A Bayesian analysis of racial differences in treatment among breast-cancer patients

Balgobin Nandram[1], Dhiman Bhadra[21] and Yiwei Liu[3]

[1]*Department of Mathematical Sciences, WPI, Worcester, MA 01609*

[2]*Production and Quantitative Methods Area, Indian Institute of Management Ahmedabad, Gujarat 380015, India*

[3]*Fishing Partnership Support Services, 30 Chestnut Ave, Burlington MA 01803*

Abstract

It is a well known fact that race and ethnicity specific variations exist in the treatment and survival of cancer patients. Studies based on breast cancer patients admitted to community hospitals in U.S depicted that there is significant difference in patterns of care between black and white breast cancer patients with blacks receiving lower quality and quantity of care. In this study, we look at this problem from a different perspective, treating the hospitals as small areas, and employing Bayesian techniques for parameter estimation. Two separate models are constructed to estimate the odds ratio of receiving liver scan (a pattern of care) for blacks and whites. The first model uses hospital-specific information while the second one uses pooled hospital data by borrowing strength from neighbouring hospitals. We have used the non-central hypergeometric distribution as the basis for constructing the likelihood while estimation has been carried out using the griddy Metropolis-Hastings sampler. We apply our methodology on a National Cancer Institute (NCI) database. Although our results corroborate some of the observations from previous studies, it proposes a computationally attractive alternative to the established procedures in formulating and analyzing this problem.

Keywords : Breast cancer; Bayesian statistics; Hypergeometric; Metropolis-Hastings; small areas

[1]Corresponding author: e-mail: dhiman@iimahd.ernet.in, Phone: +91-79-6632-4893, Fax: +91-79-6632-6896

1 Introduction

It is a well documented fact that survival rates of women with breast cancer in the U.S is lower in blacks than whites. An obvious question that comes up in this context is whether this difference can be attributed, at least partly, to differences in treatment and diagnostic methods. In fact, some studies have indicated that, in general, blacks face more barriers to access and lower quality (and quantity) of treatment and care compared to whites (Flaherty et al., 1980; Yergan et al., 1987; Diehr at al., 1989). This discrimination has also been reflected in some studies specific to breast cancer. In one such study (Axtell and Myers, 1978), it was observed that twice as many blacks as whites had no treatment for their breast cancer between 1960 and 1966 while the five-year survival rates of whites were significantly higher than blacks (65% vs. 51%) across all stages of the disease between 1967-73. Another study (McWhorter and Mayer, 1987) reported that blacks were more likely to have their breast cancer go untreated (or treated surgically) compared to whites. These results were unchanged after adjustment for age, stage and histology. The five year survival rate was also found to be associated with treatment type (and favoured whites) after controlling for age, stage of disease and histology. Based on data collected by the National Cancer Institute Surveillance, Epidemiology and End Results [NCISEER], it was concluded that blacks had lower relative survival rates than all other racial groups except American Indians (Young et al., 1984).

One of the studies mentioned above (Diehr at al., 1989) used a dataset based on a National Cancer Institute funded program to explore whether there are significant differences in patterns of care for black and white breast cancer patients. The program generated data on 7,781 breast cancer patients treated at 107 community hospitals in the US between 1982 and 1985. This study explored whether these differences can be attributed to age, stage of disease, insurance, hospital type and physician characteristics. It was observed that, after controlling for age and stage of disease, black patients received significantly less appropriate care than whites in four out of the ten patterns of care originally considered (progesterone receptor assay, rehabilitation services, liver scan and radiation therapy). The race differences for the last three patterns remained significant even after controlling for hospital, insurance and physician variables. Last but not the least, race and care pattern were significantly related to insurance, physician and hospital characteristics in most of the cases; nevertheless, race difference remained significant even after controlling for these factors. This work employed classical statistical approaches, based on logistic regression, to arrive at the above conclusions. However, the sample sizes in the hospitals were quite restrictive (considering that only 19 of the 107 hospitals had a significant number of black

patients).

Recent articles have also corroborated race and ethnicity-specific variations in treatment and survival of cancer patients. A review article (Shavers and Brown, 2002) looked at 87 studies carried out between 1990 and 2001 concerning racial/ethnic variation in treatment given to breast, cervical, colorectal, prostate and lung cancer patients. The overall observations were : (i) African-Americans have a 33% higher risk of dying (from cancer) compared to whites while this value is 100% (two times) when compared to Asian/Pacific islanders, American Indians and Hispanics; (ii) 5-year survival rates for African-Americans are lower than those for whites for all of the major cancer types; (iii) racial disparities exist in the receipt of definitive primary therapy, conservative therapy and adjuvant therapy; (iv) although mortality rates have been decreasing among both African-Americans and whites, the decrease generally have been smaller and less consistent among the former while racial disparities have persisted; (v) African-American women with stage I or II breast cancer have shorter overall survival than white women controlling for treatment and chemotherapy rates; (vii) elderly African-American women generally are two times more likely than whites not to receive radiation therapy after breast-conserving surgery; (viii) African-American women are significantly less likely (than whites) to have a mammogram after diagnosis and treatment of breast cancer and finally (ix) controlling for tumour size and comorbidity, African American women were significantly less likely than whites to receive breast conserving surgery and radiation therapy.

A study (Shavers et al., 2002) based on the SEER (Surveillance, Epidemiology and End Results) program data for 1990-1998 too found racial and ethnic variation in treatment and survival of breast cancer patients under age 35. Specifically for this age group, the cancer incidence rate for African-American women was more than twice that for white women of similar age while the mortality rate was more than three times higher. Also African-American and Hispanic women had poorer overall survival after controlling for clinical and demographic characteristics and treatment type. Finally, an article (Smigal et al., 2006) by researchers at the American Cancer Society concluded that breast cancer incidence rates increased rapidly among women of all races from 1980-1987 but increased at a much slower pace between 1987-2002. It was also observed that trends in incidence vary by age, race, socioeconomic status and stage and lastly that death rates for African-American women remain 37% higher than in whites. Their findings were based on incidence data from 1975-2002.

In this paper, we have viewed the above problem from an altogether different perspective by treating the hospitals as small areas. Small area estimation techniques have been widely used for estimating various features of small domains - domains for which the sample size is prohibitively small to yield direct survey-based estimates of adequate precision (Ghosh and

Rao, 1994; Rao, 2003). Thus, classical statistical procedures, when applied in this context, may yield erroneous conclusions and thus can fail altogether. Small domains can be specific regions like a state, county or school district (hospitals in our case) or can even be identified by particular socio-demographic characteristics like specific racial or ethnic groups. These estimates play an increasingly vital role in government policy making, administration of federal programs and allocation of federal funds to local jurisdictions. For example, state level estimates of median income for four-person families are needed by the U.S Department of Health and Human Services (HHS) in order to formulate its energy assistance program to low income families. Due to smallness of the sample sizes, a popular approach in this estimation procedure is to "borrow strength" from neighbouring areas to improve the accuracy of estimates for a given area. This entails the need to develop and incorporate alternative estimation procedures, one of them being the hierarchical Bayesian methodology.

In the present study, we have viewed hospitals as small domains. Patients in a given hospital are considered a sub-population of all hospital patients. For the purpose of our study, we have considered 1,856 patients from 19 of the 107 hospitals with the highest percentage (more than 10%) of black patients. We have used Bayesian methodology to infer about the population of interest, which, in our case, are the breast cancer patients categorized by race (blacks vs. whites) who received a liver scan (a pattern of care; Sec 2 has details). Specifically, we estimate the odds ratio of receiving liver scan for the two groups of patients in a hierarchical Bayesian framework with the goal of identifying a valid method of obtaining parameter estimates from small areas in a context like ours. For reasons that will be detailed later, we believe that our proposed approach is computationally advantageous compared to some of the existing ones.

The rest of the paper is organized as follows. In Section 2, we describe the breast cancer dataset that motivated our study. Section 3 gives an overview of some classical statistical procedures that forms the foundation of the proposed Bayesian approach. Section 4 details the hierarchical Bayesian procedure that we have used to analyze this data. Section 5 outlines the analysis of the breast cancer data using the proposed methodology including details on MCMC computation. We end with a discussion in Section 6.

2 Data description

The dataset used in this article is based on a study (Diehr et al., 1989) to assess the relationship between race and patterns of care for 7,781 breast cancer patients treated at 107 community hospitals in the US between 1982 and 1985. This study was initiated and funded by the National Cancer Institute (NCI) through the Community Clinical Oncology Program

(CCOP) in 1983 with a view to increase community physicians participation in national clinical trials. To assess the effectiveness of the above program, the NCI also funded the Community Cancer Care Evaluation (CCCE) project (Young at al., 1984).

A patterns-of-care study was conducted as part of this project to observe changes in patient care. Based on expert recommendation, a set of ten (10) patterns of care for breast cancer was defined which were intended to be sensitive to the effect of an increase in physician participation in the trials. These were further categorized as "more appropriate" or "less appropriate" according to whether a pattern was expected to occur more (less) frequently over time given an increased physician participation. An example of a pattern is that the clinical size of the tumour should be recorded; this is a "more appropriate" pattern since it was expected to increase over time under the influence of protocol participation. Treatment information during the first four months after diagnosis was obtained from a systematic sample of inpatient and outpatient records of female breast cancer patients. The final data set comprised patients admitted on October 1, 1982 onwards till September 30, 1985. This information was used to determine the pattern of care followed for each patient.

In this study, we have only dealt with 1,856 patients from 19 of the 107 hospitals with the highest percentage (more than 10%) of black patients. As per pattern of care is concerned, we have only concentrated on liver scan which is a specialized radiological procedure used to examine the liver to identify certain conditions or to assess its function. It is basically a nuclear radiological procedure that uses a radiological substance called a radionuclide, which, when absorbed by a liver tissue, emits gamma radiation. The radiation is detected by a scanner which transforms it into an image of the liver. Physicians can measure the behaviour of the radionuclide in the body during a scan and assess and diagnose various conditions such as tumors, abscesses, hematomas, organ enlargement or cysts. Liver scan is a "less appropriate" procedure since "Liver scans and CT scans are not routinely required for a patient with local or regional disease because the likelihood of finding an abnormality is low in the absence of abnormal liver chemistries or hepatomegaly" (Diehr et al., 1989, pg 951, lines 11-15).

We are well aware of the fact that the dataset is old. However, as mentioned in Section 1, the main thrust of the paper is to propose an alternative estimation procedure for the odds ratios based on a Bayesian paradigm that will be computationally more efficient and advantageous than the existing ones.

3 Review of classical statistical approaches

The liver scan data is a classic example of a two-sample problem where each sample is binomial with a given number of successes and failures. Clearly, the samples consist of black and white patients admitted to a particular hospital while success (failure) signifies receiving (not receiving) a liver scan. Thus, the number of successes and failures vary across hospitals. Hence, this data can be conveniently represented in a 2×2 contingency table. For example, Table 1 depicts the data for hospital 1.

TABLE 1: 2×2 contingency table for hospital 1

Hospital 1	Liver scan	No liver scan	Total
Black patients	4	9	13
White patients	12	34	46
Total	16	43	59

Analysing 2×2 tables boils down to testing for the association between two categorical variables, in this case, race and treatment, each having two categories. Thus, we would like to see whether the race of an individual has a significant effect on the chances of his/her receiving a liver scan. This is synonymous to testing for the equality of two proportions i.e the proportion of blacks (p_b) and whites (p_w) receiving a liver scan. Thus, we can test the hypotheses,

$$H_0 : p_b = p_w \quad vs \quad H_a : p_b > p_w, \tag{1}$$

with the alternative stating that the population proportion of blacks receiving a liver scan is higher than that of whites. The above hypotheses can be alternatively stated in terms of odds ratios as

$$H_0 : \Gamma = 1 \quad vs \quad H_a : \Gamma > 1, \tag{2}$$

where $\Gamma = \dfrac{p_b}{1-p_b} \Big/ \dfrac{p_w}{1-p_w}$, the numerator (denominator) respectively being the odds of a black (white) patient receiving liver scan.

Probably, the most well known procedure for performing the above test is the celebrated Pearson chi-square test which is based on comparing the observed and expected counts of the contingency table. However, this test is governed by some strict conditions with regard to the sample size in each cell (specifically, expected sample size in each cell should be no less than 5). However, this condition is often violated as in the case of the above breast

cancer data (expected frequency for cell (1, 1) in Table 1 is 3.53). In situations like this, an analogous test is the Fisher's exact test.

Fisher's exact test of independence conditions on both margins (not naturally fixed) of a 2×2 table leading to a hypergeometric distribution under independence. The "exactness" of the Fisher's test is attributable to the fact that the probabilities of the hypergeometric distribution can be calculated exactly under the hypotheses of independence. However, due to the discreteness of the hypergeometric distribution, it is not possible to obtain a specific significance level (say .05); as a result, both Fisher's exact test and Pearson chi-squared tests are conservative, i.e. their actual significance level falls short of the nominal level. Fisher's exact test has also been criticized due to the fact that its marginal totals are informative, which results in smaller variability and hence, smaller p-values (Berkson, 1978). In this article too, we will use the hypergeometric distribution as a basis of our proposed framework (see Sec 4).

Recently, some research has been carried out towards obtaining exact tests of independence for sparse tables (ie. expected counts less than 5). In one such work (Nandram et al., 2015), a likelihood ratio test of quasi-independence is proposed for sparse two-way contingency tables in which many cells have observed zero counts. It uses a truncated multinomial distribution to model the positive counts and Monte Carlo methods to derive the parameter estimates.

Although there have been some classic works on the Bayesian analysis of contingency tables (Lindley, 1964; Altham, 1969), not much attention has been given to the analysis of 2×2 tables, specially for small domains. However, one study (Hashemi et al., 1997) analyzed this problem in a Bayesian paradigm and obtained highest posterior density intervals (both exact and approximate) for three commonly used measures *viz.* relative risk, odds ratio and attributable risk. Bayes factors were also obtained to test whether the two binomial proportions are similar or not. The methodology was applied to the Worcester Heart Attack Study to test for gender differences in the therapeutic management of patients with acute myocardial infarction by demographic and clinical characteristics.

A situation may as well arise where the above test needs to be carried out for a series of 2×2 contingency tables - in fact, this was one of the limitations of the above study (Hashemi et al., 1997). For example, in our case, it may be of interest to test the hypotheses in (2) for each of the 19 hospitals in question. Clearly this would entail an extension of the Fisher's exact test. One such test for analyzing sets of 2×2 contingency tables is the Mantel-Haenszel test (Mantel and Haenszel, 1959). In order to implement this test, we first need to construct a 2×2 table out of each of the 19 hospitals in question. Each of those tables would allow us to compare the proportions of black and white patients receiving a liver scan in a particular

hospital. Table 2 depicts the 2×2 table for the i^{th} hospital

TABLE 2: 2×2 table for hospital i

	Successes	Failures	Total
Black patients	x_i	$n_{1i} - x_i$	n_{1i}
White patients	$m_i - x_i$	$n_{2i} - m_i + x_i$	n_{2i}
Total	m_i	$N_i - m_i$	N_i

Here m_i is the total number of patients (blacks and whites combined) receiving a liver scan in hospital i, n_{1i} and n_{2i} are respectively the number of blacks and whites admitted to hospital i, while $N_i = n_{1i} + n_{2i}$ is the total number of patients admitted to hospital i, $(i = 1, ..., 19)$. Let $p_b^{(i)}(p_w^{(i)})$ be the probability that a black (white) patient in hospital i receives a liver scan. The Mantel-Haenszel procedure tests the null hypotheses that within each hospital, the probabilities of receiving a liver scan are the same for black and white patients against the alternative that in at least one hospital, a black patient is more likely to receive a liver scan i.e,

$$H_0 \ : \ p_b^{(i)} = p_w^{(i)} \quad i = 1, ..., 19 \ \ vs \ \ H_a : p_b^{(i)} > p_w^{(i)} \quad \text{for at least one } \ i, \ \ i = 1, ..., 19. \tag{3}$$

Alternatively, the above hypotheses can be written in terms of the odds ratios (for hospital i) as follows

$$H_0 : \Gamma_i = 1 \ \ i = 1, ..., 19 \ \ vs \quad H_a : \Gamma_i > 1 \quad \text{for at least one } \ i, \ \ i = 1, ..., 19, \tag{4}$$

where $\Gamma_i = \dfrac{p_b^{(i)}}{1 - p_b^{(i)}} \Big/ \dfrac{p_w^{(i)}}{1 - p_w^{(i)}}$ is the odds ratio for hospital i.

In the present article, we propose an altogether different approach for estimating Γ_i, based on a Bayesian framework. We hope that the proposed framework will be easier to implement and computationally advantageous than the classical approaches detailed above (which often suffer from intractable analytical forms of the test statistics, thus leading to a cumbersome estimation process). The proposed methodology is based on the non-central hypergeometric distribution. One of the important applications of this distribution has been to model and analyze counts in a general two-way contingency table in the presence of covert biases (Rosenbaum, 2002). This is an important problem, specially when the odds ratio is not the parameter of interest. In the next section, we will discuss the non-central hypergeometric distribution.

4 Proposed methodology

4.1 Noncentral hypergeometric distribution

The noncentral hypergeometric distribution arises quite naturally out of a 2×2 contingency table setup and hence we use it as a basis of our proposed Bayesian framework. Let n_1 and n_2 be the number of subjects in the two groups to be compared while the corresponding counts of successes be Y_1 and Y_2 respectively. We assume that $Y_i \sim \text{Bin}(n_i, p_i)$ be independent binomial random variables with success probabilities $p_i, i = 1, 2$. Assuming $Y_1 + Y_2 = m$, Y_1 can be shown to have a noncentral hypergeometric distribution given by

$$P(Y_1 = x | Y_1 + Y_2 = m) = \frac{\binom{n_1}{x}\binom{n_2}{m-x}\Gamma^x}{\sum_{x \in \mathfrak{J}} \binom{n_1}{x}\binom{n_2}{m-x}\Gamma^x}, x \in \mathfrak{J}, \tag{5}$$

where $\mathfrak{J} : \{x : \max(0, m - n_2) \leq x \leq \min(n_1, m)\}$.

In the context of our data set, let $Y_{1i}(Y_{2i})$ be the number of black (white) patients who received a liver scan at the i^{th} hospital ($i = 1, 2, ..., 19$). Thus we have

$$P(Y_{1i} = x_i | Y_{1i} + Y_{2i} = m_i) = \frac{\binom{n_{1i}}{x_i}\binom{n_{2i}}{m_i - x_i}\Gamma_i^{x_i}}{\sum_{x_i \in \mathfrak{J}_i} \binom{n_{1i}}{x_i}\binom{n_{2i}}{m_i - x_i}\Gamma_i^{x_i}}, x_i \in \mathfrak{J}_i, \tag{6}$$

where $\mathfrak{J}_i : \{x_i : \max(0, m_i - n_{2i}) \leq x_i \leq \min(n_{1i}, m_i)\}$ and Γ_i is the odds ratio for the i^{th} hospital given by $\frac{p_{1i}}{1-p_{1i}} / \frac{p_{2i}}{1-p_{2i}}$.

As mentioned before, inferences regarding the difference between the two groups with respect to treatment outcomes (for the i^{th} hospital) are based on the odds ratios Γ_i. If the credible interval of the odds ratios contain 1, that would imply no significant difference in the likelihood of receiving liver scans between black and white patients. On the other hand, if the credible interval contain values greater than 1, that would imply a significantly higher chance of a black patient receiving liver scans compared to whites. The present study is concerned with the estimation of Γ_i given data from the local hospitals. We will also propose a general framework for drawing inferences on Γ_i using the combined (or pooled) hospital data by borrowing strength across hospitals.

4.2 Prior construction

Here the parameter of interest is Γ, the ratio of the odds of blacks and whites receiving a liver scan. Thus, the posterior of Γ would be

$$p(\Gamma|x) \propto p(x|\Gamma)p(\Gamma), \tag{7}$$

where $p(x|\Gamma)$ is given in (5). Thus, in order to implement our proposed methodology, we need to figure out an appropriate prior density of Γ.

Let $p_1(p_2)$ be the population probability of receiving liver scans for black (white) patients respectively. We assume (p_1, p_2) to be independent with $p_i \sim U(0,1), i = 1,2$. Let $y_1 = p_1/(1-p_1)$ and $y_2 = p_2/(1-p_2)$; then the probability density function of y_i can be shown to be

$$f(y_i) = \begin{cases} \dfrac{1}{(1+y_i)^2} & \text{if } 0 < y_i < \infty \\ 0 & \text{if } y_i = 0, i = 1,2. \end{cases}$$

This leads to the cumulative distribution function (cdf) of $\Gamma\left(\dfrac{p_1}{1-p_1}/\dfrac{p_2}{1-p_2}\right)$ to be

$$F_\Gamma(v) = \int_0^\infty \frac{1}{(1+y_2)^2}\frac{vy_2}{(1+vy_2)}dy_2, \tag{8}$$

while the probability density function (pdf) of Γ will be

$$f_\Gamma(v) = \int_0^\infty \frac{1}{(1+y_2)^2}\frac{y_2}{(1+vy_2)^2}dy_2. \tag{9}$$

By transforming y_2 to a bounded variable, say ϕ, where $y_2 = \dfrac{\phi}{1+\phi}$, (9) can be rewritten as

$$f_\Gamma(v) = E\left[\frac{\phi(1-\phi)}{(1-\phi+v\phi)^2}\right], \tag{10}$$

where expectation is taken over $\phi \sim U(0,1)$ and $E(\phi) = 0.5$. Now, using a first order Taylor expansion, we can approximate $f_\Gamma(v)$ as

$$\begin{aligned} f_\Gamma(v) &\approx \frac{E(\phi)E(1-\phi)}{(1-E(\phi)+vE(\phi))^2}, \\ &= \frac{1}{(1+v)^2}, v \geq 0. \end{aligned} \tag{11}$$

Based on the above derivations, we think that $f_\Gamma(v)$ above will be a reasonable candidate for the prior density of Γ.

In order to support the above choice of prior, we numerically compared the Riemann sum of (9) to $f_\Gamma(v)$. The numerical difference between the two was negligible. This was also reflected in the plot of the two quantities which was virtually a straight line with slope 1, as shown in Figure 1. Monte Carlo methods were also used to check our results which were in accordance with the Reimann summation.

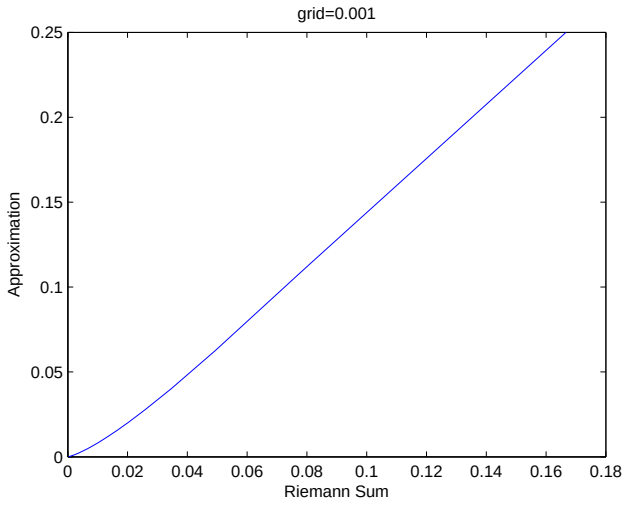

FIGURE 1: Comparison of Riemann sum of exact and approximated prior densities

5 Data analysis

5.1 Hospital specific estimation

As mentioned in Section 1, our principle target of inference are the odds ratios, Γ_i's (measuring the relative likelihood of black and white patients receiving liver scans) for each of the 19 hospitals and of all the hospitals taken together. Towards that purpose, we first explain the hospital specific estimation procedure. Here the goal is to generate samples from the posterior distribution of the Γ_i's and make inferences about the same. The procedure is explained next.

5.1.1 The grid method

As discussed in Section 4.1, the likelihood function of Γ_i $(i = 1, ..., 19)$ has the following form

$$f(x_i|\Gamma_i) = \frac{\binom{n_{1i}}{x_i}\binom{n_{2i}}{m_i - x_i}\Gamma_i^{x_i}}{\sum_{x_i \in \mathfrak{I}_i}\binom{n_{1i}}{x_i}\binom{n_{2i}}{m_i - x_i}\Gamma_i^{x_i}}, x_i \in \mathfrak{I}_i. \tag{12}$$

Combining this with the prior distribution of Γ_i in (11), we have the following posterior distribution of Γ_i

$$f(\Gamma_i|x_i) = \frac{\binom{n_{1i}}{x_i}\binom{n_{2i}}{m_i - x_i}\Gamma_i^{x_i}}{\sum_{x_i \in \mathfrak{I}_i}\binom{n_{1i}}{x_i}\binom{n_{2i}}{m_i - x_i}\Gamma_i^{x_i}}\frac{1}{(1+\Gamma_i)^2}, \quad \Gamma_i > 0. \tag{13}$$

For computational convenience, we transform Γ as $\theta = \Gamma/(1+\Gamma)$ such that $\theta \in (0, 1)$. This results in the following transformed posterior

$$f(\theta_i|x_i) = \frac{\binom{n_{1i}}{x_i}\binom{n_{2i}}{m_i - x_i}\left(\frac{\theta_i}{1-\theta_i}\right)^{x_i}}{\sum_{x_i \in \mathfrak{I}_i}\binom{n_{1i}}{x_i}\binom{n_{2i}}{m_i - x_i}\left(\frac{\theta_i}{1-\theta_i}\right)^{x_i}}, \quad \theta_i \in (0, 1).$$

Since θ is computationally stabler than Γ (being bounded in (0, 1)), hereafter we will use θ for computational purposes. (Like the odds ratio Γ, θ has a straightforward interpretation; for example, $\theta = 0.5$ would imply independence of the two variables).

Since our new parameter of interest θ is bounded in (0, 1), we will take advantage of the grid method to draw a sample of fixed size (say 10,000) from its posterior distribution. Towards that purpose, we first partition the interval (0, 1) into 100 segments. Let r_k be the mid-point of the k^{th} segment and $F(r_k|x)$ be the cumulative probability distribution of r_k. Now we draw a sequence of 10,000 numbers from a uniform distribution, u_k being the k^{th} number. Then we compare each u_k with the cumulative probability $F(r_k|x)$; if $u_k \leq F(r_k|x) < u_{k+1}$, we select r_{k+1}. Since the cdf of a random variable follows a uniform distribution in (0, 1), the sequence of grids r_k generated as above can be assumed to be samples from the posterior distribution $f(\theta|x)$.

Table 3 shows the mean, standard deviation and 95% credible intervals of θ while Figure 2 depict the histograms and the corresponding kernel density estimates of θ for the 19 hospitals. In both the above illustrations, the last row (plot) depicts the respective summaries (density estimates) corresponding to the combined hospital data.

In Table 3, the credible intervals of θ for hospitals 8, 18 and 19 contain values exceeding 0.5. Thus the odds ratios (of receiving a liver scan) at these hospitals are significantly greater

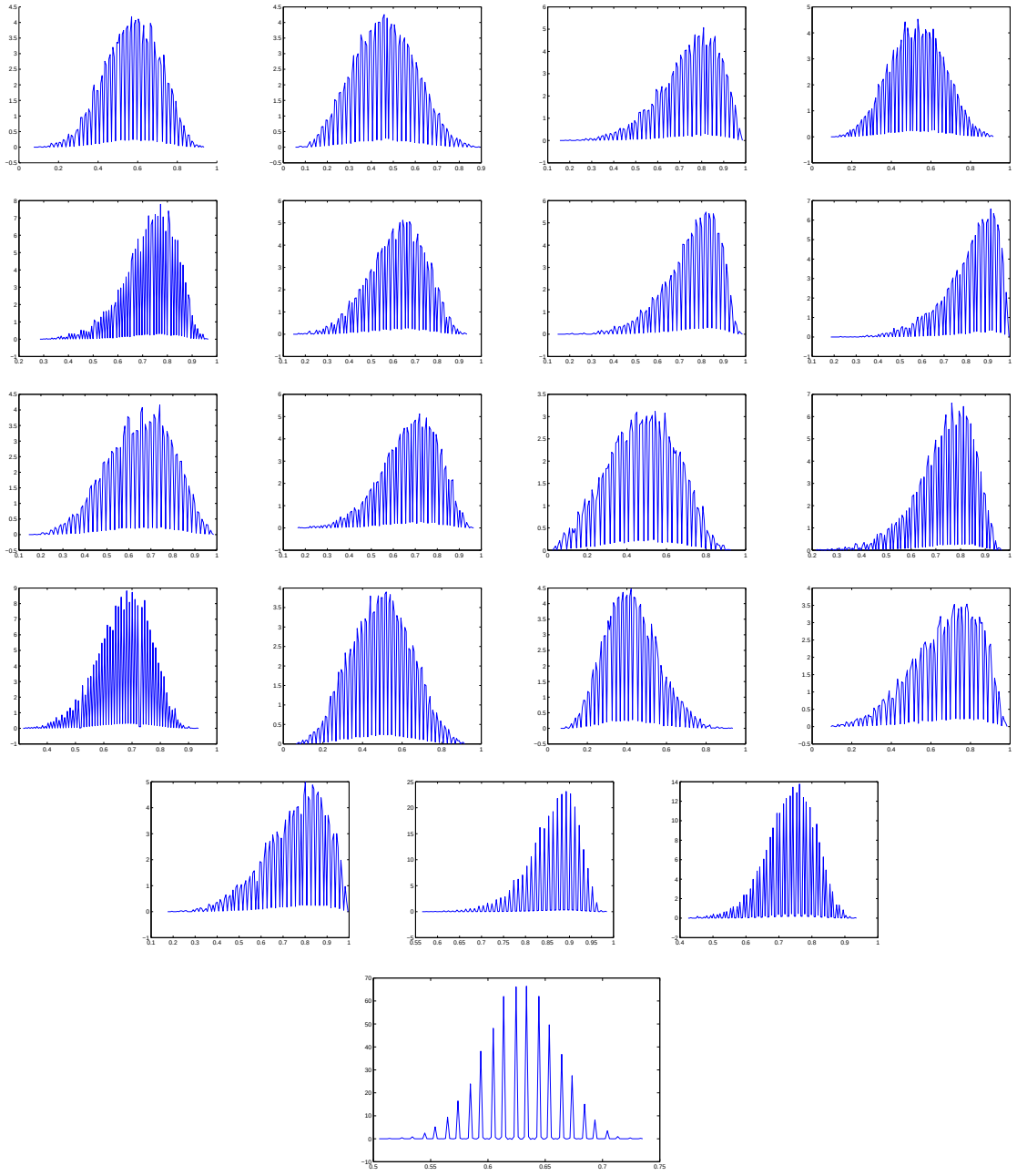

FIGURE 2: Kernel density plots of θ obtained through hospital-specific estimation procedure (last plot for combined hospital data)

TABLE 3: Hospital specific and pooled posterior summaries of θ obtained through the grid method

Hospital	Posterior mean	Standard deviation	95% Credible interval
1	0.573	0.136	(0.295, 0.815)
2	0.463	0.139	(0.195, 0.735)
3	0.746	0.135	(0.435, 0.945)
4	0.532	0.137	(0.215, 0.785)
5	0.732	0.104	(0.495, 0.895)
6	0.620	0.126	(0.355, 0.835)
7	0.758	0.123	(0.465, 0.935)
8	0.825	0.119	(0.525, 0.975)
9	0.652	0.146	(0.345, 0.905)
10	0.677	0.127	(0.395, 0.885)
11	0.480	0.170	(0.145, 0.785)
12	0.739	0.116	(0.465, 0.915)
13	0.673	0.093	(0.475, 0.835)
14	0.499	0.148	(0.215, 0.785)
15	0.426	0.131	(0.195, 0.705)
16	0.678	0.165	(0.315, 0.925)
17	0.757	0.139	(0.445, 0.965)
18	0.863	0.058	(0.725, 0.945)
19	0.736	0.074	(0.575, 0.865)
Overall	0.629	0.032	(0.565, 0.685)

than 1 implying that black patients have a significantly higher odds of receiving a liver scan at any of these hospitals compared to whites. The credible interval for the combined hospital data also exceeds 0.5 indicating the validity of the above conclusion (higher odds of liver scan for blacks) for all the 19 hospitals taken together. The same hospital data was used in another study (Hollander and Wolfe, 1999) to perform a Mantel-Haenszel test. The resulting p-value was very small (less than 0.0002) indicating that the common odds ratio (for all the hospitals combined) is significantly greater than 1 i.e a similar conclusion as above.

5.1.2 Validation of the grid method

It would be of interest to test the validity of the above approach (of generating θ) with a frequentist one; we do so by comparing the coverage probabilities of the intervals obtained above with that obtained using a frequentist method i.e. for each hospital, we generate a large number of 95% credible intervals and look at the proportion of those containing the true θ.

However, before generating the credible intervals, we need to generate contingency tables as well. The mechanism of generating a contingency table for a given θ is similar to the way we generated θ in Sec 5.1.1. As shown in Sec 4.1, for the i^{th} hospital, the observed number of success for group 1 (i.e. x_i) lies in the range $\mathfrak{I}_i : \{x_i : \max(0, m_i - n_{2i}) \leq x_i \leq \min(n_{1i}, m_i)\}$. For each x_i lying in the above range, we calculated the cumulative probability recursively using (12). Thus, for the $(i+1)^{th}$ iteration, we have

$$f(x_{i+1}|\Gamma_i) = \frac{\binom{n_{1i}}{x_{i+1}}\binom{n_{2i}}{m_i - x_{i+1}}\Gamma_i^{x_{i+1}}}{\sum_{x_{i+1}\in\mathfrak{I}_i}\binom{n_{1i}}{x_{i+1}}\binom{n_{2i}}{m_i - x_{i+1}}\Gamma_i^{x_{i+1}}}.$$

Taking the ratio of $f(x_{i+1})$ to $f(x_i)$, we have

$$\frac{f(x_{i+1}|\Gamma)}{f(x_i|\Gamma)} = \frac{n_{1i} - i}{i+1}\frac{m_i - i}{n_{2i} - m_i + i + 1}\Gamma.$$

where $\Gamma = \theta/1 - \theta$. After calculating the cumulative probability for each x_i, we drew 1000 random numbers $u_i, i = 1, 2, ..., 1000$ from $U \sim (0, 1)$. Now using the inverse CDF sampling technique, we generated 1000 new contingency tables for each hospital. From each of the new tables thus generated, we generated 10,000 values of θ and calculated the corresponding 95% credible intervals. We reproduced each interval 1000 times and calculated the proportion of those intervals which contain the true value of θ. The hospital-specific coverage proportions are displayed in Table 4 below.

TABLE 4: Hospital specific coverage probabilities of frequentist intervals

Hospital	1	2	3	4	5	6	7	8	9	10
Coverage	96.0%	78.3%	97.1%	100%	96.2%	98.3%	96.0%	97.5%	99.2%	98.7%
Hospital	11	12	13	14	15	16	17	18	19	Overall
Coverage	96.7%	98.0%	97.1%	94.7%	95.6%	99.4%	98.9%	95.9%	95.9%	96.2%

We notice that for 17 of the 19 hospitals, as well as the combined hospital data, the proportions (of credible intervals containing the true θ) are greater than 95%. For hospital

2, this percentage was unacceptably low at 78.3% while for hospital 14 it is 94.7% which can be approximated to 95%. Thus, except for hospitals 2 and 14, all other credible intervals are too conservative. This indicates that there is a need of a different approach for estimating θ - we achieve this using the pooled hospital data, as explained in the next section.

5.2 Pooled estimation

In the hospital specific estimation procedure, the credible intervals of θ were found to be quite conservative (as per their coverage probabilities). In this section, we carry out a similar Bayesian approach by pooling data across all the hospitals. Our aim in doing so is to improve the estimates for a particular hospital by borrowing strength across other hospitals. Along that line, the modified likelihood function of θ is given by

$$f(x_i|\theta_i) = \frac{\binom{n_{1i}}{x_i}\binom{n_{2i}}{m_i - x_i}\left(\frac{\theta_i}{1-\theta_i}\right)^{x_i}}{\sum\limits_{x_i \in \mathfrak{J}_i}\binom{n_{1i}}{x_i}\binom{n_{2i}}{m_i - x_i}\left(\frac{\theta_i}{1-\theta_i}\right)^{x_i}}, \quad i = 1, 2, ..., 19 \;\; \theta_i \in (0,1), \tag{14}$$

where $\mathfrak{J}_i : \{x_i : \max(0, m_i - n_{2i}) \leq x_i \leq \min(n_{1i}, m_i)\}$. We assume the following prior for θ

$$f(\theta_i|\mu,\tau) = \text{Beta}(\mu\tau, (1-\mu)\tau), \;\; 0 < \mu < 1,$$

and for (μ, τ), we assume the following density

$$f(\mu,\tau) = \frac{1}{(1+\tau)^2}, \;\; \tau > 0.$$

Thus, the joint density of $\boldsymbol{\theta}(= \theta_1, ..., \theta_{19}), \mu, \tau$ given $\boldsymbol{x}(= x_1, ..., x_{19})$ is given by

$$\begin{aligned} f(\boldsymbol{\theta}, \mu, \tau|\boldsymbol{x}) &\propto f(\mu,\tau)\prod_{i=1}^{19}\{f(x_i|\theta_i)f(\theta_i|\mu,\tau)\}, \\ &= \frac{1}{(1+\tau)^2}\prod_{i=1}^{19}\left\{\frac{\binom{n_{1i}}{x_i}\binom{n_{2i}}{m_i - x_i}\left(\frac{\theta_i}{1-\theta_i}\right)^{x_i}}{\sum\limits_{x_i \in \mathfrak{J}_i}\binom{n_{1i}}{x_i}\binom{n_{2i}}{m_i - x_i}\left(\frac{\theta_i}{1-\theta_i}\right)^{x_i}}\frac{\theta_i^{\mu\tau-1}(1-\theta_i)^{(1-\mu)\tau-1}}{B(\mu\tau,(1-\mu)\tau)}\right\}. \end{aligned}$$

Thus the full conditional distribution of θ_i will be as follows

$$f(\theta_i|\mu,\tau,x_i) \propto \frac{\binom{n_{1i}}{x_i}\binom{n_{2i}}{m_i - x_i}\left(\frac{\theta_i}{1-\theta_i}\right)^{x_i}}{\sum\limits_{x_i \in \mathfrak{J}_i}\binom{n_{1i}}{x_i}\binom{n_{2i}}{m_i - x_i}\left(\frac{\theta_i}{1-\theta_i}\right)^{x_i}}\theta_i^{\mu\tau-1}(1-\theta_i)^{(1-\mu)\tau-1}, \;\; i = 1, 2, ..., 19, \tag{15}$$

where given (μ, τ, x_i), θ_i's are assumed to be independent for $i = 1, 2, ..., 19$.
The full conditionals of μ and τ will be as follows

$$f(\mu|\tau, \boldsymbol{\theta}, \boldsymbol{x}) \propto \prod_{i=1}^{19} \frac{\theta_i^{\mu\tau-1}(1-\theta_i)^{(1-\mu)\tau-1}}{B(\mu\tau, (1-\mu)\tau)}, \ 0 < \mu < 1, \tag{16}$$

$$f(\tau|\mu, \boldsymbol{\theta}, \boldsymbol{x}) \propto \frac{1}{(1+\tau)^2} \prod_{i=1}^{19} \frac{\theta_i^{\mu\tau-1}(1-\theta_i)^{(1-\mu)\tau-1}}{B(\mu\tau, (1-\mu)\tau)}, \ \tau > 0. \tag{17}$$

For computational convenience and stability, we transform τ to [0, 1] as $\phi = \tau/1+\tau$ resulting in the modified conditional

$$f(\phi|\mu, \boldsymbol{\theta}, \boldsymbol{x}) \propto \prod_{i=1}^{19} \frac{\theta_i^{\mu\phi/(1-\phi)-1}(1-\theta_i)^{(1-\mu)\phi/(1-\phi)-1}}{B(\mu\phi/(1-\phi), (1-\mu)\phi/(1-\phi))}, 0 < \phi < 1. \tag{18}$$

5.2.1 Posterior sampling

Since the conditional density of θ given in (15) has a complicated form, it is not computationally efficient to generate samples from it using the grid method. This prompted us to experiment with the Metropolis-Hastings (M-H) algorithm.

For the purpose of M-H algorithm, the target densities are the 19 conditional distributions of θ_i given $\theta_j (i \neq j), \mu$ and τ while the corresponding proposal densities are the approximated posterior distributions of θ_i's given by

$$\begin{aligned} f(\theta_i|\mu, \tau, \boldsymbol{x}_i) &= \text{Beta}(\mu_0\tau_0, (1-\mu_0)\tau_0)\theta_i^{\mu\tau-1}(1-\theta_i)^{(1-\mu)\tau-1}, \\ &= \frac{\theta_i^{\mu_0\tau_0-1}(1-\theta_0)^{(1-\mu_0)\tau_0-1}}{B(\mu_0\tau_0, (1-\mu_0)\tau_0)}\theta_i^{\mu\tau-1}(1-\theta_i)^{(1-\mu)\tau-1}, \\ &= \text{Beta}(\mu_0\tau_0 + \mu\tau - 1, (1-\mu_0)\tau_0 + (1-\mu)\tau - 1), i = 1, ..., 19. \end{aligned} \tag{19}$$

The griddy M-H sampler iteratively uses the full set of univariate conditionals which eventually converge to the true posterior distribution. At each iteration, posterior samples of $\theta_i (i = 1, ..., 19)$ are drawn. In fact, the grid method embedded within this procedure enabled us to draw samples of μ and τ more efficiently. For the purpose of implementing the griddy M-H sampler, we used the univariate conditionals as given in (15)-(17), while rewriting (16) and (17) as

$$\begin{aligned} f(\mu|\tau, \boldsymbol{\theta}, \boldsymbol{x}) &= \left[\frac{G_1^{\mu\tau-1}G_2^{(1-\mu)\tau-1}}{B(\mu\tau, (1-\mu)\tau)}\right]^{19}, 0 < \mu < 1, \\ f(\tau|\mu, \boldsymbol{\theta}, \boldsymbol{x}) &= \frac{1}{(1+\tau)^2}\left[\frac{G_1^{\mu\tau-1}G_2^{(1-\mu)\tau-1}}{B(\mu\tau, (1-\mu)\tau)}\right]^{19}, \tau > 0, \end{aligned}$$

where G_1 and G_2 are respectively the geometric means of θ_i and $(1-\theta_i)$ respectively $(i = 1, ..., 19)$.

TABLE 5: Posterior summaries of θ obtained through pooled data estimation

Hospital	Posterior mean	Standard deviation	95% Credible interval
1	0.557	0.099	(0.360, 0.741)
2	0.500	0.099	(0.305, 0.688)
3	0.633	0.098	(0.422, 0.808)
4	0.536	0.098	(0.345, 0.723)
5	0.657	0.083	(0.482, 0.805)
6	0.582	0.095	(0.391, 0.755)
7	0.566	0.100	(0.366, 0.756)
8	0.654	0.102	(0.443, 0.835)
9	0.587	0.101	(0.380, 0.773)
10	0.610	0.097	(0.410, 0.789)
11	0.520	0.110	(0.303, 0.722)
12	0.643	0.097	(0.439, 0.814)
13	0.631	0.077	(0.478, 0.773)
14	0.522	0.101	(0.324, 0.717)
15	0.519	0.102	(0.317, 0.714)
16	0.647	0.090	(0.460, 0.809)
17	0.641	0.091	(0.454, 0.810)
18	0.776	0.057	(0.655, 0.876)
19	0.686	0.065	(0.553, 0.805)
Overall	0.662	0.031	(0.602, 0.726)

5.2.2 Analytical results

Table 5 shows the estimates and credible intervals of θ for each hospital and across all hospitals obtained from the above mentioned pooled data estimation technique. The posterior mean of θ is greater than 0.5 for all but hospital 2, for which it is exactly 0.5. Moreover, the 95% credible intervals of θ were significant (i.e contains values exceeding 0.5) for hospitals 18, 19 and for the combined data. Since a value of 0.5 for θ implies a value of 1 for the odds ratio $\Gamma = 1$ (i.e. independence), we conclude that overall the odds of a black patient receiving liver scan is significantly higher than whites.

Two important things to notice while running an M-H algorithm are the acceptance rate

TABLE 6: Acceptance rates (AR) for M-H algorithm for pooled data estimation

Hospital	1	2	3	4	5	6	7	8	9	10
AR	0.836	0.789	0.835	0.841	0.361	0.863	0.575	0.660	0.835	0.796
Hospital	11	12	13	14	15	16	17	18	19	
AR	0.766	0.437	0.449	0.820	0.803	0.810	0.837	0.058	0.457	

and auto-correlation of the simulated data. As a rule of thumb, the acceptance rates should lie within 0.5 and 0.75. Table 6 depicts the hospital-specific acceptance rates those were obtained from the pooled data estimation technique. Figure 3 depicts the autocorrelation plots of the sampled values of θ obtained using 20 lags. Based on this, it is evident that independence has been achieved. The trace plots of θ's (not shown here) also corroborate convergence of the sampled θ's.

5.3 Comparison of hospital-specific and pooled data

In order to compare the hospital specific and pooled data estimation methods, we examined the 95% credible intervals of the odds ratio Γ for each of the 19 hospitals generated by each of these procedures. Table 7 has the details. It is quite obvious from the results that the pooled method leads to narrower intervals for Γ. For example, the pooled intervals corresponding to hospitals 8 and 17 are respectively (0.8087, 5.3778) and (0.7462, 4.4577) while the same for the hospital specific procedure are much wider ([1.1053, 39] and [0.77, 27.5714]) respectively. A similar trend was noticed for the other hospitals too.

An obvious reason for the shortening of the credible intervals based on the pooled method is the lessening of the values of standard deviations corresponding to this method. Table 8 displays the percent difference of the standard deviations between the hospital specific and pooled methods. Clearly, the pooled method resulted in smaller standard deviations of the means (of Γ) and thus lesser variability compared to the hospital-specific estimation procedure. This supports the notion that the pooled estimation procedure is more precise (hence reliable) than the hospital-specific one.

For both methods however, the posterior means of Γ for all but hospital 15 are greater than 1, indicating that some of the odds ratios may in fact be greater than 1. Based on the intervals from the hospital-specific estimation procedure, hospitals 5, 8, 18, 19 and the combined (1-19) data had odds ratios significantly greater than one. On the other hand, the pooled data estimation method yielded only 3 intervals (for hospitals 18, 19 and the combined data) for which the odds ratios were significantly greater than one. Since the

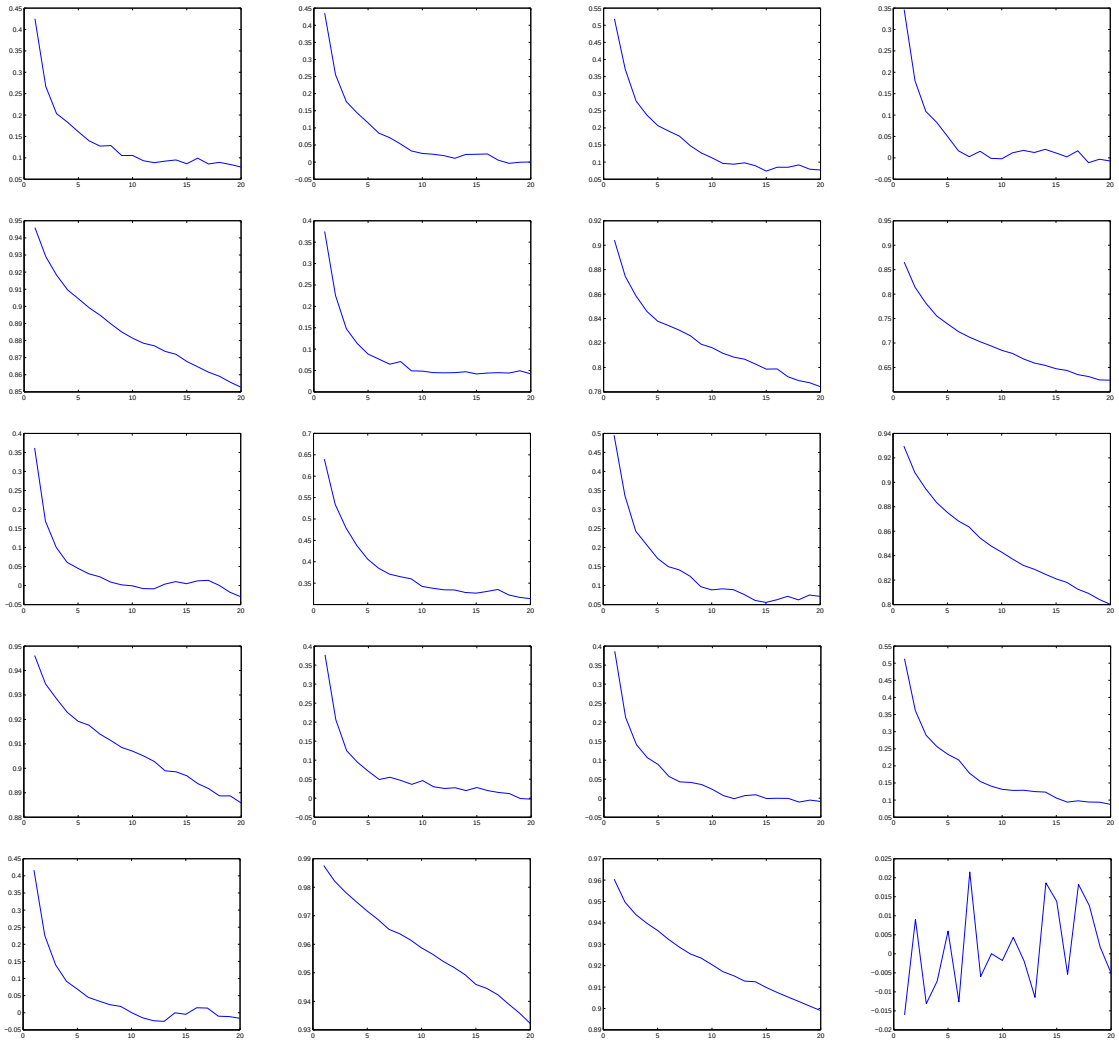

FIGURE 3: Autocorrelation plots of θ obtained through pooled data estimation procedure (last plot is for combined hospital data)

pooled estimation method seems to have lesser variability (hence higher reliability), it is likely that only hospitals 18 and 19 were significantly more likely to give liver scans to blacks than to whites. More importantly however, since the combined data (hospitals 1-19 inclusive) resulted in odds ratios significantly greater than 1, we can conclude that overall black patients had higher odds of receiving liver scans than whites. Finally Figure 4 is a pictorial representation of a comparison of estimates of Γ obtained by the two methods.

TABLE 7: Means, standard deviations and 95% credible intervals of Γ corresponding to the hospital-specific and pooled data methods

	Mean		Standard Deviation		95% Credible Interval	
Hospital	Specific	Pooled	Specific	Pooled	Specific	Pooled
1	1.647	1.404	1.041	0.610	(0.418, 4.405)	(0.578, 2.876)
2	1.058	1.085	0.673	0.475	(0.290, 2.774)	(0.416, 2.246)
3	4.752	1.947	5.730	1.013	(0.770, 17.182)	(0.702, 4.450)
4	1.391	1.276	0.911	0.571	(0.399, 3.878)	(0.506, 2.710)
5	3.384	2.104	2.044	0.901	(1.062, 8.524)	(0.892, 4.294)
6	2.047	1.551	1.280	0.666	(0.626, 5.452)	(0.645, 3.187)
7	4.466	2.064	3.579	1.038	(0.835, 14.385)	(0.719, 4.660)
8	10.191	2.268	17.845	1.211	(1.105, 39.000)	(0.809, 5.378)
9	2.726	1.592	3.068	0.785	(0.527, 9.526)	(0.616, 3.561)
10	2.707	1.761	1.879	0.800	(0.653, 7.696)	(0.684, 3.755)
11	1.180	1.213	0.958	0.581	(0.183, 3.651)	(0.452, 2.657)
12	3.761	2.087	2.473	1.000	(0.905, 10.765)	(0.766, 4.559)
13	2.337	1.853	1.050	0.687	(0.905, 5.061)	(0.844, 3.489)
14	1.235	1.206	0.916	0.546	(0.274, 3.444)	(0.474, 2.559)
15	0.853	0.998	0.558	0.438	(0.242, 2.279)	(0.409, 2.085)
16	3.416	1.686	3.723	0.911	(0.439, 12.333)	(0.587, 3.969)
17	6.002	1.958	10.993	1.011	(0.770, 27.571)	(0.746, 4.458)
18	7.834	3.871	4.318	1.638	(2.774, 17.182)	(1.704, 7.825)
19	3.114	2.357	1.270	0.793	(1.353, 6.407)	(1.183, 4.285)
1-19	1.717	2.301	0.238	0.195	(1.299, 2.279)	(1.251, 2.206)

TABLE 8: Percent difference in standard deviations between pooled data and hospital-specific estimation procedures

Hospital	1	2	3	4	5	6	7	8	9	10
% Diff	41.4	29.4	82.3	37.3	55.9	48.0	71.0	93.2	74.4	57.4
Hospital	11	12	13	14	15	16	17	18	19	1-19
% Diff	39.4	59.6	34.6	40.4	21.5	75.5	90.8	62.1	37.6	18.1

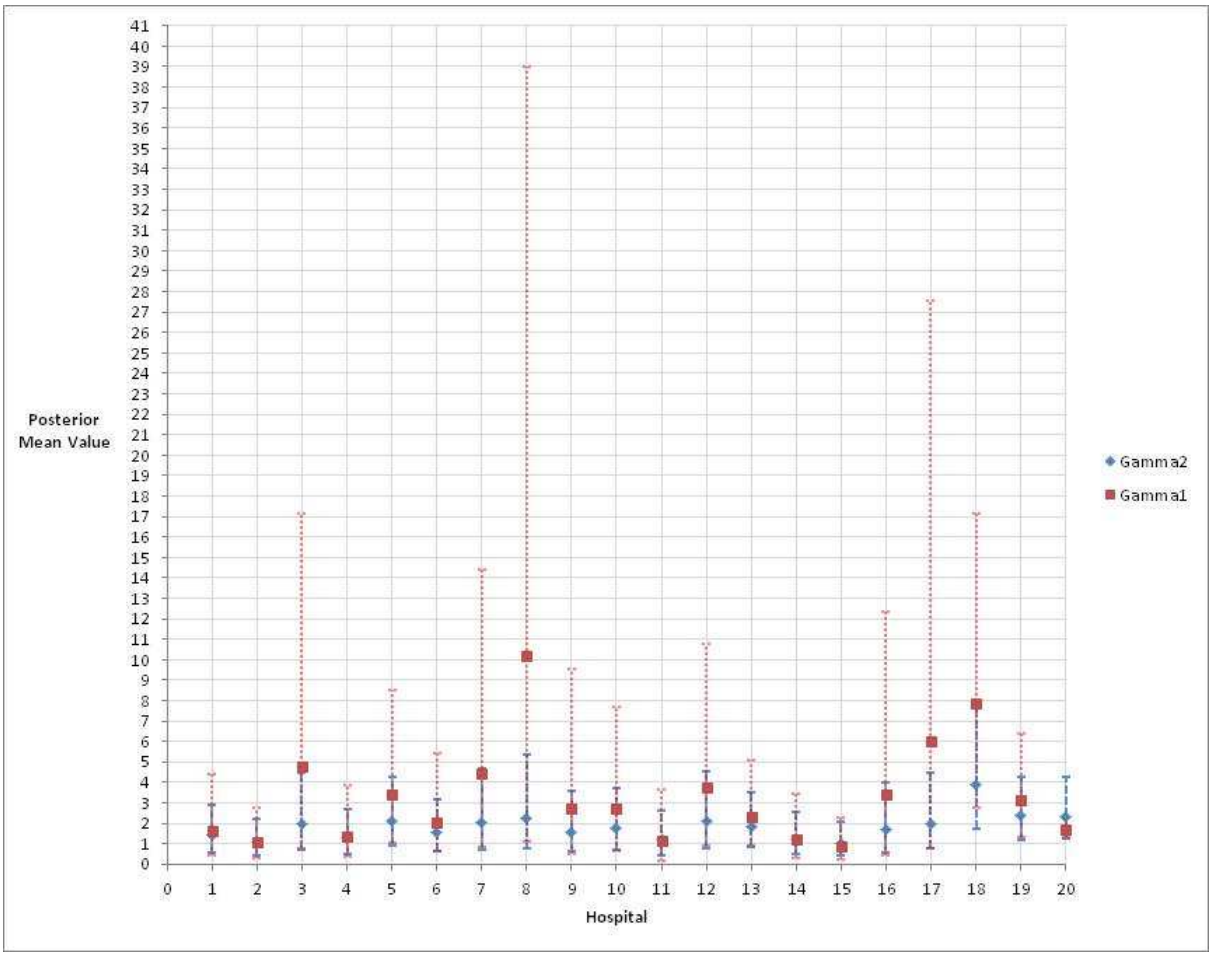

FIGURE 4: Pictorial comparison of the estimates of Γ obtained by the two methods)

6 Discussion

In this study, we have revisited and reanalyzed a well documented problem dealing with the perceived difference in survival rates between black and white breast cancer patients admitted to community hospitals in the US. Research done thus far have established that black breast cancer patients tend to have lower survival rates compared to their white counterparts. One of the principal reasons behind that is the significantly lower quality and quantity of care (or patterns of care, like treatment and diagnostic procedures) available to the blacks compared to whites. In this study, we look at this problem from a different perspective in terms of treating the hospitals as small areas and using a Bayesian paradigm to analyze the same.

Over the last decade or so, small area estimation has slowly but steadily emerged as a distinct statistical estimation procedure to tackle problems, mainly in survey sampling, where the domain-specific sample sizes are too small for classical statistical approaches to yield

regular survey based estimates of adequate precision. This procedure is now being extensively used in Government agencies across the world (and specifically in US) for developing estimates of various socio-economic characteristics (household level income for instance) for states, counties or high school districts across the US. In fact, small area estimates like these are increasingly playing a vital role in the formulation of critical policies by the Government and for the administration of federal funds to local jurisdictions. Due to the small sample sizes, a popular approach in this estimation procedure is to "borrow strength" from neighbouring areas to improve the accuracy of estimates for a given area. This entails the need to develop and incorporate alternative estimation procedures, one of them being the hierarchical Bayesian methodology, which we have used in this study.

The data used in our study goes back to a program funded by the National Cancer Institute in 1983 with regard to increasing physician participation in community hospitals in national clinical trials. It originally dealt with 7,781 breast cancer patients admitted to 107 community hospitals in US between 1982 and 1985. However, in our present study, we have worked with a subset of that, comprising of 1,856 patients from 19 (of the 107) hospitals with the highest percentage (more than 10%) of black breast cancer patients.

The main contribution of our work lies in formulating two distinct modeling frameworks for estimating the odds ratios (of receiving liver scan) for blacks and whites. The first of these models uses hospital-specific information while the second one uses pooled hospital data by borrowing strength across neighbouring hospitals. In doing so, we have used the non-central hypergeometric distribution as the basis for modeling the likelihood function of success (receiving a liver scan) for the blacks and whites. Finally, estimation has been carried out through a Bayesian route while a griddy Metropolis-Hastings sampler has been used for sampling. To our knowledge, such an analysis using the non-central hypergeometric distribution has not been done before.

Our findings confirm some of the claims made in previous literature (Diehr et al., 1989 for instance) that black breast cancer patients were more likely to receive a liver scan than the whites, when a liver scan is deemed as a less appropriate pattern of care. This was evident as both the modeling frameworks yielded odds ratios significantly higher than 1 for some of the hospitals considered in the study. In fact, the pooled data estimation procedure seemed to be more reliable since it consistently produced tighter credible intervals of the odds ratios for all the hospitals. This procedure (i.e. the idea of pooling) is based on the premise that data within each hospital are similar (or homogeneous). Thus, pooling was done to reduce variation, thereby improving the quality of information extracted from the data itself. Having said that, there are potential shortcomings too. For example, pooling tends to draw all the values closer to the mean while those with higher variability receiving a stronger

pull. This is one of the issues to be careful about while applying this procedure. However, as mentioned before, the aim of our study was to formulate a novel modeling framework to estimate the odds of receiving liver scans and subsequently comparing those. Based on the results we have got, we believe that we have been able to achieve our goals.

7 Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

8 References

Altham, P. M. E. (1969). Exact Bayesian analysis of a 2×2 contingency table, and Fisher's "Exact" significance test. *Journal of the Royal Statistical Society, Series B (Methodological)* **31**, 261-269.

Axtell, L. M. and Myers, M. H. (1978). Contrasts in survival of black and white cancer patients 1960-73. *Journal of the National Cancer Institute* **60**, 1209-1215.

Berkson, J. (1978). In dispraise of Fisher's exact test: do the marginal totals of the 2×2 table contain relevant information respecting the table proportions. *Journal of Statistical Planning and Inference* **2**, 27-42.

Diehr, P., Yergan, J., Chu, J., Feigl, P., Glaefke, G., Moe, R., Bergner, M. and Rodenbaugh, J. (1989). Treatment modality and quality differences for black and white breast cancer patients treated in community hospitals. *Medical Care* **27**, 942-958.

Flaherty, J. A. and Meagher, R. (1980). Measuring racial bias in inpatient treatment. *Amer ican Journal of Psychiatry* **137**, 679-682.

Ghosh, M. and Rao, J. N. K. (1994). Small area estimation : An apprisal. *Statistical Science* **9**, 55-76.

Hashemi, L., Nandram, B. and Goldberg, R. (1997). Bayesian analysis for a single 2×2 table. *Statistics in Medicine* **16**, 1311-1328.

Hollander, M. and Wolfe, D. A. (1999). Nonparametric statistical methods, 2nd Edition. *Wiley Interscience.*

Lindley, D. V. (1964). The Bayesian analysis of contingency tables. *Annals of Mathematical Statistics* **35**, 1622-1643.

Mantel, N. and Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute* **22**, 719-748.

McWhorter, W. P and Mayer, W. J. (1987). Black/white differences in type of initial breast

cancer treatment and implications for survival. *American Journal of Public Health* **77**, 1515-1517.

Nandram, B., Bhatta, D. R. and Bhadra, D. (2015). A likelihood ratio test of quasi-independence for sparse two-way contingency tables. *Journal of Statistical Computation and Simulation* **85**, 284-304.

Rao, J. N. K. (2003). Small area estimation : Methods and Applications. *John Wiley.*

Rosenbaum, P. R. (2002). Observational studies, 2nd Edition. *Springer.*

Shavers, V. L. and Brown, M. L. (2002). Racial and ethnic disparities in the receipt of cancer treatment. *Journal of the National Cancer Institute* **94**, 334-357.

Shavers, V.L., Harlan, L. C. and Stevens, J. L. (2002). Racial/ethnic variation in clinical presentation, treatment, and survival among breast cancer patients under age 35. *Cancer* **97**, 134-147.

Smigal, C., Jemal, A., Ward, E., Cokkinides, V., Smith, R., Howe, H. L. and Thun, M. (2006). Trends in breast cancer by race and ethnicity : update 2006. *CA: A Cancer Journal for Clinicians* **56**, 168-183.

Yergan, J., Flood, A. B., LoGerfo, J. P. and Diehr, P. (1987). Relationship between patient race and intensity of hospital services. *Medical Care* **25**, 592-603.

Young, J. L., Ries, L. G. and Pollock, E. S. (1984). Cancer patient survival among ethnic groups in the United States. *Journal of the National Cancer Institute* **73**, 341-352.