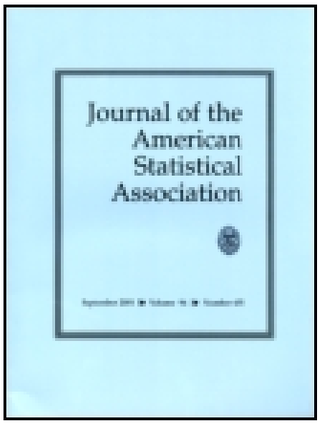


Journal of the American Statistical Association
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/uasa20

# Empirical Bayes Procedures for Stabilizing Maps of U.S. Cancer Mortality Rates

Kenneth G. Manton [a], Max A. Woodbury [a], Eric Stallard [a], Wilson B. Riggan [b], John P. Creason [c] & Alvin C. Pellom [d]

[a] Center for Demographic Studies, Duke University, Durham, NC, 27706, USA

[b] Biostatistical Branch, Biometry Division, Health Effects Research Laboratory, USA

[c] Biostatistical Branch, Health Effects Research Laboratory, U.S. Environmental Protection Agency (EPA), Research Triangle Park, NC, 27711, USA

[d] Integrated Laboratory Systems, Research Triangle Park, NC, 27709, USA

Published online: 12 Mar 2012.

To cite this article: Kenneth G. Manton, Max A. Woodbury, Eric Stallard, Wilson B. Riggan, John P. Creason & Alvin C. Pellom (1989) Empirical Bayes Procedures for Stabilizing Maps of U.S. Cancer Mortality Rates, Journal of the American Statistical Association, 84:407, 637-650

To link to this article: http://dx.doi.org/10.1080/01621459.1989.10478816

# Empirical Bayes Procedures for Stabilizing Maps of U.S. Cancer Mortality Rates

KENNETH G. MANTON, MAX A. WOODBURY, ERIC STALLARD, WILSON B. RIGGAN, JOHN P. CREASON, and ALVIN C. PELLOM*

The geographic mapping of age-standardized, cause-specific death rates is a powerful tool for identifying possible etiologic factors, because the spatial distribution of mortality risks can be examined for correlations with the spatial distribution of disease-specific risk factors. This article presents a two-stage empirical Bayes procedure for calculating age-standardized cancer death rates, for use in mapping, which are adjusted for the stochasticity of rates in small area populations. Using the adjusted rates helps isolate and identify spatial patterns in the rates. The model is applied to sex-specific data on U.S. county cancer mortality in the white population for 15 cancer sites for three decades: 1950–1959, 1960–1969, and 1970–1979. Selected results are presented as maps of county death rates for white males.

KEY WORDS: Age-standardized death rates; Negative binomial distribution.



## 1. INTRODUCTION

The mapping of age-standardized death rates (ASDR's) for cancers of specific anatomic sites is an important tool in assessing the environmental determinants of different types of cancer and in forming hypotheses about cancer risks for specific subpopulations to be tested in specially designed epidemiologic population studies (Mason, McKay, Hoover, Blot, and Fraumeni 1975). For example, maps of county and state economic area (SEA) cancer death rates have produced important insights about etiologic factors for lung cancer (i.e., occupational exposure to asbestos), cancers of the oral cavity (i.e., "smokeless" tobacco), and breast and cervical cancer (i.e., racial differences in sexual activity) (Blot, Fraumeni, Mason, and Hoover 1979; National Cancer Institute 1987).

An important stimulus for this research was the cancer mortality atlases prepared by Mason et al. (1975, 1976) for U.S. counties and SEA's for the 20-year period 1950–1969. Two subsequent publications have independently updated these atlases by extending the period covered from 1950–1969 to 1950–1979 and by presenting decade-specific maps for the three decades 1950–1959, 1960–1969, and 1970–1979 (Pickle, Mason, Howard, Hoover, and Fraumeni 1987; Riggan et al. 1987).

Pickle et al. (1987) applied a five-class choropleth map format, similar to that of Mason et al. (1975), to SEA maps for the white population, using a two-way classification rule based on (a) rank ordering of the SEA ASDR's and (b) tests of the significance of each ASDR vis-à-vis the U.S. ASDR.

Riggan et al. (1987) mapped ASDR's for counties and SEA's using a linearly ordered five-class choropleth map. These ASDR's were not based on the method of direct standardization (Fleiss 1973; Kitagawa 1955; Kleinman 1977). Instead, these ASDR's were estimated using a two-stage empirical Bayes (EB) procedure to stabilize rates by removing variation due to the small numbers of deaths in some counties.

To be comparable with Mason et al. (1975, 1976), Riggan et al. (1987) presented county maps for 15 of the most common sites for the white population and SEA maps for 34 sites for the nonwhite population and for the least common 19 sites for the white population. Maps for 4 of these 19 sites were deleted from the atlas because of the absence of significant variation in the SEA death rates (Riggan et al. 1987).

The purpose of this article is to describe the EB procedures used by Riggan et al. (1987) and to present the statistical analyses of the geographic variation of the 15 site-specific cancers that were mapped at the county level for the white population. These analyses update and extend the previous county analyses of Mason et al. (1975).

In the following five sections we present details of the application of the EB procedures to the mapping of cancer ASDR's for the United States during the three decades 1950–1959, 1960–1969, and 1970–1979. Section 2 introduces basic notation. Section 3 presents assumptions of the two-stage model. Section 4 describes the likelihood estimation procedures and shows how they yield the EB rate estimator as a weighted composite of the direct, indirect, and marginal ASDR's. Section 5 summarizes the county-level results for the 15 cancer sites, for the white population, for the three decades. Section 6 comments on the significance of those results.

## 2. PRELIMINARIES

Our data can be arranged as a 7-way table consisting of 2 sexes, 2 races (white and nonwhite), 3 decades (1950–1959, 1960–1969, and 1970–1979), 3,061 counties, 34 cancer sites, 18 age groups (0–4, 5–9, . . . , 80–84, and 85

* Kenneth G. Manton is Research Professor of Demography and Assistant Director, Center for Demographic Studies, Max A. Woodbury is Professor Emeritus, Center for Demographic Studies, and Eric Stallard is Statistical Research Analyst, Center for Demographic Studies, Duke University, Durham, NC 27706. Wilson B. Riggan is Mathematical Statistician, Biostatistical Branch, Biometry Division, Health Effects Research Laboratory, and John P. Creason is Chief, Biostatistical Branch, Health Effects Research Laboratory, U.S. Environmental Protection Agency (EPA), Research Triangle Park, NC 27711. Alvin C. Pellom is Systems Analyst, Integrated Laboratory Systems, Research Triangle Park, NC 27709. Manton's, Woodbury's, and Stallard's research was supported by National Institute on Aging Grant 5-R01-AG01159 and EPA Cooperative Agreement 811090-01-0. Special thanks go to Thomas Mason, U.S. National Cancer Institute, H. Dennis Tolley, Brigham Young University, and Burton Singer, Yale University, for reviewing drafts of this article. Any remaining errors are the responsibility of the authors.

years or older at last birthday), and a category indicating vital status (i.e., deaths and exposed-population counts). Because each map is sex, race, decade, and cancer-site specific, we suppress those dimensions and develop notation for two county × age-group tables defined by the numbers of deaths (denoted by $y_{ij}$) and person-years of exposure (denoted by $n_{ij}$), where $i = 1, \ldots, 3{,}061$ indexes counties and $j = 1, \ldots, 18$ indexes age groups.

The county is the lowest level of aggregation used by the U.S. National Center for Health Statistics in reporting mortality statistics. We use the term *county* throughout to designate a geographic mapping unit, but the method applies to any higher level of aggregation (e.g., SEA's) of the counties, as well as to county equivalents, such as parishes in Louisiana and independent cities in Virginia. Several adjustments were required to ensure consistent county definitions for the three decades (for details see Riggan et al. 1987).

Table 1 lists the principal symbols and their meanings. Cell $(i, j)$ refers to the cross-classification of county $i$ and age-group $j$, and $m_{ij}$ is the observed death rate in that cell; that is,

$$m_{ij} = y_{ij}/n_{ij}. \tag{2.1}$$

To be precise, $y_{ij}$ is the total number of deaths in cell $(i, j)$ in the specified decade and $n_{ij}$ is the sum of 10 annual estimates of midyear population sizes derived from adjacent censuses. Each component of $n_{ij}$ is a midyear population count; hence, $y_{ij}$ includes deaths to some persons not counted as part of $n_{ij}$. Under the assumption that $n_{ij}$ is known, it can be shown that $y_{ij}$ is approximately Poisson $(\mu_{ij}n_{ij})$, where $\mu_{ij}$ is the true death rate in cell $(i, j)$. The adequacy of this Poisson approximation is assessed in Manton and Stallard (1984, pp. 285–287), Manton, Woodbury, Stallard, and Dowd (1986, p. 185), and Brillinger (1986; including discussion by Manton). Under the most extreme assumptions, the approximation is no worse than the standard binomial approximation, and under more reasonable assumptions the approximation is substantially better. For example, Manton and Stallard (1984) shows that the relative error in the Poisson variance is less than 7 parts in 10,000 when the death rate is 5% or less, and less than 3 parts in 100,000 when the death rate is 1% or less. Hence, with the highest age-specific death rates for the combination of all cancer sites on the order of 2%, the Poisson approximation is acceptable for analysis.

Under this approximation, the maximum likelihood estimator (MLE) of $\mu_{ij}$ is $m_{ij}$. Using + to denote a marginal count and * to denote a marginal rate (see Table 1), it follows that

$$y_{i+} \sim \text{Poisson}(\mu_{i\cdot}n_{i+}), \tag{2.2}$$

where the MLE of $\mu_{i\cdot}$ is $m_{i\cdot}$,

$$m_{i\cdot} = y_{i+}/n_{i+}, \tag{2.3}$$

and $m_{i\cdot}$ is the marginal county-specific death rate in county $i$. Similarly,

$$y_{+j} \sim \text{Poisson}(\mu_{\cdot j}n_{+j}), \tag{2.4}$$

where the MLE of $\mu_{\cdot j}$ is $m_{\cdot j}$,

$$m_{\cdot j} = y_{+j}/n_{+j}, \tag{2.5}$$

and $m_{\cdot j}$ is the marginal age-specific death rate in age-group $j$.

We use $\{m_{\cdot j}\}$ to denote a set of age-specific death rates and refer to this set as a rate schedule. To denote the observed rate schedule for a specific county, say $i$, we use $\{m_{ij} \mid i\}$, where the vertical bar is followed by indexes that are assumed to be fixed inside the braces.

