Research Article

Statistics in Medicine

Received 17 May 2010, Accepted 28 February 2011 Published online 17 May 2011 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.4249

Defining and ranking effects of individual agents based on survival times of cancer patients treated with combination chemotherapies

Peter F. Thall,[a*†] Diane D. Liu,[a] Su G. Berrak[b] and Johannes E. Wolff[c]

An important problem in oncology is comparing chemotherapy (chemo) agents in terms of their effects on survival or progression-free survival time. When the goal is to evaluate individual agents, a difficulty commonly encountered with observational data is that many patients receive a chemo combination including two or more agents. Because agents given in combination may interact, quantifying the contribution of each individual agent to the combination's overall effect is problematic. Still, if on average combinations including a particular agent confer longer survival, then that agent may be considered superior to agents whose combinations confer shorter survival. Motivated by this idea, we propose a definition of individual agent effects based on observational survival data from patients treated with many different chemo combinations. We define an individual agent effect as the average of the effects of the chemo combinations that include the agent. Similarly, we define the effect of each pair of agents as the average of the effects of the combinations including the pair. Under a Bayesian regression model for survival time in which the chemo combination effects follow a hierarchical structure, these definitions are used as a basis for estimating the posterior effects and ranks of the individual agents, and of all pairs of agents. The methods are illustrated by a data set arising from 224 pediatric brain tumor patients treated with over 27 different chemo combinations involving seven chemo agents.

Keywords: Bayesian analysis; brain tumors; hierarchical model; ranking; survival analysis

1. Introduction

An important problem in oncology is comparison of chemotherapy (chemo) agents in terms of their effects on overall survival (OS) time or progression-free survival time. A difficulty often encountered when addressing this problem based on observational data is that while some patients are treated with a single agent, others receive a chemo combination including two or more agents. Because agents given in combination may interact, quantifying the contribution of each individual agent to the combination's overall effect is problematic. Still, if on average the combinations that include a particular agent are associated with longer survival time compared with combinations that do not include the agent then, intuitively, it seems likely that the common agent may be what is conferring longer survival. In this paper, we formalize this idea in order to define, estimate, and rank the effects of individual agents based on survival time data in which a large number of different combination chemotherapies were used. We use a similar approach for pairs of agents given together as part of some combination.

[a]*Department of Biostatistics, M.D. Anderson Cancer Center, Houston, TX, U.S.A.*
[b]*Department of Pediatrics, The Children's Hospital at Monmouth Medical Center, Long Branch, NJ 07740, U.S.A.*
[c]*Division of Hematology-Oncology, Tufts Medical Center, Boston, MA, U.S.A.*
[*]*Correspondence to: Peter F. Thall, Department of Biostatistics, M.D. Anderson Cancer Center, Houston, TX, U.S.A.*
[†]*E-mail: rex@mdanderson.org*

We are motivated by an observational data set, obtained by a literature review, on 224 patients with choroid plexus carcinoma (CPC), the most malignant subtype of choroid plexus tumor, a family of rare brain tumors. CPCs occur most often in children; the median age of the patients in the data set was 1.75 years, with range 0–64. The five-year OS rate of patients with CPC is about 40 per cent [1], and surviving patients often have long-term sequelae that include reduced psychomotor function as a result of the disease or treatment [1–3]. Because CPCs are rare, organizing a large-scale prospective randomized clinical trial is not considered feasible. Owing to the lack of prospective clinical trials, observational data obtained from single case experiences are the only available source of information on the clinical efficacy of various chemo combinations that have been used to treat CPCs. Consequently, CPC patients are treated following the personal judgments and opinions of attending physicians. Such expert opinion typically is generated from experience with other types of pediatric brain tumors, which may not be necessarily a valid basis for choosing treatments for CPC.

The literature review was conducted in several stages following specific questions building upon each other and thereby growing the database and validating previous entries [4–6]. Pubmed was the main data source using 'choroid plexus, CPC, CPP, APP' as search words. Every publication in which individual patients could be identified was included, excluding only those in which the same patients were described in multiple publications from the same group to avoid counting patients more than once. Variables in the final data set included year of publication, journal and page number, country of origin, age, gender, tumor location, histological grade, surgery result, radiation dose and radiation field, chemotherapy drugs, and OS. For our analyses, countries were combined to form country categories. Other variables that were originally intended to be included but could not be extracted from the literature in a meaningful way included chemo dose schedule, quality of life, and event-free survival time.

While many different treatments are used for CPCs, therapy typically includes some combination of surgery, radiation, and one or more chemo agents. Slightly over half (122, 54 per cent) of the patients in our data set received some form of chemo, consisting of either a single agent or a combination. The remaining 102 patients received surgery, radiation, or both, but no chemo. Each chemo combination consisted of some subset of seven individual chemo agents commonly used to treat this disease, given in Table I. The agents are denoted by CARBO=Carboplatin, CYC=Cyclophosphamide, VP16=Etoposide, IFOS=Ifosfamide, PRC=Procarbazine, VCR=Vincristine, and CDDP=Cisplatin. 'Unk' denotes that it is unknown whether the agent was included in the combination. Table I shows that a total of 23 different chemo combinations with all agents identified were given. For 7 patients, the agents comprising the combination that they received were only partially known, and for 32 patients it was known that they received chemo, but no agents were identified. Of the 27 combinations having at least one agent known, 23 (85 per cent) were given to five or fewer patients. In addition to chemo combination and indicators for radiation and surgery category, patient covariates included age, gender, tumor location, geographic location, and publication year of the paper reporting the data. An additional complication was that covariate data were partially missing for 19 (8.5 per cent) of the 224 patients.

We define each individual agent effect as the average of the effects of the combinations in which it was included. Similarly, we define the effect of each pair of agents given together as the average of the effects of all combinations including the pair. To obtain estimates and rankings, we first fit a Bayesian regression model for survival time [7] in which a hierarchical structure [8, 9] is assumed for the effects of the chemo combinations. The fact that this model borrows strength across the chemo combinations is particularly useful for this data set because most of the chemo combinations were given to very few patients. The posteriors of the chemo combination effects thus provide posteriors for the ranks of the individual agent effects, and of the agent pair effects. We also assume a hierarchical structure for the country category effects, although these are of secondary interest.

In Section 2, we describe the data structure and hierarchical survival time regression model. Definitions of the effects of individual agents and agent pairs are given in Section 3. Corresponding definitions of the ranks of these effects are given in Section 4. In Section 5, we apply the methods to the CPC data, including prior specification, imputation of missing values, and goodness-of-fit analyses. The rankings of individual agents and agent pairs are given in Section 6, including sensitivity analyses to assess the effects on the estimated posterior ranks of prior parameter values and the hierarchical structure assumed for the chemo combination effects. We close with a discussion in Section 7.

Table I. Chemo combinations used to treat the 224 CPC patients.

Comb.	VP16	VCR	CDDP	CARBO	CYC	IFOS	PRC	No. of patients	No. of deaths
0	−	−	−	−	−	−	−	102	62
1	−	−	−	+	−	−	−	1	0
2	−	−	−	+	−	+	−	2	2
3	−	−	+	−	−	−	−	1	0
4	−	+	−	−	−	−	−	7	4
5	−	+	−	−	−	−	+	1	0
6	−	+	−	−	+	−	−	2	1
7	−	+	−	+	−	−	+	1	0
8	−	+	+	−	−	−	−	1	1
9	−	+	+	−	+	−	−	1	0
10	−	+	+	−	+	−	+	4	1
11	−	+	+	+	+	−	−	1	0
12	+	−	−	−	−	−	−	1	1
13	+	−	−	+	−	−	−	4	3
14	+	−	−	+	−	+	−	14	6
15	+	−	−	+	+	−	−	5	0
16	+	−	+	−	−	−	−	4	2
17	+	−	+	−	−	+	−	2	2
18	+	+	−	−	+	−	−	1	0
19	+	+	−	+	−	+	−	1	0
20	+	+	+	−	+	−	−	14	6
21	+	+	+	−	+	+	−	1	1
22	+	+	+	+	+	−	−	1	1
23	+	+	+	+	+	−	+	13	7
24	+	Unk	Unk	+	Unk	+	Unk	4	4
25	Unk	+	+	+	+	Unk	Unk	1	0
26	Unk	+	+	Unk	+	Unk	+	1	1
27	Unk	+	Unk	Unk	Unk	Unk	Unk	1	1
28	Unk	Unk	Unk	Unk	Unk	Unk	Unk	32	12
N_a	13	17	13	12	12	6	5		

Inclusion of an agent in a combination is denoted by a '+' and absence by a '−'. The combination index 0 corresponds to patients who did not receive chemotherapy. The number of combinations including each agent $a = 1, \ldots, 7$ denoted by is N_a.

2. Data structure and survival time models

Let T denote OS time and T° the observed time to death or right censoring, with $\delta = I(T^{\circ} = T)$. As noted above, a complication with this data set is that, for the 39 patients who received the combinations numbered 24–28 in Table I, the particular single agents in the combination that they received were either partially or completely unknown. To account for this, we index patients by $i = 1, \ldots, n$ ($n = 224$), and combinations by $c = 0, 1, \ldots, K+1$ ($K = 27$), and define the indicator $W_{i,c}$ that patient i received chemo combination c, where $c = 0$ corresponds to no chemo and $c = K+1$ indexes the case where it is known that the patient received chemo but no agents were identified. The chemo combination indicator vector of patient i is denoted by $\boldsymbol{W}_i = (W_{i,0}, \ldots, W_{i,K+1})$, which has one entry 1 and all other entries 0 (Table I).

The countries in which the studies were conducted were grouped into 13 categories in order to obtain meaningful numbers. Geographical locations and cultures of leading schools of medical care were used as guidelines in this grouping. For example, USA and Canada were combined as 'North America', Mexico, Central and South American countries were combined as 'South America', Austria and Germany were combined as 'Germanic', and Sweden, Finland, and Norway were combined as 'Nordic'. While the UK–Israel–India category may seem odd in that its three countries are not geographically close, this group reflects the fact that the medical oncology environments for treating pediatric brain tumors in these three countries are very similar. The country categories are given in Table II, which includes estimated median survival time and 1-year survival probability for each country category, and also shows that the country categories of 10 patients were unknown. For $r = 1, \ldots, R$ ($R = 13$), we define the indicator $X_{i,r}$ that patient i was in a study from country category r, with $X_{i,R+1}$ indicating

Table II. Distribution of 224 choroid plexus carcinoma patients, and estimated median and one-year survival probabilities, by country category.

			Survival estimates (years)*		
Country category	No. of patients	Percent	Times†	Median	Pr($T > 1$ year)
Australia	1	0.45	4.5	4.5	—
Slavic	7	3.13		0.9	0.43
Denmark	1	0.45	1.0	1.0	—
Far East	15	6.70		—	0.66
France	32	14.29		2.1	0.74
Germanic	28	12.50		4.0	0.70
North America	78	34.82		4.8	0.75
Nordic	2	0.89	11.0^+, 11.1^+	—	—
Poland	1	0.45	1.0^+	—	—
South America	8	3.57		0.9	0.38
Southern Europe	12	5.36		7.6	0.58
Turkey	2	0.89	0.7^+, 1.0^+	—	—
UK–Israel–India	27	12.05		0.4	0.37
Missing	10	4.46		1.1	0.50

*Values that could not be computed are represented by '—'.

†For country categories with 1 or 2 subjects, the times of death are given by T and the times of right censoring by T^+.

that the country category was missing. We denote the country category indicator vector of patient i by $\boldsymbol{X}_i = (X_{i,1}, \ldots, X_{i,R+1})$, which has one entry 1 and all other entries 0. We denote by $\boldsymbol{Z}_i = (Z_{i,1}, \ldots, Z_{i,q})$ the vector of all baseline prognostic covariates and treatment variables of patient i not accounted for by either $\boldsymbol{W}_i$ or $\boldsymbol{X}_i$. In the sequel, for brevity we will suppress the index i when no meaning is lost.

To obtain a good fit to the CPC data, we consider several different parametric distributions for the survival time regression model of $[T \mid \boldsymbol{Z}, \boldsymbol{X}, \boldsymbol{W}, \boldsymbol{\theta}]$, where $\boldsymbol{\theta}$ is the vector of model parameters. In each model, we assume that the linear term of patient i takes the form

$$\eta_i = \sum_{j=1}^{q} \beta_j Z_{i,j} + \sum_{c=1}^{K} \gamma_c W_{i,c} + \gamma_{K+1} W_{i,K+1} + \sum_{r=1}^{R} \alpha_r X_{i,r} + \alpha_{R+1} X_{i,R+1}. \tag{1}$$

