive appropriate marginal distributions of contamination of 1, 2, ... eggs. This last step would involve integrations in high dimensions, a costly procedure.

4. Implications for Data Requirements

The methods for dietary exposure assessment have implications for the collection and design of data which would support analysis of food risk, risk monitoring, and economic evaluations of food hazards. For chronic exposure, the methods for estimating usual intake or exposure distributions require multiple days of observed intake for individuals. Preferably, intake should be observed on days sufficiently apart in time so that independence of day within individuals can be assumed. This was the case for the 1985 CSFII data used in our initial analysis. Larger samples of individuals will contain more information on variation between individuals, and more days per individual provides more information on intra-individual variation.

The data requirements for acute exposure differ. The information required is for portions consumed in a given day. Only one day of intake data is required for each individual in the sample, since an estimate of intra-individual variation is not necessary. If more days are available per individual, the entire data set can be pooled together. Additional

information on cooking methods, food handling techniques, and eating habits may be required.

It is important to recognize that the methods used for dietary exposure assessment have implications both for the design of federal regulatory and monitoring activities, and for the implementation of risk management strategies. Incorrectly estimating levels of risk associated with intake of specific food constituents will not only lead to misallocation of resources to risk-reduction activities, but may also alter governmental priorities in reducing risk. The methods we suggest will reduce the error incurred in the estimation of exposure of populations to hazardous food constituents. As illustrated in Figure 1, this error can be considerable. It is not clear how this error compares in magnitude to those that result from the other steps in the risk assessment process. However, the use of appropriate methods of modeling human health risk from foods will lead to improved overall assessments and risk management strategies, and properly guide the development of federal data collection efforts.

References

Ahlberg, J.H., E.N. Nilson, and J.L. Walsh. 1967. The Theory of Splines and Their Application. Academic Press, New York.

Barry, T.M. 1987. An Overview of Health Risk Analysis in the Environmental Protection Agency. In Proceedings of the Third Conference on Risk Analysis and Management of Natural and Manmade Hazards, eds. Y. Haimes and E. Stakhiv, American Society of Civil Engineering. Santa Barbara, CA. November.

Battese, G.E., S.M. Nusser, and W.A. Fuller. 1988. Estimation of Usual Intakes for Selected Dietary Components. CARD Technical Report 88-TR5, Center for Agricultural and Rural Development, Iowa State University, Ames. October.

Jensen, H.H., S.M. Nusser, W.A. Fuller, and S.R. Johnson. 1989. Estimating the Distribution of Usual Nutrient Intake. Paper presented at the AIN Fall Meeting on Nutrition Monitoring and Nutrition Status Assessment, Charleston, S.C., December 8-10.

Kramer, C.S. 1990. Food Safety: The Consumer Side of the Environmental Issue. National Center for Food and Agricultural Policy. Discussion Paper Series, No. FAP 90-06, March.

Lindgren, B.W. 1976. Statistical Theory. Macmillan, New York.

Mauskopf, J.A. 1990. Risk Characterization for Food Additives and Contaminants. Food Marketing Policy Center Research Report No. 8. Dept. of Agricultural Economics and Rural Sociology, The University of Connecticut.

Nusser, S.M. and W.A. Fuller. 1990. A Transformation Approach to Generating Pseudo Usual Intakes. Progress Report for the Human Nutrition and Information Service, USDA, for Research Support Agreement No. 58-3198--9-032.

Nusser, S.M., A.L. Carriquiry, H.H. Jensen, and W.A. Fuller. 1990. A Transformation Approach to Estimating Usual Intake Distribution. Proceedings of Second Annual Kansas State Conference on Statistics Applied to Agriculture (in press). Also available as CARD Working Paper 90-WP59, May 1990.

Pennington, J.A.T., and E.L. Gunderson. 1986. History of the Food and Drug Administration's Total Diet Study - 1961-1987. Journal of the Association of Official Analytic Chemists. 70:772-782.

Technical Assessment Systems, Inc. 1985. Chronic Dietary Exposure Analysis. Washington, D.C.

U.S. Department of Agriculture, Human Nutrition Information Service, Nutrition Monitoring Division. 1987. Women 19-50 Years and Their Children 1-5 Years, 4 Days. Nationwide Food Consumption Survey Continuing Survey of Food Intakes by Individuals. NFCS, CSFII: Report No. 85-4.

U.S. Environmental Protection Agency. 1987. The Risk Assessment Guidelines of 1986. Office of Health and Environmental Assessment, NTIS. EPA/8-87/045, Washington, D.C.

Figure 1. Estimates of the Percent of Population Whose Intake is Greater than 23 Units: