Cancer surveillance research often begins with a rate matrix, also called a Lexis diagram, of cancer incidence derived from cancer registry and census data. Lexis diagrams with 3- or 5-year intervals for age group and for calendar year of diagnosis are often considered. This simple smoothing approach suffers from a significant limitation; important details useful in understanding time trends may be lost in the averaging process involved in generating a summary rate, and cohort-specific rates by age of diagnosis (Carstensen 2007).

The purpose of this article is to present a smoothed Lexis diagram and use it to study the effects of age, period, and cohort on incidence rates in terms of straightforward graphical displays. These include the age-specific rates by year of birth, age-specific rates by year of diagnosis, year-specific rates by age of diagnosis, and cohort-specific rates by age of diagnosis. We illustrate our approach by studying the trends in lung and breast cancer incidence in Taiwan. We find that for nearly every age group the incidence rates for lung adenocarcinoma and female invasive breast cancer increased rapidly in the past two decades and those for male lung squamous cell carcinoma started to decrease, which is consistent with the decline in the male smoking rate that began in 1985. Since the analyses indicate strong age, period, and cohort effects, it seems that both lung cancer and breast cancer will become more important public health problems in Taiwan. Supplementary materials for this article are available online.
or calendar year. This also has something to do with cancer screening projects, which we comment on in the discussion.

To simplify the discussion, we assume that the unit of time is a year, which is a positive integer, and that we consider a specific cancer. Let $M(x, y)$ denote the number of people not having been diagnosed with this cancer before age $x$ and year $y$, and $N(x, y)$ denote the number of people who are newly diagnosed with this cancer at age $x$ and year $y$. With proper transformation, we assume both $x$ and $y$ are represented by fraction numbers in $[0, 1]$. Let $F(x, y)$ denote the probability that an individual at age $x$ will be a newly diagnosed case in year $y$. $F(x, y)$ will be referred to as incidence rate. The main assumption of this article is that $N(x, y)$ is a Poisson random variable having parameter $M(x, y)F(x, y)$; thus,

$$P(N(x, y) = z) = \frac{[M(x, y)F(x, y)]^z}{z!} e^{-[M(x, y)F(x, y)]}.$$ 

We note that when the incidence rate $F(x, y)$ is small and the risk set size $M(x, y)$ is large, the Poisson model approximates the binomial model with parameters $M(x, y)$ and $F(x, y)$; this is the well-known law of rare events (Taylor and Karlin 1984). We also note that, while the binomial model is more intuitive, the Poisson model is computationally more favorable when $M(x, y)$ is large and $F(x, y)$ is small.

Assuming that the incidence rate $F$ varies smoothly with year of diagnosis and age of diagnosis, our main task is to estimate $F$ based on $M(x_k, y_k)$ and $N(x_k, y_k)$ for $k = 1, \ldots, K$. In applications, $(x_k, y_k)$ are often lattice points in $[0, 1] \times [0, 1]$, $M(x_k, y_k)$ and $N(x_k, y_k)$ can be obtained from data in census bureau and cancer registry, and $K$ is the total number of lattice points considered.

Since $F$ is a bounded nonnegative function, this is a shape-restricted inference problem. We will take a Bayesian approach that puts a prior on the space of bounded nonnegative smooth functions on $[0, 1] \times [0, 1]$. The age-$x$-specific incidence rate by calendar year $y$ of diagnosis will be reported based on the posterior distribution of $F(x, y)$. The trends of the incidence rate can be studied by the estimates of $F(x, y)$. We note that since Bernstein polynomials are differentiable, we can even explore the trends of the incidence rate by their partial derivatives, which often provide additional insights.

The prior will be specified in terms of two-dimensional Bernstein polynomials. We note that one-dimensional Bernstein polynomials have been successfully used in Bayesian shape-restricted inference problems regarding a smooth function on $[0, 1]$, using the facts that priors on Bernstein polynomials can be introduced through their coefficients, the geometry of a Bernstein polynomial can be read off readily from its coefficients, and the support of the space of Bernstein polynomials can be large (Chang et al. 2005, 2007; Chien et al. 2009). We now indicate that these properties can be extended to two-dimensional Bernstein polynomials and can be used to introduce priors for Bayesian shape-restricted inference for smooth functions on the unit square.

This article is organized as follows. Section 2 presents the statistical model using two-dimensional Bernstein polynomials, proposes the prior distribution, and describes the main observations used in the algorithm for sampling the posterior distribution. The algorithm itself is given in Appendix A. Section 3 presents a simulation study, which mimics one of the real datasets and paves the way for the real data analysis in Section 4. To illustrate our method, we study the trends of lung and breast cancers in Taiwan in Section 4. More detailed information about these cancers in Taiwan is given in Appendixes B–F. Section 5 provides a brief discussion. All Appendices are given in the supplementary materials.

2. BAYESIAN SHAPE-RESTRICTED INFERENCE

2.1 Modeling by Two-Dimensional Bernstein Polynomials

For integers $0 \leq i \leq n$, let $\varphi_{i,n}(t) = C_n^i t^i (1 - t)^{n-i}$, where $C_n^i = n!/(i!(n-i)!)$ and $0 \leq t \leq 1$. The set $\{\varphi_{i,n}(t) | i = 0, \ldots, n\}$ is called the Bernstein basis for polynomials of order $n$. Let $F_a(t_1, t_2) = \sum_{i=0}^{n} \sum_{j=0}^{n} a_{i,j} \varphi_{i,n}(t_1) \varphi_{j,n}(t_2)$, where $a = (a_{0,0}, \ldots, a_{n,n})$: $F_a$ is called a two-dimensional Bernstein polynomial. Using arguments in Chang et al. (2007), we obtain immediately

**Proposition 1.** If $R_1 \leq a_{i,j} \leq R_2$, for every $i = 0, \ldots, n$, $j = 0, \ldots, n$, then $R_1 \leq F_a(t_1, t_2) \leq R_2$ for every $t_1 \in [0, 1]$ and $t_2 \in [0, 1]$.

Let $B_a = \mathbb{R}^{(n+1) \times (n+1)}$ and $I_n = \{F_a | a \in B_a, R_1 \leq a_{i,j} \leq R_2, \text{for every } i = 0, \ldots, n, j = 0, \ldots, n, \}$. Let $I$ denote the set of all continuous functions on $[0, 1] \times [0, 1]$ having values between $R_1$ and $R_2$. Then, using Proposition 1 and the two-dimensional Bernstein polynomial approximation theorem (Altomare and Campitelli 1994), we have

**Proposition 2.** The closure of $\bigcup_{n=0}^{\infty} I_n$ is equal to $I$ in uniform norm.

Let $B = \bigcup_{n=0}^{\infty} (n) \times B_n$. Let $F : B \times [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ be defined by

$$F(n, b_{0,0,n}, \ldots, b_{n,n,n}, t_1, t_2) = \sum_{i=0}^{n} \sum_{j=0}^{n} b_{i,j,n} \varphi_{i,n}(t_1) \varphi_{j,n}(t_2),$$

where $(n, b_{0,0,n}, \ldots, b_{n,n,n}) \in B, t_1 \in [0, 1]$, and $t_2 \in [0, 1]$. We also denote (1) by $F_{b_n}(t_1, t_2)$ if $b_n = (b_{0,0,n}, \ldots, b_{n,n,n})$. $F$ gives a one-to-one correspondence between $B$ and the set of two-dimensional Bernstein polynomials.

With Proposition 1, Proposition 2, and (1), we are motivated to introduce a prior distribution on $I$ by defining a probability distribution on $B$. Let $\pi$ be a probability measure on $B$. $\pi$ is called a Bernstein prior. A natural way to introduce a prior $\pi$ is to consider $\pi(n, b_n) = \pi_n(b_n) p(n)$, with $\sum_{n=0}^{\infty} p(n) = 1$ and $\pi_n$ a density function on $B_n$; in fact, $\pi_n(\bullet)$ is the conditional density of $\pi$ on $B_n$ and also denoted by $\pi(\bullet | [n] \times B_n)$. The following proposition shows that the support of the Bernstein priors can be quite large.

**Proposition 3.** Let $\bar{B}_n = \{b_n \in B_n : F_{b_n} \in I_n\}$. Assume $p(n) > 0$ and the conditional density $\pi(b_n | [n] \times \bar{B}_n)$ has support $\bar{B}_n$ for infinitely many $n$. Let $F$ be a continuous real-valued function defined on $[0, 1] \times [0, 1]$ with $R_1 \leq F \leq R_2$. Then $\pi([n] \times b_n) \in \bigcup_{n=0}^{\infty} (r) \times \bar{B}_n : ||F_{b_n} - F||_{\infty} < \varepsilon > 0$ for every $\varepsilon > 0$.

The proof for Proposition 3 is similar to that for the Proposition 3 in Chang et al. (2007) and hence omitted.
As discussed in Section 1, we assume that given \((n, b_n) \in B\), \(N(x_1, y_1), \ldots, N(x_K, y_K)\) are independent Poisson random variables satisfying

\[
P(N(x_k, y_k) = z_k) = \left\{ \frac{[M(x_k, y_k)F_{bn}(x_k, y_k)]^a}{z_k!} \right\} e^{-[M(x_k, y_k)F_{bn}(x_k, y_k)]}. \]

Here \(F_{bn}(*, *)\) is the function on \([0, 1] \times [0, 1]\) associated with \((n, b_n) \in B\) defined in (1). Given the data \(\{(x_k, y_k), z_k, M(x_k, y_k)\} | k = 1, \ldots, K\), the posterior density \(\nu\) of the parameter \((n, b_n)\) is proportional to

\[
\prod_{k=1}^{K} \frac{[M(x_k, y_k)F_{bn}(x_k, y_k)]^a}{z_k!} e^{-[M(x_k, y_k)F_{bn}(x_k, y_k)]} \pi_n(b_n) p(n). \]

### 2.2 Priors Specification

We now describe the way we introduce the prior. The prior on \(p\) is motivated by convenience. We consider two situations: the case where \(p(0) = 0\) and the case where \(p(0) = p(1) = 0\). For the former, let \(p(1) = (1 + \alpha) e^{-\alpha}, p(n) = \alpha^n e^{-\alpha}/n!\) for \(n = 2, \ldots, n_0 - 1\), and \(p(n_0) = 1 - \sum_{n=1}^{n_0-1} p(n)\) for a suitable positive integer \(n_0\). For the latter, let \(p(2) = (1 + \alpha + \alpha^2/2) e^{-\alpha}, p(n) = \alpha^n e^{-\alpha}/n!\) for \(n = 3, \ldots, n_0 - 1\), and \(p(n_0) = 1 - \sum_{n=1}^{n_0-2} p(n)\) for a suitable positive integer \(n_0\).

The prior on conditional density \(\pi_n\) is motivated by two observations regarding the model and the data structure. The first is that the values of \(F_{bn}\) on the boundary of the unit square are closely related to its “boundary coefficients” \(\{b_{i,j,n} | i = 0, n or j = 0, n\}\): namely, \(F_{bn}(0,0) = b_{0,0,n}\), \(F_{bn}(0,1) = b_{0,0,n}, F_{bn}(1,0) = b_{0,0,n}, F_{bn}(1,1) = b_{0,0,n}, F_{bn}(0,1) = \sum_{j=0}^{n} b_{j,0,n} \varphi_j(n), F_{bn}(1,0) = \sum_{j=0}^{n} b_{0,j,n} \varphi_j(n), F_{bn}(1,1) = \sum_{j=0}^{n} b_{0,0,n} \varphi_j(n),\) and \(F_{bn}(1,1) = \sum_{j=0}^{n} b_{0,0,n} \varphi_j(n)\). Note that the last four equations are one-dimensional Bernstein polynomials whose bounds are identical to the bounds of their coefficients; see, for example, Chang et al. (2007).

The second observation is that if \(F_{bn}(x, y)\) is close to the true incidence \(F(x, y)\), then the raw incidence rates (Lexis diagram) \(N(x_k, y_k)/M(x_k, y_k)\) for \(k = 1, \ldots, K\) provide some idea of the range of the values of \(F_{bn}\). This, in turn, can be used to propose priors on the coefficients, \(b_n = \{b_{0,0,n}, \ldots, b_{n,n,n}\}\), in view of Proposition 1.

The conditional density \(\pi_n\) will be defined by assuming coordinates of \(b_n = \{b_{0,0,n}, \ldots, b_{n,n,n}\}\) are independent and \(b_{0,0,n}, b_{0,n,n}, b_{n,0,n}, b_{n,n,n}\) have respective prior distributions \(q_1, q_2, q_3, q_4\), which do not change with polynomial order. Elements in \(\{b_{i,j,n} | i = 1, n - 1, j = 1, n - 1\}\), \(\{b_{i,n,n} | i = 1, n - 1\}\), \(\{b_{0,j,n} | i = 1, n - 1\}\), and \(\{b_{n,n,n} | i = 1, n - 1\}\) have priors Uniform(M1, l1), Uniform(M1, l2), Uniform(M1, l3), and Uniform(l1, l4), respectively. Those in \(\{b_{i,j,n} | i = 1, n - 1, j = 1, n - 1\}\) have priors Uniform(M1, M2). Here l1, l2, l3, l4, M1, and M2 are considered constant and their values depend on the specific Lexis diagram to be analyzed; we will illustrate one method to choose these constants in the simulation studies and the real data examples.

### 2.3 Sampling the Posterior Distributions

Since the parameter space consists of subspaces of different dimension, we will use a Metropolis–Hastings reversible jump (MHR) algorithm (Green 1995) to generate data from the posterior distributions. While we will postpone the algorithm to Appendix A, which is technical and long, we will present here one observation that motivates the way we update the coefficients of the Bernstein polynomials. Since the state space of our Markov chain Monte Carlo (MCMC) algorithm consists of Bernstein polynomials, the closeness concept suggested by the Bernstein polynomial approximation theorem is useful in designing the MCMC. Let \(F_{bn}(*, *)\) denote the current state of the Markov chain. If \(F_{bn}(*, *)\) is close to the true incidence rate function \(F(*, *)\), then \(F_{bn}(*, *)\) should be also close to the Bernstein polynomial of \(F\), \(\sum_{m=0}^{M} \sum_{j=0}^{M} F(i/m, j/m) \varphi_{i,m}(x) \varphi_{j,m}(y)\), even for \(m \neq n\). This suggests that in updating the Markov chain, it is a good idea to make the new Bernstein polynomial “close” to \(F_{bn}(*, *)\) even if \(m \neq n\). There are several obvious methods to use this closeness concept in the algorithm. The one in our algorithm seems to be more computationally efficient, according to our preliminary simulation studies (not reported).

### 3. SIMULATION STUDY

We now explore the numerical performance of the Bayesian method in a simulation study. To make this examination more relevant to the real data examples on the incidence of lung cancer and breast cancer in Taiwan, the data in the simulation study are generated using an incidence rate function \(F_T\) that resembles the one from the female lung ADC example in Section 4 and a number of people at risk \(M(x, y)\) that approximates the number of females not having been diagnosed with lung ADC before age \(x\) and before year \(y\) in Taiwan in the period from 1988 to 2007. In fact, we consider

\[
F_T(x, y) = 0.00148 \sin(0.5 \pi xy) + 0.00002 \tag{2}
\]

and

\[
M(x, y) = y_0 + y_1 x + y_2 y + x y_3 + y_4 x y + y_5 y^2 + y_6 x^3 + y_7 x^2 y + y_8 x y^2 + y_9 y^3 \]

with \(y_0 = 152, 040, y_1 = -285, 270, y_2 = 110, 410, y_3 = 173, 900, y_4 = -49, 950, y_5 = -33, 630, y_6 = -19, 530, y_7 = -110, 330, y_8 = 88, 840,\) and \(y_9 = -7990\). We note that \(M(x, y)\) is a least-square estimate based on census data in Taiwan.

In this simulation study, \(K = 1020\). The design points \((x_k, y_k) | k = 1, \ldots, 1020\) form a lattice in \([0, 1] \times [0, 1]\). Specifically, \((X_{51, X+1}, Y_{51, Y+1}), (X_{51, X+2}, Y_{51, Y+2}), \ldots, (X_{51, X+51}, Y_{51, Y+51}) = \{(0, 1/19), (1/19, 1/19), \ldots, (1, 1/19)\}\) for every \(J = 0, \ldots, 19\).

Given \((x_k, y_k), M(x_k, y_k), F_T(x_k, y_k) | k = 1, \ldots, 1020\), the values of \(\{z_k | k = 1, \ldots, 1020\}\) are independently generated according to Poisson distributions with means \(\{M(x_k, y_k)F_T(x_k, y_k) | k = 1, \ldots, 1020\}\), respectively. Let \(r_k = z_k/M(x_k, y_k)\) for \(k = 1, \ldots, 1020\).

The specific priors we use are as follows. Let \(I_1 = \max \{r_k | k = 51 \times J + 1\}, J = 0, \ldots, 19\), \(I_2 = \max \{r_k | k = (51 \times J + 51), J = 0, \ldots, 19\}, I_3 = \max \{r_k | k = 1, \ldots, 51\}, I_4 = \max \{r_k | k = 970, \ldots, 1020\},\) and \(M_2 = \)
max\{r_k | k = 1, \ldots, 1020\}. Let q_1 and q_2 be Uniform(0, l_1), and q_3 and q_4 be Uniform(0, l_2). We consider four sets of priors in this simulation study so as to conduct some sensitivity analysis. Let M_1 be 0. The first set assumes l_1 = 2l_1, l_2 = 2l_2, l_3 = 2l_3, l_4 = 2l_4, M_2 = 2M_2, n_0 = 10, and \alpha = 5. The second set assumes l_1 = 2l_1, l_2 = 2l_2, l_3 = 2l_3, l_4 = 2l_4, M_2 = 2M_2, n_0 = 20, and \alpha = 10. The third set assumes l_1 = 4l_1, l_2 = 4l_2, l_3 = 4l_3, l_4 = 4l_4, M_2 = 4M_2, n_0 = 10, and \alpha = 5. The fourth set assumes l_1 = 4l_1, l_2 = 4l_2, l_3 = 4l_3, l_4 = 4l_4, M_2 = 4M_2, n_0 = 20, and \alpha = 10.

The prior in this simulation study and the following real data analyses assume p(0) = p(1) = 0, which means only Bernstein polynomials of order 2 or higher are included. This is because our preliminary analyses showed that the results obtained under the assumption p(0) = p(1) = 0 are all very similar to those under p(0) = 0, but the Markov chain converges much faster under p(0) = p(1) = 0.

We use the MHRJ algorithm with c = 0.35 to generate the posterior distribution. This value, c, controls the probabilities of changing the order of the polynomial in updating the chain; and c = 0.35 allows a fairly large such probability.

In this article, we take the following strategy to sample from the posterior distribution. Given a set of priors, we run five MCMC with initial values chosen randomly from the priors and monitor convergence by the Gelman–Rubin statistic \( R \), suggested in Gelman and Rubin (1992) and Gelman et al. (2004). The Gelman–Rubin statistic \( R \) is calculated for the volume under the surface. We first run each chain with 200,000 iterations and use the second halves of the five chains to compute \( R \). If \( R \) is smaller than 1.1, we stop. Otherwise, we double all chain lengths and again use the second halves to compute \( R \). We repeat the procedure until \( R \) is less than 1.1; the updates from the second halves of these five sequences are considered sample from the posterior distribution, which form the basis for inference.

The estimate of the incidence rate function is defined to be the posterior mean and is denoted by \( \hat{F} \). One naive way to obtain \( \hat{F} \) is to compute, for each of the 1020 grids, the mean of the values at the grid of the Bernstein polynomials in the updates. In case the orders of the polynomials in the posterior distribution are constant, \( \hat{F} \) is simply the mean of these Bernstein polynomials. We note that in all the Bayesian analyses in this article the polynomial order in the posterior distribution is no larger than 3; in most of the analyses, the polynomial order is exactly 2.

Table 1 describes certain properties of the estimates in the simulation studies. It indicates that the numerical performance of the Bayesian method in this study is quite good. Each of the rows presents the performance under one set of priors. Columns 2 and 3 give, respectively, the \( L_1 \)-norm and sup-norm of \( |\hat{F} - F_T| \times 10^5 \), where \( |\cdot| \) represents the absolute value. Columns 4 and 5 give, respectively, the posterior mean and maximum of \( \hat{F} \times 10^5 \). Columns 6, 7, and 8 report, respectively, the Gelman-Rubin statistic \( R \), the number of iterations that achieve it, and the CPU time for this computation. Column 9 reports the posterior predictive \( p \)-value, where the discrepancy function is the maximum of the difference between the observed incidence rates and the incidence rates specified by the model. There are other reasonable discrepancy functions to use, for example, we may use the one that measures the average of the difference between the observed rates and the rates specified by the model. We note that a large posterior predictive \( p \)-value suggests no inconsistency between the model and the dataset; see Gelman et al. (2004).

Table 1 shows that our method provides a good estimate of the true incidence rates, is insensitive to the choice of prior distributions, and is computationally efficient. In fact, using the criterion \( T = \max(|\hat{F}_{\text{prior}}(x_k, y_k) - \hat{F}_{\text{posterior}}(x_k, y_k)| | k = 1, \ldots, 1020) \times 10^5 \), the largest difference happens between set of priors 1 and 4 and has the value of 4.711.

Figure 1 shows the simulation study results with \( F_T \) given in (2). The upper panel shows the year-specific incidence rates by age of diagnosis (left); age-specific incidence rates by year of diagnosis (right). The lower panel shows the absolute value of the difference of year-specific incidence rates by age of diagnosis between the estimate \( \hat{F} \) and the true incidence rate \( F_T \) multiplied by 10^5 (left); that of age-specific incidence rates by year of diagnosis between the estimate \( \hat{F} \) and the true incidence rate \( F_T \) multiplied by 10^5 (right).

We now use this simulation study to compare our Bayesian method to the standard Lexis diagram in estimating the true incidence rate \( F_T \). The 1-year tabulated standard Lexis diagram consists of the \( r_k = z_k / M(x_k, y_k) \), for \( k = 1, \ldots, 1020 \), described above. Based on the 1-year tabulated standard Lexis diagram, we construct 3-, 5-, and 7-year tabulated standard Lexis diagrams in the same manner as that described in Carstensen (2007); see Table 1 and Figure 2 in Carstensen (2007). We may shorten the name of these Lexis diagrams, for example, the 5-year tabulated standard Lexis diagram may be referred to as the 5-year Lexis diagram. These standard Lexis diagrams provide estimates \( \hat{F} \) of the true incidence rate \( F_T \); it is reasonable to regard them as estimates at the middle points of the Lexis diagram. For the 1-year standard Lexis diagram, the middle points are defined to be the 1020 lattice points mentioned above. For other standard Lexis diagrams, the coordinates of the middle points are defined to be the averages of the coordinates of the lattice points involved in the averaging process that constructs the standard Lexis diagram. (This definition of middle points is slightly different from that in Carstensen (2007), but is more appropriate here in view of our data generation procedure.) Thus, when comparing our Bayesian method to a standard Lexis diagram, we consider the \( L_1 \)-norm and sup-norm of \( |\hat{F} - F_T| \times 10^5 \) over the middle points defined by each respective Lexis diagram; these comparison results are reported in Table 2. For this comparison, we use the set of priors number 4 in Table 1, because it has the largest posterior predictive \( p \)-value. Table 2 indicates clearly that our Bayesian method outperforms standard Lexis diagrams in providing estimates of the incidence rates. Our Bayesian method has the additional advantage of producing smooth estimates at every point in the unit square. We note that Figure 1(b) suggests that the largest error in the estimation of \( \hat{F} \) occurs at the boundary.

The results shown in Tables 1 and 2 are quite general. In fact, we carried out several such studies with different \( F_T \) and \( M(x, y) \) and obtained similar results. In the interest of conciseness, we present in Tables 3 and 4 one more study with the same \( M(x, y) \) but a different \( F_T \) defined by

\[
F_T(x, y) = 0.00148 \sin(0.5\pi x(y + 0.2)) + 0.00002. \tag{3}
\]
Table 1. Performance of the Bayesian estimate in the simulation study, with $F_T$ given in (2)

| Set of priors | $L_1$-norm | Sup-norm | Mean† | Max† | $\hat{R}$ | Iterations | Time (min.) | $p$-Value |
|---------------|-------------|----------|-------|-------|-----------|------------|-------------|-----------|
| 1             | 0.676       | 7.842    | 53.96 | 157.8 | 1.016     | 200,000    | 6           | 0.739     |
| 2             | 0.751       | 10.99    | 54.00 | 161.0 | 1.001     | 200,000    | 6           | 0.770     |
| 3             | 0.741       | 11.23    | 54.03 | 161.2 | 1.002     | 200,000    | 6           | 0.753     |
| 4             | 0.872       | 7.929    | 54.08 | 157.9 | 1.061     | 200,000    | 6           | 0.808     |

NOTE: †The mean and max values of the true incidence rate multiplied by $10^5$ are 54.07 and 150, respectively.

In fact, the only difference between these two studies is that the true incidence rate function $F_T$ in (3) is not monotone in $y$. Tables 3 and 4 are, respectively, the counterparts of Tables 1 and 2. Together, Tables 3 and 4 show again that our Bayesian method outperforms standard Lexis diagrams.

4. TRENDS IN TAIWAN LUNG CANCER AND BREAST CANCER

4.1 Preliminaries

Taiwan is a densely populated island with one of the fastest growing aging populations in the world and has experienced tremendous lifestyle changes in the past several decades. With cancer rates on the rise, it is of interest to investigate both the time trends and secular trends to monitor changes in the incidence rates of various cancers. For example, using the Taiwan Cancer Registry from 1980 to 2006, in which age-standardized incidence rates of most important cancers were reported, the effects of the 1995 implementation of Taiwan’s National Health Insurance on cancer incidence rates were explored (Chiang et al. 2010).

There also appeared more detailed epidemiological studies of specific cancers. For example, using the Taiwan Cancer Registry and data from SEER program, the incidence of FIBC (female invasive breast cancer) in the younger generation of Taiwanese was found to be rapidly increasing and approaching that of Caucasian Americans, suggesting a consequence of Westernization of lifestyles (Shen et al. 2005). As another example, lung ADC (adenocarcinoma) was found to be the predominant lung cancer histology subtype among Taiwan female patients; furthermore, its incidence rate was increasing, based on data from National Taiwan University Hospital (Chen et al. 2005).

As an illustration of the method of this article, we will study age, period, and cohort effects on breast cancer and lung cancer in Taiwan. Specifically, we consider people with age between 35 and 85 years in the period from 1988 to 2007, based on

![Figure 1. This figure provides graphs of the estimate $\hat{F}$ and the absolute difference between the estimate $\hat{F}$ and the true incidence rate $F_T$.](image-url)
Figure 2. Contour plots. The number on each contour curve shows the incidence rate for each point on the curve.

the Taiwan Cancer Registry and Taiwan census data. Since the incidence rate of female lung SCC (squamous cell carcinoma) is very low, we consider only FIBC, female lung ADC, male lung ADC, and male lung SCC.

The Taiwan Cancer Registry was launched in 1979; it collects information on newly diagnosed cancer patients from all hospitals in Taiwan with a capacity of more than 50 beds. Knowing that the quality of the data has improved steadily, we focused the current study on a recent 20-year period. For a more detailed description of the Taiwan Cancer Registry, see Chiang et al. (2010).

| Class | $L_1$-norm | $L_1$-norm | $L_1$-norm | $L_1$-norm |
|-------|-------------|-------------|-------------|-------------|
| 1 year| 0.872       | 7.929       | 7.532       | 67.73       |
| 3 year| 0.882       | 6.962       | 2.689       | 23.95       |
| 5 year| 0.833       | 4.794       | 1.408       | 7.700       |
| 7 year| 1.124       | 6.425       | 1.363       | 6.054       |

NOTE: Both $L_1$-norm and sup-norm of $|\hat{F} - F_T| \times 10^4$ are over the middle points of each respective Lexis diagram.

4.2 Construction of the Smoothed Lexis Diagram

We now describe the way we construct the smoothed Lexis diagram for the incidence of female lung ADC. The smoothed Lexis diagrams of other cancers are constructed similarly. Using the notation of Section 3, for the lattice point $(x_k, y_k)$ in $[0, 1] \times [0, 1]$, the number of newly diagnosed female lung ADC patients at the age $35 + 50x_k$ in the year $1988 + 19y_k$ is denoted by $N(x_k, y_k)$, which were obtained from the Taiwan Cancer Registry on June 11, 2010. $M(x_k, y_k)$ is similarly defined and was derived jointly from Taiwan census data and the Taiwan Cancer Registry. The ICD-9-CM codes for the cancer site (topography) and the ICD-O3 codes for the histology (morphology) used in this study are given in Appendix B.

Let $r_k = N(x_k, y_k) / M(x_k, y_k)$ for $k = 1, \ldots, 1020$; we analyze the real datasets in the same manner as we conduct simulation studies in Section 3. In particular, we consider the same four sets of priors, carry out the same MCMC algorithm for posterior distribution sampling, use the Gelman–Rubin statistic to monitor the convergence, and apply the posterior predictive $p$-value to assess the model fitness.

Tables C1–C4 in Appendix C summarize the MCMC computation for FIBC, female lung ADC, male lung ADC, and male lung SCC, respectively. Entries in these tables bear the same meaning as those in Table 1. To conduct sensitivity analyses, we compare the results obtained from analyses using the four
different sets of priors for each cancer dataset. Tables C1-1, C2-1, C3-1, and C4-1, also in Appendix C, report the differences between the posterior means studied in Tables C1, C2, C3, and C4, respectively. These tables indicate that, except in the case of male lung SCC, the models considered are quite consistent with the datasets and the results are not sensitive to the priors. In fact, even for male lung SCC, two of the sets of priors show acceptable posterior means studied in Tables C1, C3-1, and C4-1, also in Appendix C, report the differences between the posterior means.

The following results are based on those priors showing largest posterior predictive p-values.

Figure 2 presents the smoothed Lexis diagrams, with 2a, 2b, 2c, and 2d for FIBC, female lung ADC, male lung ADC, and male lung SCC, respectively. Contour lines with associated incidence rates are shown in these figures; by definition, incidence rates at points on the same contour line are the same; they are given in terms of number of cases per 100,000 person-years. For example, Figure 2(c) shows that in the year 1988, the incidence rate for an 83-year-old male was about 60 cases per 100,000 person-years. These contour lines help in grasping the pattern change of the incidence rate for each cancer. For example, while the incidence rate of lung ADC increases with both age and calendar year, male lung SCC incidence rate peaks around 2003 among older people. In view of the reports that smoking is a risk factor for lung cancer, more for SCC than for ADC, and that among never smoking lung cancer patients, ADC subtype is more prevalent than SCC (Sun, Schiller, and Gazdar 2007; Couraud et al. 2012), it is of interest to see if the pattern of contour lines in Figure 2 can be explained by a decrease in the smoking rate in Taiwan.

### 4.3 Age, Period, and Cohort Effects on Incidence Rate

We now explore the age, period, and cohort effects on incidence rate in terms of the four classical plots derived from a Lexis diagram (Carstensen 2007). For each of the smoothed Lexis diagrams constructed in last subsection, these four plots are similarly constructed and reported in Figures 3–6. They are, respectively, the plots for age-specific rates by year of birth (rates vs. age of diagnosis, observations within each birth-cohort are connected), age-specific rates by year of diagnosis (rates vs. age of diagnosis, observations within each year of diagnosis are connected), year-specific rates by age of diagnosis (rates vs. year of diagnosis, observations within each age group are connected), and cohort-specific rates by age of diagnosis (rates vs. birth-cohort, observations within each age group are connected).

Each of these figures has four parts, indexed by a, b, c, and d, for FIBC, female lung ADC, male lung ADC, and male lung SCC, respectively.

Since our main purpose is to illustrate the method, we explore some possible uses of each of these figures, even if the same conclusion may be observed from different figures.

The following observations are clear from Figure 3. Given age of diagnosis, later birth-cohorts experience a higher incidence rate for lung ADC. Given a birth-cohort and regarding incidence rate as a function of age, this function is increasing and convex for both female and male lung ADC; it is increasing for male lung SCC. It is also increasing for the younger cohort for FIBC; but it is concave with a peak for birth-cohorts around 1928 and 1923; we expect to see peaks in later cohorts if we observe the rates in longer periods. It might be of interest to see if these observations hold in other geographical regions and to provide an interpretation if they do.

Figure 3(b)–(d) indicates that our results are in line with the findings that, among different lung cancer histology types, adenocarcinoma is the most prevalent and its incidence is increasing in East Asia and West Europe (Janssen-Heijnen and Coebergh 2003; Chen et al. 2005; Devesa et al. 2005; Lian et al. 2006; Toyoda et al. 2008). The facts that incidence rate is increasing and convex suggest a rapid increase tendency. One way to stress this is to compute their derivatives. Figure D1 in Appendix D shows that these derivatives for lung ADC are not only positive but also increasing by themselves, and that those for male lung SCC are positive but not always increasing. Figures D1–D4 in Appendix D are, respectively, the derivatives for those in Figures 3–6.

Figure 4 indicates that given year of diagnosis and regarding incidence rate as a function of age, this function is increasing for lung ADC and male lung SCC, whereas it is unimodal for FIBC with a peak around 53 years of age. For FIBC and lung ADC, Figure 4 also indicates that, except for male lung SCC, the age-specific rate increases with year of diagnosis for nearly every age group.

### Table 3. Performance of the Bayesian estimate in the simulation study, with $F_T$ given in (3)

| Set of priors | $L_1$-norm | Sup-norm | Mean$^1$ | Max$^1$ | $\hat{R}$ | Iterations | Time (min.) | $p$-Value |
|---------------|-------------|-----------|----------|---------|-----------|------------|-------------|-----------|
| 1             | 1.126       | 8.647     | 71.08    | 155.7   | 1.002     | 200,000    | 3           | 0.487     |
| 2             | 1.033       | 5.439     | 71.05    | 154.9   | 1.018     | 200,000    | 3           | 0.523     |
| 3             | 1.086       | 10.45     | 71.08    | 156.1   | 1.004     | 200,000    | 3           | 0.460     |
| 4             | 1.086       | 10.03     | 71.08    | 156.0   | 1.004     | 200,000    | 8           | 0.468     |

**NOTE:** The mean and max values of the true incidence rate multiplied by $10^5$ are 71.33 and 150, respectively.

### Table 4. Comparison of the Bayesian estimate and the standard Lexis diagrams based on the simulation study, with $F_T$ given in (3)

| Class | Bayesian estimate $L_1$-norm$^1$ | Sup-norm$^1$ | Standard Lexis diagram $L_1$-norm$^1$ | Sup-norm$^1$ |
|-------|---------------------------------|-------------|---------------------------------------|-------------|
| 1 year| 1.033                           | 5.439       | 8.755                                 | 85.09       |
| 3 year| 1.034                           | 4.123       | 2.791                                 | 18.29       |
| 5 year| 0.972                           | 2.926       | 1.999                                 | 8.554       |
| 7 year| 1.162                           | 4.510       | 1.977                                 | 18.20       |

**NOTE:** Both $L_1$-norm and sup-norm of $|\hat{F} - F_T| \times 10^5$ are over the middle points of each respective Lexis diagram.
Figure 3. Age-specific rates by year of birth (rates vs. age of diagnosis, observations within each birth-cohort are connected).

Figure 4. Age-specific rates by year of diagnosis (rates vs. age of diagnosis, observations within each year are connected).
Figure 5. Year-specific rates by age of diagnosis (rates vs. year of diagnosis, observations within each age group are connected).

Figure 6. Cohort-specific rates by age of diagnosis (rates vs. birth-cohort, observations within each age group are connected).
Figure 5 indicates that for every age specification, incidence rate as a function of year of diagnosis increases for lung ADC; for some age groups, this does not hold for FIBC and male lung SCC. For example, Figure 5(d) shows that the male lung SCC incidence rate first increases and then decreases for older people. This is consistent with the declining smoking rate in Taiwan. We note that for the period from 1971 to 2010, the adult female smoking rate has been fluctuating around 4% and that for males decreased from 60% in 1985 to 35% in 2010, as described in Appendix E.

Figure 5 also shows that, for a given year of diagnosis, older people experience a higher incidence rate of lung cancer. For FIBC, peak incidence occurs around the age of 53, which can also be seen from Figure 4(a) and Figure D2a. In fact, Figure D2a suggests that the age that peaks increases slightly with calendar year; this phenomenon deserves further attention, in view of the remark that in Japan and Taiwan the peak age does not seem to be increasing with calendar year as in Hong Kong and the United States (Shen et al. 2005). We note that the incidence rate of FIBC in California, USA, peaks around 70 (Clarke et al. 2005).

Figure 6 shows that given an age specification, people in later cohorts experience higher rates, except for male lung SCC and FIBC. For example, Figure 6(d) shows that for males 75 years of age, the 1928 cohort experiences higher lung SCC rates than nearby cohorts. This is also consistent with the declining smoking rate in Taiwan. For FIBC, among females 85 years of age, the 1918 cohort experiences higher FIBC rates than nearby cohorts, but among the younger cohorts, later ones experience higher rates.

4.4 Comparison With Standard Lexis Diagram

4.4.1 General Observations. We illustrate the use of smoothed Lexis diagrams in studying lung and breast cancer trends in Taiwan by presenting the four classical plots that depict the age, period, and cohort effects on incidence rates. Except for male lung SCC, the smoothed Lexis diagrams are based on the Bayesian models that fit the datasets well and the posterior distributions are not sensitive to the priors (which enhances our confidence in the four classical plots). We now compare those plots with ones obtained from standard Lexis diagrams to examine the gains and limitations of our approach.

For this comparison, we obtained standard Lexis diagrams with 1-, 3-, 5-, 7-, 9-, 11-, and 13-year tabulation periods in the manner described in Section 3. Observations from the comparison are based on mostly 1-, 3-, and 5-year Lexis diagrams, although plots from only 1- and 3-year Lexis diagrams are shown. The four classical plots from standard Lexis diagrams with 1-year and 3-year intervals are presented in Appendix F. Considering the smooth plots for age-specific rates by year of birth in Figure 3, we present in Figure F1-1 (F1-3) their counterparts based on standard Lexis diagrams with 1 (3)-year intervals. Likewise, Figure 4 has counterparts Figure F2-1 (F2-3), Figure 5 has counterparts Figure F3-1 (F3-3), and Figure 6 has counterparts Figure F4-1 (F4-3). F1-1, F1-3, F2-1, F2-3, F3-1, F3-3, F4-1, and F4-3 have four plots each: for FIBC, female lung ADC, male lung ADC, and male lung SCC.

Here are some general remarks from this comparison. One-year tabulated standard Lexis diagrams are volatile, and the four classical plots based on them are typically noisy. In general, the longer the tabulation period of a standard Lexis diagram is, the smoother the corresponding classical plots, the shorter the domains of definition of these classical plots, and the fewer the number of the classical plots that can be obtained. However, our smoothed Lexis diagram does not suffer any such drawbacks.

The main observation of the comparison based on the four cancers in Taiwan is that our methods reveal cancer incidence trends more clearly than standard Lexis diagrams; if the trends do not involve the oldest age group, then they are not in contradiction with the results from standard Lexis diagrams, which may or may not reveal any trends; for trends regarding the oldest age group, there may exist contradictions, for which extra efforts may offer some insights.

We now present two examples to portray the above observations. The first regards age-specific rates by year of diagnosis for breast cancer. The smoothed Lexis diagram shows that, around the year 2006, people around the age of 53 experienced the highest incidence rate among all age groups, while the 3-year standard Lexis diagram shows that, around the same period, people between the ages of 50 and 65 experienced the highest incidence rate; see Figure 4(a) and Figure F2-3a in the supplementary materials. We note that the age range showing the highest incidence rate gets smaller as the tabulation period gets longer and that the 13-year standard Lexis diagram reports that the highest incidence rate is around 55-years of age (data not shown). We also note that the 13-year Lexis diagram for our data is a $2 \times 4$ matrix.

4.4.2 Choosing the Age Range. The second example regards a possible discrepancy shown in the plots for age-specific rates by year of diagnosis among the oldest male lung SCC patients. The smoothed Lexis diagram indicates that the incidence rate increases with age for each year of diagnosis, whereas the standard Lexis diagram indicates that for certain years of diagnosis, it increases with age only before age 75 (see Figure 4(d) and Figure F2-3d). Thus, we can more confidently claim that the male lung SCC incidence rate increases with age until at least age 75. To gain some insights into this discrepancy, we examine several standard Lexis diagrams simultaneously. Based on the 3-year Lexis diagram, it seems that incidence rates increase over larger age ranges for more recent years of diagnosis than for earlier years of diagnosis. This pattern becomes more prominent for 5- and 7-year period tabulated standard Lexis diagrams, as well as for longer periods; eventually, the incidence rate increases everywhere in its domain for 13-year Lexis diagram (data not shown). These plots show that the agreement between our results and those from standard Lexis diagrams is higher for more recent years of diagnosis.

The above discrepancy seems to be in line with the small predictive $p$-values shown in Table C4 and mentioned in Section 4.2. Small predictive $p$-values suggest inconsistency between model and dataset and may serve as warnings that prompt for more detailed studies. We think one possible source for this discrepancy regards the fact that incidence rate estimation error is likely to be biggest for the oldest people in 1988, because Taiwan is a fast-growing aging population with the number
of male at age 85 being 4103 for 1988 and being 19,630 for 2007. To avoid this type of difficulties, we considered smoothed Lexis diagrams for people with different age ranges, including \([35, 65]\), \([35, 70]\), \([35, 75]\), \([65, 85]\), and \([70, 85]\). We find the posterior predictive \(p\)-values are not small and the classical plots based on them are more satisfactory (details are omitted).

4.5 Remarks on the Taiwan Lung and Breast Cancer Trends

The single most prominent finding in the above smoothed Lexis diagrams is that disease burden, as a function of year of diagnosis, is steadily increasing in lung ADC and FIBC, and decreasing in male lung SCC, for the period from 1988 to 2007. While the latter is consistent with the reduction of smoking rate, the former deserves serious attention from the point of view of public health.

Although the incidence rates of both lung ADC and FIBC are on the rise, they show different patterns. Figure 4 shows that considering incidence rate as a function of age of diagnosis, given year of diagnosis, lung ADC, and FIBC show very different patterns, with the former increasing in a much larger age range. In our study, FIBC peaks between 50 and 55 years of age; but in the United States it peaks at about age 70 (Shen et al. 2005), which prompts questions on the role of environmental contributions to FIBC susceptibility. In view of the discussion of the impact of Westernization of lifestyles on FIBC incidence in Taiwan (Shen et al. 2005) and of the main molecular subtypes of FIBC in Taiwan (Lin et al. 2009), we think it of interest to learn why FIBC peaks at such an early age in Taiwan. From Figure 3(a), we know that except for the 1923–1928 birth cohorts, age-specific incidence rate increases with age for each birth cohort in the period from 1988 to 2007; for birth cohort around 1923–1928, it increases until age 75. We think age-specific rates by year of birth are relevant in addressing the impact of Westernization of lifestyles on FIBC incidence rates and we need to obtain them from other periods and countries to better understand this issue.

Given that the smoking rate among males is decreasing, that the smoking rate among females has remained small, and that the incidence rate of lung ADC is rapidly increasing among both males and females, it is of interest to look for other risk factors for the development of lung ADC in Taiwan (Hsiung et al. 2010; Lan et al. 2012; Lo et al. 2013). Although there have been advances in these directions, further studies are needed to address this public health issue.

5. DISCUSSION

We use a Poisson model to describe the relationship between the number of new cases, the number of people at risk, and a smoothly varying incidence rate for the study of the incidence rate function. Based on the Poisson model, we use the standard Lexis diagram to introduce priors through the coefficients of Bernstein polynomials and propose a Bayesian approach to construct a smoothed Lexis diagram. We outline a fairly standard procedure for implementing the MCMC algorithm. For the four cancers studied in this article, we generally find that the models fit the datasets well, the results do not seem to be sensitive to the priors, and it is a good idea to construct the four classical plots to describe directly the effects of age, period, and birth-cohort on cancer incidence rates. While these classical plots from smoothed Lexis diagrams are revealing, comparing them with their counterparts based on standard Lexis diagrams and considering smoothed Lexis diagrams based on different age range helps to gain more confidence in the results or to explore subtle patterns regarding older people.

The prior distributions consist of two parts: the polynomial order and the polynomial coefficients. As discussed in Section 2, the geometric properties of Bernstein polynomials suggest we introduce the prior on polynomial coefficients through the raw rate matrix. As for the prior on the polynomial order, we used a truncated Poisson distribution out of convenience. For the simulation study and the applications, it turns out that the posterior distribution of the polynomial order is exactly 2 for most of the studies and is never larger than 3, which are much smaller than the polynomial order in the prior. We think it is a good idea to try several priors on the polynomial order to see if the results are sensitive to the choice of the prior.

Our approach assumes a smoothly varying incidence rate on the age range and period range under study. This assumption may not hold when screening rate changes due to factors such as media coverage of a famous person with cancer (Lane, Poledenak, and Burg 1989), a publicity campaign to increase screening (Cram et al. 2003), implementation of a specific cancer screening project (Wu et al. 2012), or a nationwide health insurance policy (Chiang et al. 2010). Although some of these screening effects may not affect long-term trends and smoothing over them may not cause too much concern (Feuer and Wun 1992), it is advisable to always pay attention to this assumption when applying the method of this article. We note that our Bayesian analysis includes posterior predictive \(p\)-values to assess the fit of the model to the data and small predictive \(p\)-values may serve as warnings that further investigations are needed, which may be useful in the choice of age range and period range under study.

Taiwan’s national health insurance program started in 1995; it is of interest to see if this event might affect the otherwise possibly smoothly varying incidence rate function and consequently the results in this article. One possible way to deal with this issue is to compare smoothed Lexis diagrams of periods before and after 1995 for each cancer. It would also be of interest to compare these patterns across different cancers. While our example on male lung SCC exemplifies the importance of considering Lexis diagrams covering different ranges of age of diagnosis, this discussion suggests the possible use of Lexis diagrams covering different ranges of year of diagnosis.

While this approach allows formal Bayesian inference, our illustrations focus on graphical displays. Other uses of the posterior distribution of incidence rates should be possible. For example, it might be interesting to explore the posterior distribution to propose age, period, and cohort models for closer surveillance studies.

SUPPLEMENTARY MATERIALS

The supplementary materials include Appendices A–F.
A. (2013), “Risk Factors for Primary Lung Cancer among Never Smokers by Gender in a Matched Case-Control Study,” Cancer Causes Control, 24, 567–576. [i010]

Rosenberg, P. S., and Anderson, W. F. (2011), “Age-Period-Cohort Models in Cancer Surveillance Research: Ready for Prime Time?,” Cancer Epidemiology Biomarkers and Prevention, 20, 1263–1268. [i000]

Shen, Y. C., Chang, C. J., Hsu, C., Cheng, C. C., Chiu, C. F., and Cheng, A. L. (2005), “Significant Difference in the Trends of Female Breast Cancer Incidence between Taiwanese and Caucasian Americans: Implications from Age-Period-Cohort Analysis,” Cancer Epidemiology Biomarkers and Prevention, 14, 1986–1990. [i004,i009,i010]

Sun, S., Schiller, J. H., and Gazdar, A. F. (2007), “Lung Cancer in Never Smokers—A Different Disease,” Nature Reviews Cancer, 7, 778–790. [i006]

Taylor, H. M., and Karlin, S. (1984), An Introduction to Stochastic Modeling, Orlando, FL: Academic Press. [i001]

Toyoda, Y., Nakayama, T., Ioka, A., and Tsukuma, H. (2008), “Trends in Lung Cancer Incidence by Histological Type in Osaka, Japan,” Japanese Journal of Clinical Oncology, 38, 534–539. [i006]

Wu, T. Y., Chung, S., Yeh, M. C., Chang, S. C., Hsieh, H. F., and Ha, S. J. (2012), “Understanding Breast Cancer Screening Practices in Taiwan: A Country with Universal Health Care,” Asian Pacific Journal of Cancer Prevention, 13, 4289–4294. [i010]