*Table 1. Principal Symbols and Their Meanings*

| Symbol | Meaning | First use |
|---|---|---|
| $N_j$ | Exposed population in age-group $j$ of the selected standard population | Equation (2.15) |
| $N_+$ | $\sum_{j=1}^{J} N_j$ = total standard population | Equation (2.15) |
| $n_{ij}$ | Exposed population in cell $(i, j)$ | Equation (2.1) |
| $n_{+j}$ | $\sum_{i=1}^{I} n_{ij}$ = exposed marginal population in age-group $j$ | Equation (2.4) |
| $n_{i+}$ | $\sum_{j=1}^{J} n_{ij}$ = exposed marginal population in county $i$ | Equation (2.2) |
| $y_{ij}$ | Number of deaths in cell $(i, j)$ | Equation (2.1) |
| $\bar{y}_{ij}$ | $n_{ij}m_{\cdot j} = E(y_{ij})$, under the null model | Section 2; also Equation (4.15) |
| $y_{+j}$ | $\sum_{i=1}^{I} y_{ij}$ = marginal number of deaths in age-group $j$ | Equation (2.4) |
| $y_{i+}$ | $\sum_{j=1}^{J} y_{ij}$ = marginal number of deaths in county $i$ | Equation (2.2) |
| $\bar{y}_{i+}$ | $\sum_{j=1}^{J} n_{ij}m_{\cdot j} = E(y_{i+})$, under the null model | Equation (2.13) |
| $y_{++}$ | $\sum_{i=1}^{I}\sum_{j=1}^{J} y_{ij}$ = total number of deaths in all cells | Equation (2.10b) |
| $\mu_{ij}$ | True death rate in cell $(i, j)$ | Section 2; also Equation (3.2) |
| $\tilde{\mu}_{ij}$ | Bayes estimator of $\mu_{ij}$ | Equation (3.4) |
| $\mu_{\cdot j}$ | $\sum_{i=1}^{I} n_{ij}\mu_{ij}/n_{+j}$ = true marginal death rate in age-group $j$ | Equation (2.4) |
| $\mu_{i\cdot}$ | $\sum_{j=1}^{J} n_{ij}\mu_{ij}/n_{i+}$ = true marginal death rate in county $i$ | Equation (2.2) |
| $\bar{\mu}_{i\cdot}$ | $\sum_{j=1}^{J} n_{ij}\mu_{\cdot j}/n_{i+} = E(\mu_{i\cdot})$, under the null model | Equation (2.7) |
| $\rho_i$ | $E(\mu_{ij})/\mu_{\cdot j}$ = true SMR in county $i$, fixed over index $j$ | Equations (2.11) and (3.1) |
| $\tilde{\rho}_i$ | Bayes estimator of $\rho_i$ | Equation (3.9) |
| $m_{ij}$ | Observed death rate in cell $(i, j)$ | Equation (2.1) |
| $m_{\cdot j}$ | $y_{+j}/n_{+j}$ = observed marginal death rate in age-group $j$ | Equation (2.5) |
| $m_{i\cdot}$ | $y_{i+}/n_{i+}$ = observed marginal death rate in county $i$ | Equation (2.3) |
| $\bar{m}_{i\cdot}$ | $\sum_{j=1}^{J} n_{ij}m_{\cdot j}/n_{i+}$ = estimator of $\mu_{i\cdot}$, under the null model | Equation (2.9) |
| $r_i$ | $m_{i\cdot}/\bar{m}_{i\cdot}$ = estimator of $\rho_i$ under the SMR model | Equation (2.12) |

### 2.1 Null Model

We define the null hypothesis of county uniformity of age-specific death rates as

$$\mu_{ij} = \mu_{\cdot j} \quad \text{for all } i \text{ and } j. \tag{2.6}$$

We refer to (2.6) as the null model. Under this null model

$$y_{i+} \sim \text{Poisson}(\bar{\mu}_{i\cdot} n_{i+}), \tag{2.7}$$

where $\bar{\mu}_{i\cdot}$ is the expected death rate in county $i$,

$$\bar{\mu}_{i\cdot} = \sum_{j=1}^{J} n_{ij}\mu_{\cdot j}/n_{i+}, \tag{2.8}$$

$J$ is the number of age groups, and the MLE of $\bar{\mu}_{i\cdot}$, denoted by $\bar{m}_{i\cdot}$, is obtained by replacing $\mu_{\cdot j}$ in (2.8) with its MLE, $m_{\cdot j}$, as defined in (2.5). This yields

$$\bar{m}_{i\cdot} = \sum_{j=1}^{J} n_{ij}m_{\cdot j}/n_{i+}. \tag{2.9}$$

The coefficient of variation (CV) of $\bar{m}_{i\cdot}$ is

$$\text{CV}(\bar{m}_{i\cdot}) = \left(\sum_{j=1}^{J} n_{ij}^2\mu_{\cdot j}/n_{+j}\right)^{1/2} \Big/ n_{i+}\bar{\mu}_{i\cdot} \tag{2.10a}$$

$$\approx 1/y_{++}^{1/2}, \tag{2.10b}$$

assuming that the county's age structure is approximately proportional to the marginal age structure $\{n_{+j}\}$.

Since $y_{++}$ is approximately 1.1–1.8 million deaths per decade for white males and 1.0–1.5 million for white females, $\text{CV}(\bar{m}_{i\cdot})$ is on the order of .1% for total cancer mortality. For the least frequent of the 15 site-specific cancers, $\text{CV}(\bar{m}_{i\cdot})$ is still only about 1%. Hence, it is reasonable to substitute the estimator $\bar{m}_{i\cdot}$ for the parameter $\bar{\mu}_{i\cdot}$ in certain models that follow in this section, to reduce the size of the estimation problem. This is readily accomplished by replacing $\mu_{\cdot j}$ in (2.8) with $m_{\cdot j}$ [see (2.9)] in the following development.

### 2.2 Standardized Mortality Ratio Model

An important alternative to the null model is the standardized mortality ratio (SMR) model (Breslow and Day 1975; Gail 1978). The SMR model is a variant of the proportional hazards model in which the baseline hazard rates are assumed known. Given that $\{m_{\cdot j}\}$ can be used to replace $\{\mu_{\cdot j}\}$ as discussed before, we define an SMR model for cancer death rates using

$$\mu_{ij} = \rho_i\mu_{\cdot j}, \qquad \text{assuming that } \mu_{\cdot j} \equiv m_{\cdot j}. \tag{2.11}$$

It follows from Breslow and Day (1975) that $\mu_{\cdot j} \equiv m_{\cdot j}$ implies that the MLE of $\rho_i$ is

$$r_i = m_{i\cdot}/\bar{m}_{i\cdot}. \tag{2.12}$$

We refer to (2.11) as the SMR model, and to $r_i$ in (2.12) as the observed SMR. Defining $\bar{y}_{i+}$ as

$$\bar{y}_{i+} = \sum_{j=1}^{J} n_{ij}m_{\cdot j}, \tag{2.13}$$

one can use (2.9) to rewrite (2.12) in the form

$$r_i = y_{i+}/\bar{y}_{i+}, \tag{2.14}$$

where $\bar{y}_{i+}$ is the expectation of $y_{i+}$ under the null model with $\mu_{\cdot j} \equiv m_{\cdot j}$. In the following, the overbars on $\bar{\mu}_{i\cdot}$, $\bar{m}_{i\cdot}$, $\bar{y}_{i+}$, and $\bar{y}_{ij}$ (see Table 1) refer to the null model (2.6) with $\mu_{\cdot j}$ replaced by $m_{\cdot j}$. A caret (^) will be used for all estimators except for Bayes estimators, which are indicated by a tilde (~).

### 2.3 Age-Standardized Death Rates

ASDR's are weighted averages of age-specific death rates, where the weights are independent of the county. Hence, they satisfy the general formula

$$\text{ASDR}(\{\mu_{ij} \mid i\}) = \sum_{j=1}^{J} N_j\mu_{ij} \Big/ \sum_{j=1}^{J} N_j = \sum_{j=1}^{J} \frac{N_j}{N_+}\mu_{ij}, \tag{2.15}$$

where the weights $N_j$ ($j = 1, \ldots, J$) are nonnegative. We refer to $\{N_j\}$ as the standard population.

It is conventional to select as the standard population either (a) the marginal population derived from the aggregate of the counties under analysis or (b) an external standard whose age structure is similar to the marginal population (Kitagawa 1955). For this article we selected the 1970 U.S. population as the standard that is close to the marginal population for the two decades 1960–1969 and 1970–1979. Other studies have used the 1940 (e.g., Klebba 1982) or 1960 (e.g., Mason et al 1975; Pickle et al. 1987) U.S. population as a standard. Thus ASDR's obtained from the different studies may not be fully comparable.

## 3. MODEL ASSUMPTIONS

The two stages of our EB procedure consist of the following:

1. generation of the model-based empirical Bayes estimator (EBE) of $\rho_i$ in (4.12) (see Sec. 4), which compromises between the SMR estimator, $r_i$, in (2.12) and the unit estimator of the null model in (2.6)
2. testing of the goodness of fit of the first-stage EBE's using methods proposed by Collings and Margolin (1985) for detecting negative binomial departures from the basic Poisson model of $y_{ij}$.

Thus the first-stage model is an EB generalization of the proportional hazards SMR model in which the EBE's of $\rho_i$ are "close" to $r_i$ in large counties, but shrink toward the mean value (unity) in small counties where the expected death counts are small. The second-stage model is required because standard methods of testing goodness of fit are inappropriate for sparse contingency tables of the type we are analyzing, where the expected cell counts frequently are less than .1 deaths per cell (Haberman 1988).

Collings and Margolin (1985, p. 411) emphasized that the detection of extra-Poisson variation in the second stage signals the need to develop a scientifically credible alternative model. This is important and, for the one cancer site where the test is significant (lung cancer; see Sec. 5),

we cite evidence from auxiliary analyses to support an alternative model. For the purpose of developing rate estimators for mapping, however, we use a second-stage EBE that compromises between the MLE for cell $(i, j)$, $m_{ij}$, in (2.1), and the EBE, $\hat{\rho}_i m_{\cdot j}$, of the first stage. This use of EB principles to adjust for lack of fit of the first-stage estimates was discussed by Morris (1983a).

### 3.1 Assumptions

*Assumption 1.* The parameter $\rho_i$ is independently and identically distributed over counties $\{i\}$ with gamma density $h(\rho_i)$, where

$$h(\rho_i) = \left(\frac{\rho_i}{\beta}\right)^{\gamma} \frac{e^{-\rho_i/\beta}}{\rho_i \Gamma(\gamma)}, \qquad 0 < \gamma \equiv \beta^{-1}. \tag{3.1}$$

*Assumption 2.* Given $\{\rho_i\}$ and $\{n_{ij}\}$, the parameters $\{\mu_{ij}\}$ are independently distributed with gamma density $g(\mu_{ij} \mid \rho_i)$, where

$$g(\mu_{ij} \mid \rho_i) = \left(\frac{n_{ij}\mu_{ij}}{\alpha}\right)^{n_{ij}\rho_i\mu_{\cdot j}/\alpha} \frac{e^{-n_{ij}\mu_{ij}/\alpha}}{\mu_{ij}\Gamma(n_{ij}\rho_i\mu_{\cdot j}/\alpha)},$$

$$\mu_{\cdot j} \equiv m_{\cdot j}, \quad 0 \le \alpha, \rho_i, \tag{3.2}$$

which is independent of all elements of $\{\rho_i\}$ and $\{n_{ij}\}$ except $\rho_i$ and $n_{ij}$.

*Assumption 3.* Given $\{\rho_i\}$, $\{\mu_{ij}\}$, and $\{n_{ij}\}$, the death counts $\{y_{ij}\}$ are independently distributed with Poisson probability function $P(y_{ij} \mid \mu_{ij})$, where

$$P(y_{ij} \mid \mu_{ij}) = (n_{ij}\mu_{ij})^{y_{ij}} e^{-n_{ij}\mu_{ij}}/y_{ij}!, \qquad 0 \le \mu_{ij}, \tag{3.3}$$

which is independent of $\{\rho_i\}$ and all elements of $\{\mu_{ij}\}$ and $\{n_{ij}\}$ except $\mu_{ij}$ and $n_{ij}$.

### 3.2 Rationale

The rationale for the specific forms of Assumptions 1–3 are discussed here, along with the specific properties they impart to the model.

Assumption 3 represents the Poisson approximation to the probability of $y_{ij}$, given $\mu_{ij}$, and was justified following (2.1).

Assumption 2 implies that $\mu_{ij}$ is gamma distributed with mean $\rho_i\mu_{\cdot j}$ (as in the SMR model) and variance $(\mathrm{var}(\mu_{ij} \mid \rho_i))\alpha\rho_i\mu_{\cdot j}/n_{ij}$. Thus $\mathrm{var}(\mu_{ij} \mid \rho_i)$ is proportional to $\alpha$ and is 0 if the generalized SMR model holds. For $\alpha > 0$, $\mathrm{var}(\mu_{ij} \mid \rho_i)$ is proportional to $n_{ij}^{-1}$ and tends to 0 in large samples.

We employ a two-stage estimation procedure for the parameters of (3.1) and (3.2). In this case, Assumption 2 provides a more general probabilistic mechanism for the stochasticity of $y_{ij}$ than the usual Poisson model (e.g., Breslow and Day 1975) obtained by substituting the right side of (2.11) for $\mu_{ij}$ in (3.3). This use of the gamma-mixed Poisson as a model of the residual variation of $y_{ij}$, given the mean rate $\rho_i\mu_{\cdot j}$, was recommended by Collings and Margolin (1985). We extend their procedure, however, by using the (second-stage) EBE of $\mu_{ij}$ whenever $\hat{\alpha} > 0$.

Because the gamma density in (3.2) is conjugate to the Poisson distribution, our rate estimators are approximately "EB $G_2$-minimax" for expected squared error loss within the class of models for $g(\mu_{ij} \mid \rho_i)$ with the same means and variances (Morris 1983b, p. 525, theorem 5.5). The qualifier *approximately* is used for the general case where both the means and variances are unknown (Morris 1983a). In our model, however, the mean, $\rho_i\mu_{\cdot j}$, is "known" prior to the second-stage test of the parameter $\alpha$. Hence, only the variance is unknown. It is given by $\alpha\rho_i\mu_{\cdot j}/n_{ij}$ so that any reasonably accurate estimator of $\alpha$ is sufficient to ensure that the approximation is good and that the EBE of $\mu_{ij}$ is "robust."

Morris (1983b) referred to $n_{ij}$ as the convolution parameter because it specifies a family of gamma-distributed random variables that is closed under convolution. Under (3.2), the aggregate of any two counties, say $i_1$ and $i_2$, is characterized by distributions of $\{\mu_{ij} \mid i\}$ with convolution parameters $\{n_{ij} \mid i\}$ obtained by summation of each pair of age-specific component values, $n_{i_1 j}$ and $n_{i_2 j}$, where $n_{ij} = n_{i_1 j} + n_{i_2 j}$. Strictly speaking, this convolution closure property requires one to reparameterize the means, $\rho_{i_1}\mu_{\cdot j}$ and $\rho_{i_2}\mu_{\cdot j}$, to the nonparametric forms, $v_{i_1 j}$ and $v_{i_2 j}$, prior to convolution. The resulting mean $v_{ij}$ is a weighted average of $v_{i_1 j}$ and $v_{i_2 j}$, with weights proportional to $n_{i_1 j}$ and $n_{i_2 j}$. Upon parameterization of $v_{ij} \equiv \rho_i\mu_{\cdot j}$, one again obtains (3.2).

Since the counties themselves are aggregates of smaller areas (e.g., census tracts) we can reverse the convolution operation to characterize the distribution of $\{\mu_{ij} \mid i\}$ within census tracts by redefining the index set $\{i\}$. Thus the gamma distribution (3.2) is consistent with a variety of aggregation schemes ranging from census tracts to SEA's.

The convolution closure property in (3.2) also applies to aggregation of age-group categories. Thus, (3.2) is consistent with age categorizations other than the five-year groupings used in our analyses. It follows that the county marginal densities, $g(\mu_{i\cdot} \mid \rho_i)$, are also gamma with $n_{i+}$ and $\bar{\mu}_{i\cdot}$ replacing $n_{ij}$ and $\mu_{\cdot j}$ in (3.2).

An important aggregation property derives from convoluting $g(\mu_{i_1^*} \mid \rho_{i_1})$ and $g(\mu_{i_2^*} \mid \rho_{i_2})$ to obtain the county marginal density $g(\mu_{i^*} \mid \rho_i)$, for the aggregate of counties $i_1$ and $i_2$—namely, that $\rho_i$ is a weighted average of $\rho_{i_1}$ and $\rho_{i_2}$ with weights proportional to $\bar{y}_{i_1^+} = n_{i_1^+}\bar{\mu}_{i_1^*}$ and $\bar{y}_{i_2^+} = n_{i_2^+}\bar{\mu}_{i_2^*}$. In this case the convolution closure property applies to the parameterized form of the gamma marginal density.

Assumption 1 implies that $\rho_i$ is gamma distributed with mean 1 and variance $\beta$. This gamma distribution yields EBE's of $\rho_i$ that are reasonably consistent with geographic aggregation, in the sense that aggregating counties $i_1$ and $i_2$ yields estimators that are weighted averages of $\rho_{i_1}$ and $\rho_{i_2}$, with the weights approximately proportional to $\bar{y}_{i_1^+}$ and $\bar{y}_{i_2^+}$ [see Eq. (4.12)]. Equation (4.13) shows that the approximation is best for small counties and rare cancers, which are the usual motivations for aggregating mapping units.

This aggregation property justifies the side condition in (3.1). Specifically, the weighted mean of the observed SMR's $\{r_i\}$ is unity, for weights proportional to $\{\bar{y}_{i+}\}$. This condition holds for weights proportional to $\{\hat{B}_i\}$ in (4.13) for the general case, where (a) $\gamma \equiv \bar{\rho}/\beta$ in (3.1), (b) $\bar{\rho}$ represents the mean of $\{\rho_i\}$, and (c) $\bar{\rho}$ and $\beta$ are estimated

by ML using (4.8). This condition will hold approximately when $\gamma \equiv \beta^{-1}$ or, equivalently, $\bar{\rho} \equiv 1$ in (3.1). Therefore, since we have no a priori reason to expect that the mean of $\{\rho_i\}$, $\bar{\rho}$, is either greater or less than the weighted mean of $\{r_i\}$ (i.e., unity), it is reasonable to (a) prespecify the mean of $\rho_i$ as the value 1, by setting $\gamma \equiv \beta^{-1}$ in (3.1), and (b) estimate $\beta$ using ML, as in (4.8). This specification makes the parameters of $h(\rho_i)$ less dependent on the particular set of geographic mapping units and improves the efficiency of the estimator of $\beta$. In addition, with the mean of $\rho_i$ known, the gamma density is the (approximately) EB $G_2$-minimax choice for $h(\rho_i)$ for expected squared error loss, given that $P(y_{i+} \mid \rho_i)$ is the Poisson probability in (4.5). Thus the gamma density in Assumption 1 yields robust estimators of $\rho_i$.

Except for Assumption 1, our current model is similar to the model presented in Manton, Woodbury, and Stallard (1981). Though argued by Tsutakawa, Shoop, and Marienfeld (1985) to be inappropriate for analyses of cancer death rates in counties where no cancer deaths occur (i.e., $y_{i+} = 0$) because that model yields zero death-rate estimates for all age groups in such counties, this critique was not relevant to Manton et al. (1981) because all counties in that analysis had at least one cancer death. The parameter $\rho_i$ was one of several effects included and tested as components of a general regression function. Alternative specifications without county-specific effects, $\{\rho_i\}$, and with alternative temporal aggregations were investigated in Manton and Stallard (1981, 1984).

The issue of counties with no cancer deaths (and, hence, zero ASDR's) underlies the rationale for our two-stage procedure. Previously, two approaches were proposed to deal with zero ASDR's in mapping. First, one can use larger aggregates to eliminate the problem, as did Pickle et al. (1987), who selected SEA-level mapping units. Second, one can respecify the mean, $\rho_i \mu_{\cdot j}$, in (3.2) as a general function of observed covariates as in Manton et al. (1981), but not include county-specific effects.

For our application to mapping neither of these approaches seemed satisfactory because environmental causes of cancer are likely to be specific to location and time. Thus we wanted to avoid unnecessary aggregation that might dilute the effects of areal risk factors. Additionally, because there is currently no set of covariates that satisfactorily explains the geographic variation of each site-specific cancer, we did not specify regression functions. This required us to develop the two-stage strategy to deal with the instability of the MLE's of $\{\rho_i\}$, and the problem of observed zero ASDR's.

### 3.3 Bayes Estimators

Assumptions 1–3 parallel the assumptions used by Deely and Lindley (1981, p. 833) in formulating their Bayesian EB model. The differences are the definition of the index set $\{i\}$ used in Assumption 1 and the use of $h(\rho_i)$ in our EBE $\hat{\rho}_i$. Thus the Bayesian EB model is extended to a two-stage EB model with hyperparameters $\alpha$ and $\beta$. For later reference, we present the Bayes estimators of $\rho_i$ and $\mu_{ij}$.

From Assumptions 1–3 it follows that

$$\tilde{\mu}_{ij} = \int_0^\infty \mu_{ij} g(\mu_{ij} \mid \{y_{ij}\})\, d\mu_{ij}, \tag{3.4}$$

where [see Deely and Lindley 1981, p. 835, eq. (1.5)]

$$g(\mu_{ij} \mid \{y_{ij}\}) = \int_0^\infty g(\mu_{ij} \mid y_{ij}, \rho_i) h(\rho_i \mid \{y_{ij}\})\, d\rho_i. \tag{3.5}$$

The two terms of this integrand are

$$g(\mu_{ij} \mid y_{ij}, \rho_i) = \frac{P(y_{ij} \mid \mu_{ij}) g(\mu_{ij} \mid \rho_i)}{\int_0^\infty P(y_{ij} \mid \mu_{ij}) g(\mu_{ij} \mid \rho_i)\, d\mu_{ij}} \tag{3.6}$$

and

$$h(\rho_i \mid \{y_{ij}\}) = \frac{P(\{y_{ij} \mid i\} \mid \rho_i) h(\rho_i)}{\int_0^\infty P(\{y_{ij} \mid i\} \mid \rho_i) h(\rho_i)\, d\rho_i}, \tag{3.7}$$

where

$$P(\{y_{ij} \mid i\} \mid \rho_i) = \prod_{j=1}^{J} \int_0^\infty P(y_{ij} \mid \mu_{ij}) g(\mu_{ij} \mid \rho_i)\, d\mu_{ij}. \tag{3.8}$$

Hence the terms on the right sides of (3.6) and (3.7) derive from (3.1)–(3.3), and the Bayes estimator of $\mu_{ij}$ in (3.4) can be simplified to

$$\tilde{\mu}_{ij} = W m_{ij} + (1 - W) \tilde{\rho}_i \mu_{\cdot j}, \qquad \mu_{\cdot j} \equiv m_{\cdot j}, \tag{3.9}$$

where

$$W = \alpha/(1 + \alpha), \tag{3.10}$$

and $\tilde{\rho}_i$ is the Bayes estimator of $\rho_i$; that is,

$$\tilde{\rho}_i = \int_0^\infty \rho_i h(\rho_i \mid \{y_{ij}\})\, d\rho_i. \tag{3.11}$$

Naturally, one could take a fully Bayesian approach to our application by specifying additional prior distributions for parameters $\beta$ and $\alpha$ in (3.1) and (3.2). We did not do so, however, because our large sample size (i.e., $I = 3{,}061$ counties) ensures that our EB procedures are valid. In addition, Deely and Lindley (1981, p. 840) warned that for such large samples the use of the Bayesian approach may actually lead to a degradation of the performance characteristics of the estimators.

## 4. TWO–STAGE EMPIRICAL BAYES ESTIMATION

The two-stage EB procedure consists of a model estimation stage and a model testing stage. Nonetheless, because ML is used in both stages, each stage involves both estimation and testing of a specific parameter, $\alpha$ or $\beta$. In addition, each stage involves the generation of EBE's of other parameters, $\{\rho_i\}$ and $\{\mu_{ij}\}$. The procedure is distinct from conditional likelihood estimation (Kalbfleisch and Sprott 1970), however, because the second-stage MLE of $\alpha$ is conditional on the EBE's of $\{\rho_i\}$, rather than on the MLE of $\beta$.

In the first stage, the significance of between-county rate

differences is tested using a negative binomial model of $y_{i+}$, where the extra-Poisson variance is indexed by $\beta$ and where $\alpha \equiv 0$ by assumption. If $\beta$ is not significantly different from 0, then there is no significant between-county variation to map. If $\beta$ is significantly different from 0, then there are significant between-county differences. These differences are represented in further calculations using the EBE's of $\{\rho_i\}$ defined in (4.12). At the end of the first stage we have estimated a parsimonious model of between-county variance in death rates that is consistent with, but more general than, the SMR model.

In the second stage, we test the adequacy of this specification, conditional on the EBE's of $\{\rho_i\}$, using a negative binomial model of $y_{ij}$, where the extra-Poisson variance is indexed by $\alpha$. If $\alpha$ is significantly different from 0, then we have evidence of the failure of the generalized SMR model. In any case, when $\hat{\alpha} > 0$ the additional variance is represented through $\hat{W}$ in the EBE's of $\mu_{ij}$ defined in (4.16).

### 4.1 First Stage

We begin by considering the doubly integrated likelihood [i.e., with $\mu_{ij}$ and $\rho_i$ integrated out of (3.3)]. This likelihood function is

$$L_{\alpha,\beta} = P(\{y_{ij}\}) = \prod_{i=1}^{I} P(\{y_{ij} \mid i\}), \tag{4.1}$$

where $I$ is the number of counties,

$$P(\{y_{ij} \mid i\}) = \int_0^\infty P(\{y_{ij} \mid i\} \mid \rho_i) h(\rho_i)\, d\rho_i, \tag{4.2}$$

and the right side of (4.2) is the denominator of (3.7). Thus the likelihood is defined, but it requires extensive computation. Strategies based on profile likelihood techniques (Richards 1961) that reduce the computational burden for the simultaneous estimation of $\alpha$ and $\beta$ were discussed by Tsutakawa (1988). Additional reductions in computational effort can be achieved for our model of $\beta$ by the use of marginal likelihood estimation instead of the full ML solution (Kalbfleisch and Sprott 1970).

This approach is based on the following algebraic rearrangement of (4.2):

$$P(\{y_{ij} \mid i\}) = P(\{y_{i+}) \int_0^\infty P(\{y_{ij} \mid i\} \mid y_{i+}, \rho_i) h(\rho_i \mid \{y_{i+}\})\, d\rho_i, \tag{4.3}$$

where

$$h(\rho_i \mid \{y_{i+}\}) = \frac{P(y_{i+} \mid \rho_i) h(\rho_i)}{\int_0^\infty P(y_{i+} \mid \rho_i) h(\rho_i)\, d\rho_i} \tag{4.4a}$$

$$= P(y_{i+} \mid \rho_i) h(\rho_i) / P(y_{i+}). \tag{4.4b}$$

It follows from Assumptions 2 and 3 (jointly), with $\mu_{\cdot j} \equiv m_{\cdot j}$ and $\alpha \equiv 0$, that

$$P(y_{i+} \mid \rho_i) = (\bar{y}_{i+}\rho_i)^{y_{i+}} e^{-\bar{y}_{i+}\rho_i} / y_{i+}!, \tag{4.5}$$

a Poisson conditional probability of $y_{i+}$, given $\rho_i$, with mean and variance $\bar{y}_{i+}\rho_i$. The marginal probability of $y_{i+}$ is the integral in the denominator of (4.4a), denoted by $P(y_{i+})$ in (4.4b):

$$P(y_{i+}) = \left(\frac{1}{1 + \bar{y}_{i+}\beta}\right)^{\beta^{-1}} \frac{\Gamma(\beta^{-1} + y_{i+})}{\Gamma(\beta^{-1}) y_{i+}!} \left(\frac{\bar{y}_{i+}\beta}{1 + \bar{y}_{i+}\beta}\right)^{y_{i+}}, \tag{4.6}$$

a negative binomial probability with mean $\bar{y}_{i+}$ and variance $(1 + \bar{y}_{i+}\beta)\bar{y}_{i+}$. This variance is larger than in the null model (2.6) by the factor $(1 + \bar{y}_{i+}\beta)$.

With $\alpha \equiv 0$, the integrand in (4.3) simplifies to a multinomial probability; that is,

$$P(\{y_{ij} \mid i\} \mid y_{i+}, \rho_i) = \frac{y_{i+}!}{\prod_j y_{ij}!} \prod_j \left(\frac{n_{ij} m_{\cdot j}}{n_{i+} \bar{m}_{\cdot \cdot}}\right)^{y_{ij}}, \tag{4.7}$$

which is independent of $\{\rho_i\}$ and $\beta$. Hence, we can ignore the integral in (4.3) in estimating $\beta$.

The likelihood of $\beta$, given $\alpha \equiv 0$, is obtained from (4.1) and the first term of (4.3) as

$$L_\beta = \prod_{i=1}^{I} P(y_{i+}), \tag{4.8}$$

where $P(y_{i+})$ is defined in (4.6). Collings and Margolin (1985, p. 417) showed that continuity of (4.6) [see their eq. (5)] at $\beta = 0$ occurs on the left as well as the right, so ML estimation of $\beta$ implies no special boundary problems. The MLE of $\beta$ is obtained from (4.8) using a Newton–Raphson algorithm detailed in Manton and Stallard (1981). Tests of the hypothesis $\beta = 0$ versus the alternative $\beta > 0$ use the standard large-sample chi-squared approximation to the distribution of the likelihood ratio based on (4.6) versus the constrained likelihood with $\beta \equiv 0$. The constrained likelihood is defined by replacing $P(y_{i+})$ in (4.8) with $P(y_{i+} \mid \rho_i \equiv 1)$ in (4.5).

It follows from (4.7) that the conditional density of $\rho_i$, given $\{y_{ij}\}$, can be replaced by the conditional density of $\rho_i$, given $\{y_{i+}\}$. Hence, (3.11) can be replaced by

$$\tilde{\rho}_i = \int_0^\infty \rho_i h(\rho_i \mid \{y_{i+}\})\, d\rho_i, \tag{4.9}$$

where $h(\rho_i \mid \{y_{i+}\})$ is defined in (4.4). The Bayes estimator of $\rho_i$ is $\tilde{\rho}_i$. The MLE of $\tilde{\rho}_i$ is obtained by replacing the implicit parameter $\beta$ in (4.9) with its MLE, $\hat{\beta}$. The EBE of $\rho_i$, denoted by $\hat{\rho}_i$, is defined as the MLE of $\tilde{\rho}_i$ and is likewise obtained from (4.9). In view of (4.5), it follows that

$$h(\rho_i \mid \{y_{i+}\}) = \left(\frac{\bar{y}_{i+}\rho_i}{B_i}\right)^{\beta^{-1} + y_{i+}} \frac{e^{-\bar{y}_{i+}\rho_i / B_i}}{\rho_i \Gamma(\beta^{-1} + y_{i+})}, \tag{4.10}$$

where

$$B_i = \bar{y}_{i+}\beta / (1 + \bar{y}_{i+}\beta). \tag{4.11}$$

When (4.10) is used to replace $h(\rho_i \mid \{y_{i+}\})$ in (4.9), the EBE of $\rho_i$ is

$$\hat{\rho}_i = 1 - \hat{B}_i + \hat{B}_i r_i, \tag{4.12}$$

where $\hat{B}_i$ is the MLE of $B_i$ in (4.11); that is,

$$\hat{B}_i = \bar{y}_{i+}\hat{\beta} / (1 + \bar{y}_{i+}\hat{\beta}). \tag{4.13}$$

### 4.2 Second Stage

The second stage exploits the observation that $\hat{\rho}_i$ in the first stage is both the MLE of $\tilde{\rho}_i$ and the EBE of $\rho_i$. The first property permits $\hat{\rho}_i$ to replace $\tilde{\rho}_i$ in (3.9) prior to consideration of an estimator of $W \equiv W(\alpha)$. The second property permits $\hat{\rho}_i$ to replace $\rho_i$ in (3.2), with $\alpha$ estimated conditionally on $\rho_i \equiv \hat{\rho}_i$, all $i$.

The likelihood of $\alpha$, given $\{\hat{\rho}_i\}$, is

$$L_\alpha = P(\{y_{ij}\} \mid \{\rho_i\}) = \prod_{i=1}^{I} \prod_{j=1}^{J} P(y_{ij} \mid \rho_i), \qquad \rho_i \equiv \hat{\rho}_i, \quad \text{all } i, \quad (4.14)$$

where

$$P(y_{ij} \mid \rho_i) = \left(\frac{1}{1+\alpha}\right)^{\bar{y}_{ij}\rho_i/\alpha} \frac{\Gamma(\bar{y}_{ij}\rho_i/\alpha + y_{ij})}{\Gamma(\bar{y}_{ij}\rho_i/\alpha)y_{ij}!} \left(\frac{\alpha}{1+\alpha}\right)^{y_{ij}}, \quad (4.15)$$

a negative binomial probability with mean $\bar{y}_{ij}\rho_i$ and variance $(1 + \alpha)\bar{y}_{ij}\rho_i$. The MLE of $\alpha$ is obtained from (4.14) using a Newton–Raphson algorithm detailed in Manton and Stallard (1981). Tests of the hypothesis $\alpha = 0$ versus the alternative $\alpha > 0$ use the standard large-sample chi-squared approximation to the distribution of the likelihood ratio based on (4.14) with $\alpha > 0$ versus the constrained likelihood with $\alpha \equiv 0$, where both likelihoods are evaluated with $\rho_i \equiv \hat{\rho}_i$. The constrained likelihood with $\alpha \equiv 0$ is defined by replacing $P(y_{ij} \mid \rho_i)$ in (4.14) with $P(y_{ij} \mid \mu_{ij} \equiv \rho_i m_{\cdot j})$ in (3.3), which is the limiting Poisson form of (4.15) as $\alpha \to 0^+$.

The EBE of $\mu_{ij}$ is obtained from (3.9) and (3.10) as the MLE of $\tilde{\mu}_{ij}$:

$$\hat{\mu}_{ij} = \hat{W} m_{ij} + (1 - \hat{W})\hat{\rho}_i m_{\cdot j}, \quad (4.16)$$

where

$$\hat{W} = \hat{\alpha}/(1 + \hat{\alpha}), \quad (4.17)$$

and $\hat{\rho}_i$ is obtained from (4.12).

### 4.3 EB Age-Standardized Death Rates

To obtain the ASDR for county $i$, we replace $\mu_{ij}$ in (2.15) with the expression for $\hat{\mu}_{ij}$ on the right side of (4.16). Using $\text{EBASDR}_i$ to denote the result, we use the fact that $\text{ASDR}(\cdot)$ is a linear functional on a suitably defined space of rate schedules, $\{m_{ij} \mid i\}$, to obtain

$$\text{EBASDR}_i = \hat{W}\,\text{DASDR}_i + (1 - \hat{W})\hat{B}_i\text{IASDR}_i + (1 - \hat{W})(1 - \hat{B}_i)\text{MASDR}, \quad (4.18)$$

where $\text{DASDR}_i$, $\text{IASDR}_i$, and MASDR denote, respectively, the direct-method (observed) ASDR, the indirect-method ASDR, and the marginal ASDR defined by

$$\text{DASDR}_i = \text{ASDR}(\{m_{ij} \mid i\}), \quad (4.19)$$

$$\text{IASDR}_i = \text{ASDR}(\{r_i m_{\cdot j} \mid i\}), \quad (4.20)$$

and

$$\text{MASDR} = \text{ASDR}(\{m_{\cdot j}\}). \quad (4.21)$$

Thus the EB rate in (4.18) is a weighted composite of the direct, indirect, and marginal ASDR's. In our application, the weight $W \equiv W(\alpha)$ is independent of $i$, but this is not essential. The weight $B_i$ depends on $i$ through the expected number of deaths, $\bar{y}_{i+}$.

Recall that the large-sample confidence intervals for $\beta$ are based on the assumption that $\hat{\beta}$ is asymptotically normally distributed with variance $\sigma^2_{\hat{\beta}}$ equal to the inverse of the expected Fisher information, based on (4.8). To generate EB confidence intervals, we alter this assumption slightly, following a suggestion by Morris (1983a, p. 51) (see also Laird and Louis 1987), by assuming that $\beta$ is asymptotically normally distributed with mean $\hat{\beta}$ and variance, $\hat{\sigma}^2_{\hat{\beta}}$, equal to the inverse of the observed Fisher information (Efron and Hinkley 1978). Thus $\beta \mid \{y_{ij}\} \sim \text{Normal}(\hat{\beta}, \hat{\sigma}^2_{\hat{\beta}})$. EB confidence intervals for $\beta$ can then be obtained by treating this distribution as a Bayes marginal posterior density whose mean is the MLE (i.e., $\hat{\beta}$) and whose variance is the squared standard error of the estimator.

In this case, the posterior distribution of $\rho_i$ in (3.7) [which is approximated by Eq. (4.10)] is assumed to refer to $\rho_i \mid \{y_{ij}\}, \beta$. EB confidence intervals for $\rho_i$ are similarly defined using the marginal posterior density of $\rho_i \mid \{y_{ij}\}$.

Similarly, we assume that asymptotically $\alpha \mid \{y_{ij}\}, \beta, \{\rho_i\} \sim \text{Normal}(\hat{\alpha}, \hat{\sigma}^2_{\hat{\alpha}})$, based on (4.14), and that the posterior distribution of $\mu_{ij}$ in (3.6) refers to $\mu_{ij} \mid \{y_{ij}\}, \beta, \{\rho_i\}, \alpha$. Again, EB confidence intervals for $\mu_{ij}$ and $\alpha$ are defined using the marginal posterior densities of $\mu_{ij} \mid \{y_{ij}\}$ and $\alpha \mid \{y_{ij}\}$.

For our mapping application, we require the marginal posterior density of $\text{ASDR}(\{\mu_{ij} \mid i\}) \mid \{y_{ij}\}$ for generating EB confidence intervals. It can be shown that the mean of this distribution is $\text{ASDR}(\{\hat{\mu}_{ij} \mid i\}) = \text{EBASDR}_i$, as defined in (4.18), and the asymptotic variance is

$$\text{var}[\text{ASDR}(\{\mu_{ij} \mid i\}) \mid \{y_{ij}\}] = \hat{\sigma}^2_{\text{EBASDR}_i} \quad (4.22a)$$

$$= \frac{\hat{W}}{n_{i+}}\text{EBTASDR}_i + [(1 - \hat{W})^2 + \hat{\sigma}^2_{\hat{W}}] \times (1 - \hat{B}_i)\hat{\beta}\hat{\rho}_i\text{MASDR}^2 + \hat{\sigma}^2_{\hat{W}}(\text{DASDR}_i - \hat{\rho}_i\text{MASDR})^2 + \hat{\sigma}^2_{\hat{B}_i}[(1 - \hat{W})^2 + \hat{\sigma}^2_{\hat{W}}](r_i - 1)^2\text{MASDR}^2, \quad (4.22b)$$

where

$$\text{EBTASDR}_i = \text{ASDR}(\{t_{ij}\hat{\mu}_{ij} \mid i\}), \quad (4.23)$$

$$t_{ij} = \frac{N_j/N_+}{n_{ij}/n_{i+}}, \quad (4.24)$$

$$\hat{\sigma}^2_{\hat{W}} = (1 - \hat{W})^4\hat{\sigma}^2_{\hat{\alpha}}, \quad (4.25)$$

and

$$\hat{\sigma}^2_{\hat{B}_i} = (1 - \hat{B}_i)^4\bar{y}^2_{i+}\hat{\sigma}^2_{\hat{\beta}}. \quad (4.26)$$

In our application, both $\beta$ and $\alpha$ are estimated with a high degree of precision (see Tables 2 and 3 in Sec. 5). Thus, the third and fourth terms of (4.22b) will be neg-

ligible. In addition, since $\hat{\alpha}$ (hence, $\hat{W}$) is either small or 0, the first term of (4.22b) will be negligible and the second term will yield

$$\hat{\sigma}^2_{\text{EBASDR}_i} \approx (1 - \hat{B}_i)\hat{\beta}\hat{\rho}_i\text{MASDR}^2 \quad (4.27a)$$

$$< \hat{\beta}\hat{\rho}_i\text{MASDR}^2, \quad (4.27b)$$

which implies that, for $\hat{\alpha} = 0$,

$$\text{CV}^2(\text{EBASDR}_i) < \hat{\beta}/\hat{\rho}_i \approx \hat{\beta}. \quad (4.28)$$

Thus $\hat{\beta}$ is an approximation to an upper bound of the squared coefficient of variation of the EB rates produced by our two-stage EB estimation procedure.

These results may be contrasted with similar results for the direct-method ASDR's in (4.19) used by Mason et al. (1975, 1976) and Pickle et al. (1987). It follows from the assumption $y_{ij} \sim \text{Poisson}(\mu_{ij}n_{ij})$ that

$$\hat{\sigma}^2_{\text{DASDR}_i} = \text{ASDR}(\{t_{ij}m_{ij} \mid i\})/n_{i+}, \quad (4.29)$$

which, on average, is larger than the first term of (4.22b) by the factor $\hat{W}^{-1} = 1 + \hat{\alpha}^{-1}$.

On the other hand, if one used the indirect-method ASDR's in (4.20), based on the SMR model, one would obtain

$$\hat{\sigma}^2_{\text{IASDR}_i} = (r_i/\bar{y}_{i+})\text{MASDR}^2. \quad (4.30)$$

This can be compared with the following revision of (4.27a):

$$\hat{\sigma}^2_{\text{EBASDR}_i} \approx [\hat{\rho}_i/(\bar{y}_{i+} + \hat{\beta}^{-1})]\text{MASDR}^2. \quad (4.31)$$

On average, (4.30) is larger than (4.31) by the factor $1 + (\bar{y}_{i+}\hat{\beta})^{-1}$.

For both sets of comparisons, the relative reductions in the squared standard errors of $\{\text{EBASDR}_i\}$ implied by the parameter estimates in Tables 2 and 3 are typically on the order of 10 to 100 or more, with the upper-bound effect identified in (4.28) responsible for the largest relative reductions. An important implication is that the standard errors of $\{\text{EBASDR}_i\}$ will tend to fall in a narrow range below the boundary, so comparisons of counties with similar values of $\text{EBASDR}_i$ (hence, similar values of $\hat{\rho}_i$) will have similar standard errors, regardless of the size of the expected numbers of deaths, $\bar{y}_{i+}$, in those counties. As a consequence, we did not evaluate (4.22b) in our results section.

## 5. RESULTS

We analyzed the between-county variation in 15 sets of site-specific cancer death rates for the 48 coterminous states of the United States and the Washington, D.C., area for the three-decade period 1950–1979. The analyses were independently conducted on a decade- and sex-specific basis for the white population. Nine types of cancer were analyzed: stomach, large intestine, rectum, liver/gallbladder, pancreas, lung, bladder, kidney/ureter, and brain. Five sex-specific types of cancer were also analyzed: prostate, female breast, cervix uteri, other uterus, and ovary, along with a residual category designated "secondary sites." The sites selected for males account for 78%–79% of total cancer mortality; the sites for females, 86%–87%.

### 5.1 Parameter Estimates

Tables 2 and 3 contain parameter estimates and test statistics for white males and white females, respectively. The MLE's of $\beta$ were obtained by substituting the negative binomial function (4.6) for (4.8). The MLE's of $\alpha$ were obtained by substituting the negative binomial function (4.15) for (4.14). The test statistics in Tables 2 and 3 are $z$ scores obtained by taking the square roots of the chi-squared statistics derived from the likelihood ratio tests of the hypotheses $\beta = 0$ versus $\beta > 0$ and of $\alpha = 0$ versus $\alpha > 0$, respectively. Also in Tables 2 and 3 are the MASDR's for the site-specific and total cancers. These rates appear in Formula (4.18).

In Table 2 the $z(\beta)$ scores are all highly significant [using one-tailed significance levels (Lawless 1987)], ranging from 8.44 for kidney/ureter cancer in 1970–1979 to 121.41 for lung cancer in 1960–1969. Similar results are found in Table 3 with $z(\beta)$ ranging from 5.59 for kidney/ureter cancer in 1970–1979 to 81.12 for female breast cancer in 1960–1969. Hence, there is significant statistical variation between counties for all 15 cancer sites selected for mapping.

There is also significant variation of MASDR's between decades in Tables 2 and 3. The statistical variation of the marginal rates is negligible (CV of .1%–1.0%), so we can be confident in these temporal changes. The increase in lung cancer MASDR's in Table 2 accounts for the increase in total cancer MASDR's for white males. Table 3 shows a decrease in total cancer MASDR's for white females, despite a tripling of their lung-cancer MASDR's.

Under Assumption 1, each $\beta$ can be interpreted as a relative variance of the $\{\rho_i\}$ in each decade. The estimates of $\beta$ for total cancer indicate a large temporal decrease in relative variance for white males and a small decrease for white females. Thus the relative variance for white males for 1970–1979 is actually slightly smaller than for white females (1.67% vs. 1.76%). For lung cancer the estimates of $\beta$ decrease substantially for white males (from 17.39% to 4.99%) but increase substantially for white females (from 3.54% to 9.42%).

A second important result shown in Tables 2 and 3 is that only two of the 75 site-specific estimates of $\alpha$ (i.e., white-male lung cancer in 1960–1969 and 1970–1979) have significant $z$ scores, using a .01 one-tailed significance level of 2.33 to indicate a failure of the generalized SMR model. This was confirmed in separate analyses that showed (a) a proportional hazards model for lung cancer fits cohort data from separate geographic areas for 1950–1978, (b) the estimated proportionality constants differ systematically between birth cohorts within geographic areas, and (c) the differences between areas in cohort trend lines for the proportionality constants are too large to permit a proportional hazards model to fit the cross-sectional data for all three decades (Manton, Stallard, Creason, and Riggan 1985). These cohort differentials also explain the increase in MASDR for lung cancer from 29.6 to 64.0 $\times$ $10^{-5}$ between 1950–1959 and 1970–1979 (Manton et al. 1985). Other analyses of cohort trends support the lack of

*Table 2. County-Level Parameter Estimates Obtained Under Two-Stage Maximum Likelihood: U.S. White Males, Site-Specific Cancers, 1950–1979*

| | $\beta \times 10^2$ | | | $\hat{\alpha} \times 10^2$ | | | $MASDR \times 10^5$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Site | 1950–1959 | 1960–1969 | 1970–1979 | 1950–1959 | 1960–1969 | 1970–1979 | 1950–1959 | 1960–1969 | 1970–1979 |
| Stomach | 6.99 | 7.44 | 7.05 | 0.27 | 0.00 | 0.00 | 20.2 | 13.2 | 9.0 |
| | 74.84[a] | 72.60 | 53.25 | 0.37[b] | 0.00 | 0.00 | | | |
| Large intestine | 12.90 | 9.26 | 5.14 | .05 | 0.00 | 0.00 | 17.1 | 18.3 | 20.0 |
| | 79.85 | 82.90 | 65.79 | .14 | 0.00 | 0.00 | | | |
| Rectum | 31.68 | 20.01 | 12.67 | 0.00 | 0.00 | 0.00 | 9.0 | 7.5 | 5.8 |
| | 82.52 | 70.99 | 50.65 | 0.00 | 0.00 | 0.00 | | | |
| Liver/gallbladder | 6.41 | 4.35 | 3.34 | .34 | 0.00 | 0.00 | 3.0 | 3.7 | 4.6 |
| | 21.23 | 24.54 | 22.96 | .42 | 0.00 | 0.00 | | | |
| Pancreas | 2.29 | 1.99 | 1.07 | 1.19 | 0.00 | 0.00 | 9.3 | 10.9 | 10.9 |
| | 22.82 | 30.79 | 16.10 | 1.48 | 0.00 | 0.00 | | | |
| Lung | 17.39 | 9.57 | 4.99 | 1.08 | 3.85 | 8.35 | 29.6 | 46.8 | 64.0 |
| | 117.44 | 121.41 | 106.59 | 1.47 | 4.97 | 10.52 | | | |
| Prostate | 1.61 | 1.55 | .95 | 0.00 | .22 | 1.55 | 20.7 | 19.7 | 20.3 |
| | 21.67 | 33.23 | 22.69 | 0.00 | .33 | 2.06 | | | |
| Kidney/ureter | 4.75 | 2.04 | 1.12 | 1.49 | 1.74 | .44 | 3.7 | 4.2 | 4.6 |
| | 18.73 | 17.56 | 8.44 | 1.71 | 2.24 | .66 | | | |
| Bladder | 17.18 | 9.05 | 4.71 | 0.00 | .29 | .06 | 7.4 | 7.3 | 7.3 |
| | 47.80 | 41.75 | 27.93 | 0.00 | .62 | .07 | | | |
| Brain | 2.09 | 2.23 | 1.42 | 0.00 | .39 | .09 | 4.0 | 4.4 | 4.9 |
| | 14.75 | 19.50 | 12.74 | 0.00 | .54 | .14 | | | |
| Secondary sites | 4.97 | 5.40 | 4.94 | 0.00 | 0.00 | 0.00 | 14.5 | 13.6 | 12.8 |
| | 49.14 | 59.91 | 51.62 | 0.00 | 0.00 | 0.00 | | | |
| Total | 4.83 | 2.93 | 1.67 | 11.22 | 10.66 | 12.78 | 176.6 | 190.0 | 204.1 |
| | 180.95 | 183.20 | 120.46 | 17.76 | 17.10 | 19.49 | | | |

[a] z score—$H_0$: $\beta = 0$.
[b] z score—$H_0$: $\alpha = 0$.



sizeable nonproportional effects (Manton and Stallard 1988).

The interpretation of $\alpha$ derives from the limiting Poisson form of (4.15) as $\alpha \to 0^+$—namely, $\alpha$ is the relative increase in the variance of $y_{ij}$ over that predicted by the Poisson model. For total cancer, the estimates of $\alpha$ indicate relative increases as large as 12.98% for white males and 5.52% for white females. These estimates are highly significant for both sex groups. For the 15 site-specific cancers, the maximum estimate of $\alpha$ for white females is 1.04%. Without lung cancer, the maximum for white males is 1.74%. Thus, except for lung cancer for white males, the generalized SMR model derived from the first-stage estimation of $\{\rho_i\}$ does an excellent job in fitting the site-specific data.

The modeling of the extra-Poisson variance identified in the second stage of the procedure serves three purposes in our application. First, it permits one to test the generalized SMR model under the assumption that an adequate fit is obtained if the residual variation of $y_{ij}$ is Poisson (i.e., $\alpha = 0$). In the case when $\hat{\alpha} = 0$, $\hat{W} = 0$ in (4.18) and the EBASDR is a weighted average of the IASDR and MASDR, but not the DASDR. Second, it alerts the analyst to additional effects in the data beyond those formally included in Assumptions 1–3. As already noted, our auxiliary analysis of lung cancer identified significant cohort effects that were not respresentable in a cross-sectional proportional hazards model. Third, it provides an adjustment to the rate estimator when the generalized SMR model is inadequate [i.e., use the estimator $\hat{\alpha}$ to generate $\hat{W} > 0$ for use in (4.16) and (4.18)]. As Morris (1983a) discussed, this adjustment represents a compromise that on average will outperform either the observed death rate, $m_{ij}$, or the model-based rate, $\hat{\rho}_i m_{\cdot j}$, in (4.16). In the case when $\hat{\alpha} > 0$ (regardless of whether the test of $\hat{\alpha}$ is significant), the EBASDR is a weighted average of the IASDR, MASDR, *and* DASDR. Thus the adjustment allows one to generate optimal estimates of the county ASDR's for the mapping application, without having to specify explicitly a credible alternative model for those cancer sites for which the generalized SMR model fails.

## 5.2 Illustrative Maps

The motivation for these cancer analyses was to parameterize (4.18) to obtain EB-rate estimates $\{\text{EBASDR}_i\}$ suitable for mapping. The choropleth maps in Figures 1 and 2 display two sets of ASDR's for kidney/ureter cancer for white males in 1970–1979. The rates in Figure 1 are direct age-standardized $\{\text{DASDR}_i\}$ and shaded from white to black with changes in tone at the 75th, 90th, 95th, and 98th percentiles. In the map, the western counties appear larger. Counties in the highest two percentiles appear to be clustered in a band from west Texas to North Dakota. Figure 2 displays the corresponding map based on the EB rates. These rates are independently shaded from white

*Table 3. County-Level Parameter Estimates Obtained Under Two-Stage Maximum Likelihood: U.S. White Females, Site-Specific Cancers, 1950–1979*

| Site | $\beta \times 10^2$ | | | $\hat{\alpha} \times 10^2$ | | | $MASDR \times 10^5$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1950–1959 | 1960–1969 | 1970–1979 | 1950–1959 | 1960–1969 | 1970–1979 | 1950–1959 | 1960–1969 | 1970–1979 |
| Stomach | 7.46 | 7.62 | 6.81 | 0.00 | 0.00 | .24 | 10.7 | 6.7 | 4.3 |
| | 59.91[a] | 57.63 | 46.36 | 0.00[b] | 0.00 | .31 | | | |
| Large intestine | 7.39 | 4.37 | 3.42 | 0.00 | 0.00 | 0.00 | 18.0 | 17.1 | 16.4 |
| | 65.04 | 67.50 | 54.16 | 0.00 | 0.00 | 0.00 | | | |
| Rectum | 12.88 | 10.89 | 8.87 | 0.00 | 0.00 | 0.00 | 6.0 | 4.6 | 3.4 |
| | 47.56 | 44.99 | 37.15 | 0.00 | 0.00 | 0.00 | | | |
| Liver/gallbladder | 6.25 | 5.17 | 4.34 | .39 | .35 | 0.00 | 4.2 | 3.9 | 3.8 |
| | 23.78 | 26.84 | 25.86 | 1.11 | .47 | 0.00 | | | |
| Pancreas | 2.44 | 1.85 | 1.72 | .42 | .42 | .36 | 5.9 | 6.5 | 6.7 |
| | 20.12 | 27.26 | 23.59 | .47 | .96 | .59 | | | |
| Lung | 3.54 | 6.54 | 9.42 | .56 | 0.00 | 0.00 | 5.1 | 7.6 | 15.3 |
| | 26.25 | 45.96 | 75.19 | .69 | 0.00 | 0.00 | | | |
| Breast | 6.88 | 4.84 | 4.08 | 0.00 | 0.00 | 0.00 | 26.3 | 26.4 | 27.0 |
| | 70.12 | 81.12 | 68.89 | 0.00 | 0.00 | 0.00 | | | |
| Cervix uteri | 7.90 | 9.13 | 7.78 | 0.00 | 0.00 | .72 | 9.1 | 6.9 | 4.2 |
| | 47.59 | 54.59 | 36.00 | 0.00 | 0.00 | 1.19 | | | |
| Other uterus | 6.00 | 4.06 | 3.13 | .49 | 0.00 | 0.00 | 7.8 | 5.4 | 4.3 |
| | 34.22 | 30.85 | 20.90 | 1.01 | 0.00 | 0.00 | | | |
| Ovary | 6.10 | 2.71 | 1.97 | .09 | .99 | 0.00 | 8.6 | 8.9 | 8.8 |
| | 33.31 | 35.80 | 24.24 | .12 | 1.30 | 0.00 | | | |
| Kidney/ureter | 1.61 | 1.72 | .78 | .84 | .66 | 0.00 | 2.0 | 2.1 | 2.1 |
| | 7.87 | 10.53 | 5.59 | 1.10 | 1.19 | 0.00 | | | |
| Bladder | 3.61 | 3.18 | 2.41 | .85 | 0.00 | 0.00 | 2.9 | 2.4 | 2.1 |
| | 14.07 | 17.29 | 11.63 | .96 | 0.00 | 0.00 | | | |
| Brain | 3.65 | 3.20 | 1.73 | 1.00 | 1.04 | .92 | 2.6 | 2.9 | 3.3 |
| | 15.17 | 19.21 | 13.51 | 1.46 | 1.61 | 1.37 | | | |
| Secondary sites | 4.42 | 4.73 | 4.61 | 0.00 | 0.00 | 0.00 | 13.5 | 11.1 | 10.1 |
| | 40.44 | 56.35 | 49.39 | 0.00 | 0.00 | 0.00 | | | |
| Total | 2.30 | 1.81 | 1.76 | 4.90 | 3.58 | 5.52 | 141.6 | 132.4 | 131.7 |
| | 114.67 | 146.36 | 111.96 | 7.73 | 5.65 | 8.82 | | | |

[a] z score—$H_0$: $\beta = 0$.
[b] z score—$H_0$: $\alpha = 0$.



to black, again with changes in tone at the 75th, 90th, 95th, and 98th percentiles.

The bar charts at the lower right of Figures 1 and 2 are identical and show the age-specific death rates from kidney/ureter cancer, ranging from $.7 \times 10^{-5}$ at ages 35–39 years to $33.2 \times 10^{-5}$ at ages 80–84 years. This rate schedule yields a MASDR of $4.6 \times 10^{-5}$ in 1970–1979, an increase from $4.2 \times 10^{-5}$ in 1960–1969 and $3.7 \times 10^{-5}$ in 1950–1959 (Table 2). These can be compared with the MASDR's for total cancer: $176.6 \times 10^{-5}$ in 1950–1959, $190.0 \times 10^{-5}$ in 1960–1969, and $204.1 \times 10^{-5}$ in 1970–1979. Thus kidney/ureter cancer accounts for 2% of all cancer deaths for white males in this period.

The frequency function at the lower left of Figure 2 shows that the EB rates in the 3,061 counties range from 3.75 to $5.58 \times 10^{-5}$. The range for the DASDR's in Figure 1 is from 0.00 to $178.59 \times 10^{-5}$. Given the MASDR of $204.1 \times 10^{-5}$ for total cancer, the entire set of high rates in Figure 1 in the range from 11.79 to $178.59 \times 10^{-5}$ is suspect because of the small number of deaths represented. This is confirmed in Figure 2, where the largest EB rate ($5.58 \times 10^{-5}$; see the tick marks below the frequency function) is below the 75th percentile of the DASDR's ($5.65 \times 10^{-5}$).

Kidney/ureter cancer is sufficiently rare that the stabilization procedure makes a difference, but not so rare that county maps only represent noise. The first-stage $z(\beta)$ score is 8.44 (Table 2)—strong evidence that the marginal rates do not fit the data for all 3,061 counties. On the other hand, the second-stage $z(\alpha)$ score of .66 indicates that the residual extra-Poisson variation after fitting the first-stage model is nonsignificant.

Figure 3 compares the rankings for $\{\text{DASDR}_i\}$ with rankings for $\{\text{EBASDR}_i\}$. The square in the lower left of Figure 3 shows that counties in the 98–99 percentile in Figure 1 and the 0–74 percentile in Figure 2 are coded white. Only one county (San Miguel County, Colorado) is in this category. Similarly, counties in the 0–74 percentile in Figure 1 and the 98–99 percentile in Figure 2 are coded black. Thirteen counties are in this category, including several counties in California, Connecticut, Indiana, New Jersey, and New York with large populations.

Light gray and dark gray indicate changes of two or three classes in the five-class ranking system. Light gray

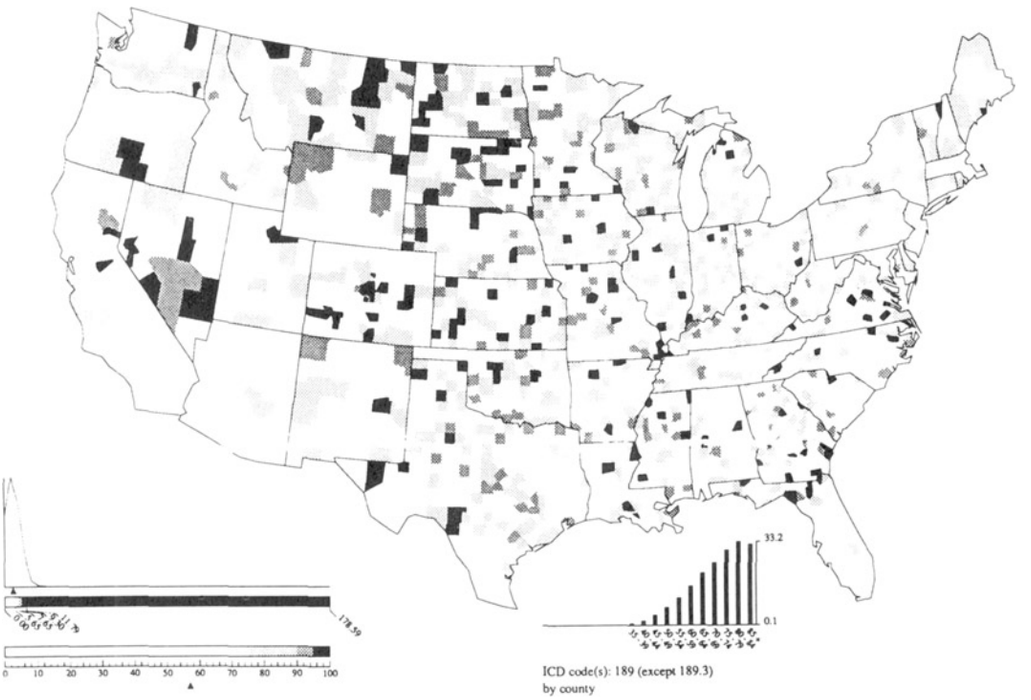


*Figure 1. Direct Age-Standardized Death Rates (DASDR's) for Cancer of Kidney/Ureter per 100,000 Exposed Population According to County: U.S. White Males, 1970–1979. The frequency function in the lower left of the figure is a graph of the unweighted frequencies of the 3,061 county-specific DASDR's. The first tone bar below the graph indicates the range of the distribution (in units of $10^{-5}$) and the locations of the 75th, 90th, 95th, and 98th percentiles, as defined on the second tone bar. The arrowheads below the graph and below the second tone bar indicate the location of the national death rate (MASDR). The bar graph in the lower right is a graph of the age-specific death rates (in units of $10^{-5}$) for ages 35–39 to 85 years and older.*



indicates lowered risk and appears primarily along the band from west Texas to North Dakota, and in Nevada and the deep south. Dark gray indicates increased risk and appears primarily in the north central, the northeast, the Missouri River area, and Texas.

This example illustrates three features of the EB approach. First, comparison of the frequency functions for the DASDR's versus EB rates shows substantial reductions in the range and variance of the distribution—especially the upper tail of the distribution. Second, the EB rates provide a good fit to the data with only two parameters (one in each stage) with the second-stage test showing that the hypothesis that the *variation in rates can be described by a proportional hazards model* cannot be rejected. Third, a substantial reordering of the counties occurs when the DASDR's are replaced by the EB rates. Given that the mapping problem involves comparing ASDR's between the 3,061 counties and that, taken as a set, the EB procedure yields the most accurate estimates of the 3,061 ASDR's (Berger 1983; Efron 1982), the map in Figure 2 is better than that in Figure 1 for identification of counties with high ASDR's.

## 6. CONCLUDING REMARKS

Although the DASDR and IASDR maps have proved useful for cancer epidemiology and environmental surveillance, there has been no accepted procedure for shifting from the DASDR's to the IASDR's because of instability of the age-specific death rates for small populations. For very small total county populations, the IASDR is also unstable.

The usual response to such instability is to aggregate the data, across either mapping units, time, or both. Such aggregation is undesirable if the effects to be identified are highly localized, for example, point-source pollutants. In the case of county mortality data, there is a wide range of population sizes and, hence, of levels of stability of site-specific cancer ASDR's. These data are thus mixtures of stable and unstable rates. This view is implicit in the mapping strategy used by Mason et al. (1975, 1976) and Pickle et al. (1987).

Frequently, unstable DASDR's and IASDR's are replaced with MASDR's derived from a national schedule of age-specific rates (Schaible, Brock, and Schnack 1977). This implies use of the null model in Equation (2.6). Age standardization of these marginal rate schedules is a special case of the indirect method with the SMR set to unity, by construction; compare Equations (4.20) and (4.21). Again, there is no general procedure for shifting from the IASDR's to the MASDR's because of instability of the indirect method for small populations.

We resolved the question of which ASDR's to use in the county maps by using a weighted composite of the three rates [Eq. (4.18)]. The weights are derived on a county-specific basis using a two-stage EB procedure that gives greater weight to the MASDR when the county population size is small and greater weight to the IASDR when the county population size is large [Eq. (4.13)]. The second

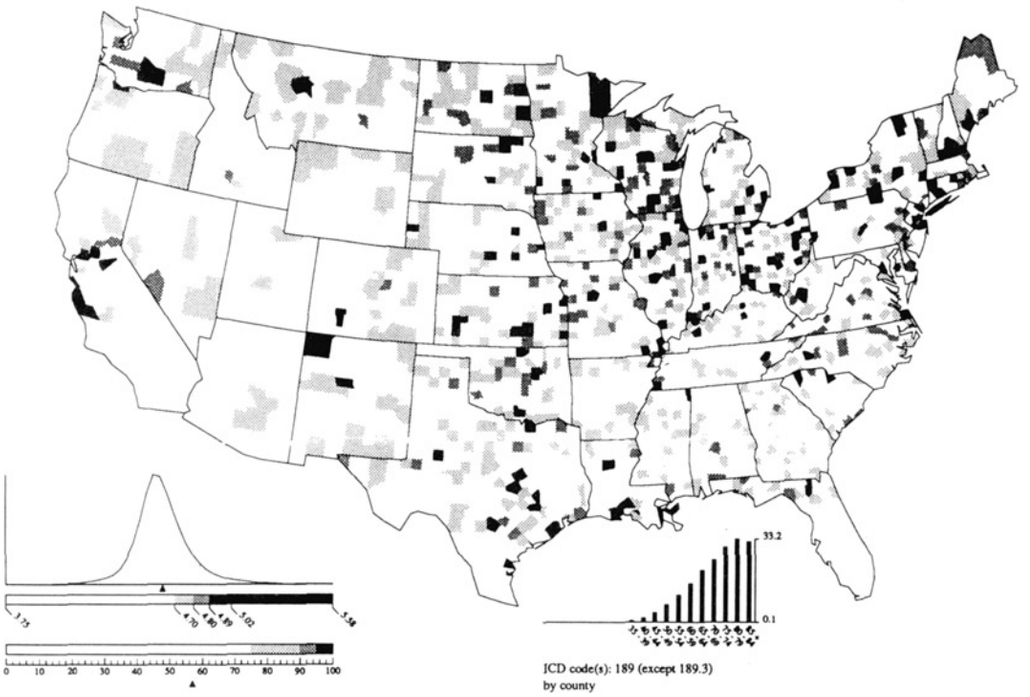


*Figure 2. Empirical Bayes Age-Standardized Death Rates (EBASDR's) for Cancer of Kidney/Ureter per 100,000 Exposed Population According to County: U.S. White Males, 1970–1979. The frequency function in the lower left of the figure is a graph of the unweighted frequencies of the 3,061 county-specific EBASDR's. The first tone bar below the graph indicates the range of the distribution (in units of $10^{-5}$) and the locations of the 75th, 90th, 95th, and 98th percentiles, as defined on the second tone bar. The arrowheads below the graph and below the second tone bar indicate the location of the national death rate (MASDR). The bar graph in the lower right is a graph of the age-specific death rates (in units of $10^{-5}$) for ages 35–39 to 85 years and older.*



stage weights-in the DASDR independent of population size according to the residual variation after fitting the proportional hazards model of the first stage [Eq. (4.17)].

The procedure is a two-stage application of EB principles (Morris 1983a). The particular model structure implied by Assumptions 1–3 has five advantages over other EB models (e.g., Gaver and O'Muircheartaigh 1987; Tsutakawa 1988; Tsutakawa et al. 1985). First, because the EB rate is a weighted composite of the three types of ASDR's, our results can be readily understood by cancer epidemiologists concerned about the instability of ASDR's. Second, because our proportional hazards assumption implies the estimation of county-specific proportionality parameters [Eq. (4.12)], our procedure generalizes the SMR model (Gail 1978) in analyzing environmental causes of cancer. In this regard, our approach is similar to that of Tsutakawa (1988).

Third, two sets of test statistics are provided to assess the fit of the model. (a) The variation in SMR's between counties is tested using the parameter $\beta$ in (3.1). If this parameter cannot be set to 0 without loss of fit, then the between-county cancer SMR variation is significant. This test was significant for all 15 cancer sites. (b) The adequacy of the proportional hazards assumption in the first stage is tested using the parameter $\alpha$ in (3.2). If this parameter can be set to 0 without loss of fit, then the residual extra-Poisson variation of the age-specific cancer death rates is not significant and the proportional hazards model is adequate. This test can reveal the presence of additional effects such as differential trends in cohort lung-cancer death rates (Manton et al. 1985) that may be hidden by age standardization.

The likelihood ratio test of $\alpha$ is conditional on $\{\hat{\rho}_i\}$ and, hence, is the most powerful test of the hypothesis $\alpha = 0$ versus the negative binomial alternative (4.15) (Collings and Margolin 1985). The second-stage EB estimation of $\{\mu_{ij}\}$ (i.e., with $\hat{W} > 0$) responds to the model's lack of fit, following Morris's (1983a) suggestion of a compromise estimator when the first-stage model does not hold.

The first-stage EB estimation of $\{\rho_i\}$ also follows Morris's (1983a) suggestion of a compromise estimator. The difference is simply that our prior expectations led us to believe that the first-stage null model would be rejected by the $z(\beta)$ test in virtually all cases (as it was), whereas we expected that the generalized SMR model would be accepted by the $z(\alpha)$ test in virtually all cases (as it was in 73 of 75 tests).

Fourth, in our analyses, the two-stage procedure yields computationally efficient closed-form expressions for the likelihoods and the posterior means. This may be contrasted with procedures (e.g., Gaver and O'Muircheartaigh 1987; Tsutakawa 1988) that involve numerical integration for both the likelihood evaluation and the posterior-means computations and that produce exact values of suboptimal estimates (based on the EB $G_2$-minimax criterion).

Fifth, the proposed model is logically closed under aggregation. Our model results are consistent whether counties or SEA's are used, whether 5-year or 10-year age categories are used, or whether data are aggregated or

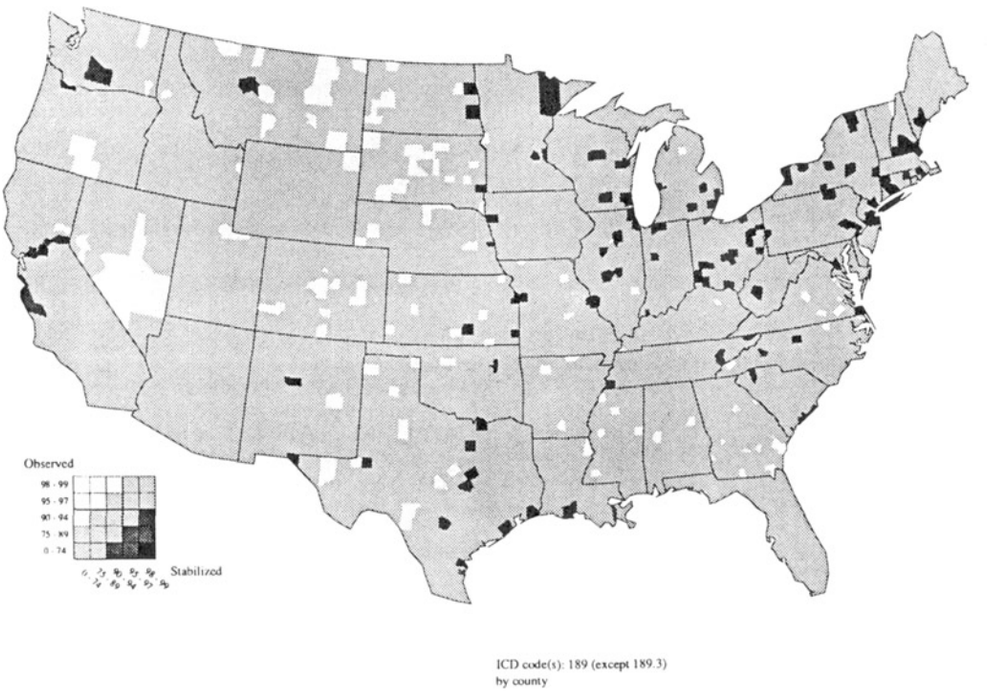


*Figure 3. Representation of the Joint Distribution of Two Sets of Age-Standardized Death Rates (ASDR's) for Cancer of Kidney/Ureter: U.S. White Males, 1970–1979. The observed ASDR's refer to the DASDR's mapped in Figure 1. The stabilized ASDR's refer to the EBASDR's mapped in Figure 2. The box in the lower left refers to the percentile classes defined in the lower tone bars of Figures 1 and 2. The tones in the box define the tones in the map. A correlation of 1.0 between the two sets of ASDR's would yield a map that is entirely middle gray in tone. Extreme tones (white or black) indicate extreme discordance between the two sets of rates.*



disaggregated over decades. Tsutakawa's (1988) model is not closed under aggregation, and the parameter estimates for different areal, temporal, age, and other demographic groupings will not be directly comparable. The fact that closure under aggregation coincides with the EB $G_2$-minimax property for the gamma-mixed Poisson distributions also means that our two-stage estimators yield good approximations to the optimal estimators.

Finally, the two-stage estimation and model testing procedure distinguishes our model from other models based on one-stage procedures. Such formal testing is essential in our application, because the substitution of the stabilized maps for the observed DASDR maps is justified only when it can be shown that the stabilized maps do not distort the data.

[*Received June 1986. Revised December 1988.*]

## REFERENCES

Berger, J. (1983), Comment on "Parametric Empirical Bayes Inference," by C. N. Morris, *Journal of the American Statistical Association,* 78, 55–57.

Blot, W. J., Fraumeni, J. F., Mason, T. J., and Hoover, R. (1979), "Developing Clues to Environmental Cancer: A Stepwise Approach With the Use of Cancer Mortality Data," *Environmental Health Perspectives,* 32, 53–58.

Breslow, N. E., and Day, N. E. (1975), "Indirect Standardization and Multiplicative Models for Rates, With Reference to the Age Adjustment of Cancer Incidence and Relative Frequency Data," *Journal of Chronic Diseases,* 28, 289–303.

Brillinger, D. R. (1986), "The Natural Variability of Vital Rates and Associated Statistics" (with discussion), *Biometrics,* 42, 693–734.

Collings, B. J., and Margolin, B. H. (1985), "Testing Goodness of Fit for the Poisson Assumption When Observations Are Not Identically Distributed," *Journal of the American Statistical Association,* 80, 411–418.

Deely, J. J., and Lindley, D. V. (1981), "Bayes Empirical Bayes," *Journal of the American Statistical Association,* 76, 833–841.

Efron, B. (1982), "Maximum Likelihood and Decision Theory," *The Annals of Statistics,* 10, 340–356.

Efron, B., and Hinkley, D. V. (1978), "Assessing the Accuracy of the Maximum Likelihood Estimator: Observed Versus Expected Fisher Information" (with discussion), *Biometrika,* 65, 457–487.

Fleiss, J. L. (1973), *Statistical Methods for Rates and Proportions,* New York: John Wiley.

Gail, M. (1978), "The Analysis of Heterogeneity for Indirect Standardized Mortality Ratios," *Journal of the Royal Statistical Society,* Ser. A, 141, 224–234.

Gaver, D. P., and O'Muircheartaigh, I. G. (1987), "Robust Empirical Bayes Analyses of Event Rates," *Technometrics,* 29, 1–15.

Haberman, S. J. (1988), "A Warning on the Use of Chi-Squared Statistics With Frequency Tables With Small Expected Cell Counts," *Journal of the American Statistical Association,* 83, 555–560.

Kalbfleisch, J. D., and Sprott, D. A. (1970), "Application of Likelihood Methods to Models Involving Large Numbers of Parameters" (with discussion), *Journal of the Royal Statistical Society,* Ser. B, 32, 175–208.

Kitagawa, E. M. (1955), "Components of a Difference Between Two Rates," *Journal of the American Statistical Association,* 50, 1168–1194.

Klebba, A. J. (1982), *Mortality From Diseases Associated With Smoking* (Vital and Health Statistics Ser. 20, No. 17; Department of Health and Human Services Publication PHS 82-1854), Washington, DC: U.S. Government Printing Office.

Kleinman, J. C. (1977), "Age-Adjusted Mortality Indexes for Small Areas: Applications to Health Planning," *American Journal of Public Health,* 67, 834–840.

Laird, N. M., and Louis, T. A. (1987), "Empirical Bayes Confidence Intervals Based on Bootstrap Samples" (with discussion), *Journal of the American Statistical Association,* 82, 739–757.

Lawless, J. F. (1987), "Regression Methods for Poisson Process Data," *Journal of the American Statistical Association,* 82, 808–815.

Manton, K. G., and Stallard, E. (1981), "Methods for the Analysis of Mortality Risks Across Heterogeneous Small Populations: Examination of Space-Time Gradients in Cancer Mortality in North Carolina Counties 1970–75," *Demography,* 18, 217–230.

——— (1984), *Recent Trends in Mortality Analysis,* Orlando, FL: Academic Press.
——— (1988), *Chronic Disease Modelling,* London: Charles W. Griffin.
Manton, K. G., Stallard, E., Creason, J. P., and Riggan, W. B. (1985), "U.S. Cancer Mortality 1950–1978: A Strategy for Analyzing Spatial and Temporal Patterns," *Environmental Health Perspectives,* 60, 369–380.
Manton, K. G., Woodbury, M. A., and Stallard, E. (1981), "A Variance Components Approach to Categorical Data Models With Heterogeneous Cell Populations: Analysis of Spatial Gradients in Lung Cancer Mortality Rates in North Carolina Counties," *Biometrics,* 37, 259–269.
Manton, K. G., Woodbury, M. A., Stallard, E., and Dowd, J. E. (1986), "A Composite Estimation Model for Producing Stabilized Health Rate Estimates for Small Areas Using Sample Surveys: Experience From Health Surveys in Ethiopia, India and Indonesia," in *Small Area Statistics,* eds. R. Platek and M. P. Singh, Ottawa: Laboratory for Research in Statistics and Probability (Carleton Univ.)/University of Ottawa, pp. 182–206.
Mason, T. J., McKay, F. W., Hoover, R., Blot, W. J., and Fraumeni, J. F. (1975), *Atlas of Cancer Mortality for U.S. Counties, 1950–1969* (Department of Health, Education, and Welfare Publication 75-780), Washington, DC: U.S. Government Printing Office.
——— (1976), *Atlas of Cancer Mortality Among U.S. Nonwhites: 1950–1969* (Department of Health, Education, and Welfare Publication 76-1204), Washington, DC: U.S. Government Printing Office.
Morris, C. N. (1983a), "Parametric Empirical Bayes Inference: Theory and Applications" (with discussion), *Journal of the American Statistical Association,* 78, 47–65.
——— (1983b), "Natural Exponential Families With Quadratic Variance Functions: Statistical Theory," *The Annals of Statistics,* 11, 515–529.
National Cancer Institute (1987), "Research Contributions Made Possible by the NCI Cancer Atlases Published in the 1970s," Office of Cancer Communications Report (Backgrounder Series), June 9, Bethesda, MD.
Pickle, L. W., Mason, T. J., Howard, N., Hoover, R., and Fraumeni, J. F. (1987), *Atlas of U.S. Cancer Mortality Among Whites, 1950–1980* (Department of Health and Human Services Publication 87-2900), Washington, DC: U.S. Government Printing Office.
Richards, F. S. (1961), "A Method of Maximum-Likelihood Estimation," *Journal of the Royal Statistical Society,* Ser. B, 23, 469–475.
Riggan, W. B., Creason, J. P., Nelson, W. C., Manton, K. G., Woodbury, M. A., Stallard, E., Pellom, A. C., and Beaubier, J. (1987), *U.S. Cancer Mortality Rates and Trends, 1950–1979, Volume IV: Maps* (Environmental Protection Agency Health Effects Research Laboratory Publication 600/1-83/015e), Washington, DC: U.S. Government Printing Office.
Schaible, W. L., Brock, D. B., and Schnack, G. A. (1977), "An Empirical Comparison of the Simple Inflation, Synthetic and Composite Estimators for Small Area Statistics," in *Proceedings of the Social Statistics Section, American Statistical Association,* pp. 1017–1021.
Tsutakawa, R. K. (1988), "Mixed Model for Analyzing Geographic Variability in Mortality Rates," *Journal of the American Statistical Association,* 83, 37–42.
Tsutakawa, R. K., Shoop, G. L., and Marienfeld, C. J. (1985), "Empirical Bayes Estimation of Cancer Mortality Rates," *Statistics in Medicine,* 4, 201–212.