The elements of the vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_q)$ are covariate parameters. The subgroup of 102 patients who did not receive any chemo ($c = 0$) is employed as a baseline group for evaluating chemo combination effects. Thus, if $W_{i,0}$=1, indicating that the patient received no chemo, then $W_{i,1} = \cdots = W_{i,K} = 0$ and $\sum_{c=1}^{K} \gamma_c W_{i,c} = 0$; hence, no γ_c appears in η_i. That is, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_K)$ is the vector of chemo combination effects compared with no chemo. In contrast, there is no baseline comparator group for country category, and the elements of the parameter vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_R)$ are country category effects. These play the role of intercept parameters in η_i, with α_r the intercept if $X_{i,r}$=1. The parameters γ_{K+1} and α_{R+1} account for the patients whose chemo combination or country category are missing, respectively. In particular, the vector $\boldsymbol{\gamma}$ does not include γ_{K+1}, and $\boldsymbol{\alpha}$ does not include α_{R+1}. Let ξ or $\boldsymbol{\xi} = (\xi_1, \xi_2)$ denote any additional scale, shape, or precision (inverse variance) parameters under the assumed model that do not appear in η_i so that $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\gamma}, \gamma_{K+1}, \boldsymbol{\alpha}, \alpha_{R+1}, \boldsymbol{\xi})$. Denoting the probability density function (pdf) and survivor function of T by f and $\bar{F}$, the likelihood for a sample of size n may be expressed as

$$\mathscr{L}(\boldsymbol{T}^{\mathrm{o}}, \boldsymbol{\delta} \mid \boldsymbol{Z}, \boldsymbol{X}, \boldsymbol{W}, \boldsymbol{\theta}) = \prod_{i=1}^{n} \{f(T_i^{\mathrm{o}}, \delta_i \mid \boldsymbol{Z}_i, \boldsymbol{X}_i, \boldsymbol{W}_i, \boldsymbol{\theta})\}^{\delta_i} \{\bar{F}(T_i^{\mathrm{o}} \mid \boldsymbol{Z}_i, \boldsymbol{X}_i, \boldsymbol{W}_i, \boldsymbol{\theta})\}^{1-\delta_i}. \tag{2}$$

We write f, $\bar{F}$, and $\mathscr{L}$ as regression functions in (2) to reflect the viewpoint that $(\boldsymbol{T}^{\mathrm{o}}, \boldsymbol{\delta})$ characterize patient outcomes while $\boldsymbol{Z}$, $\boldsymbol{X}$, and $\boldsymbol{W}$ are vectors of predictive variables.

The distributional families that we consider for T are listed in Table III. Temporarily suppress i and $\boldsymbol{Z}, \boldsymbol{X}, \boldsymbol{W}, \boldsymbol{\theta}$ for brevity, and denote precision, shape, and scale parameters by ξ or (ξ_1, ξ_2). Denoting the normal distribution with mean μ and precision parameter (inverse variance) τ by $\mathrm{N}(\mu, \tau)$, for the lognormal distribution in Table III we assume that $\log(T) \sim \mathrm{N}(\eta, \xi)$. The remaining distributions

Table III. (a) Goodness-of-fit statistics for survival time regression model distributions in the null case with no covariates. A large (small) p-value bound for the Johnson test indicates a good (poor) model fit to the data. p = number of model parameters, BIC = Bayesian information criterion, DIC = deviance information criterion. (b) DIC values for different linear terms of the log normal survival time model. Full = the full hierarchical model, NH Chemo = the model obtained by dropping the hierarchical structure for chemo combination effects.

Distribution	Johnson test p-value bound	p	BIC	DIC
(*a*)				
Log normal	1.000	2	665.4	544.5
Logistic	0.460	2	815.0	809.5
Weibull	0.076	2	679.3	558.3
Gamma	0.019	2	685.8	564.6
Exponential	<0.0001	1	722.6	604.4

Model	DIC
(*b*)	
Full	2831.9
NH Chemo, τ_γ=1.0	2843.2
NH Chemo, τ_γ=0.5	2841.0
NH Chemo, τ_γ=0.1	2859.8
NH Chemo, τ_γ=0.01	2898.4

are given by $\bar{F}(t)=[1+\exp\{\xi(t-\eta)\}]^{-1}$ for the logistic; $f(t)=\exp(\eta-te^{\eta})t^{\xi-1}/\Gamma(\xi)$ for the gamma; $\bar{F}(t)=\exp(-e^{\eta}t^{\xi})$ for the Weibull; and $\bar{F}(t)=\exp(-e^{\eta}t)$ for the exponential.

Given the distributional form, each Bayesian regression model includes two hierarchical structures, one for chemo combinations and one for country categories. We assume that $\gamma_1,\ldots,\gamma_K$ follow a conditionally independent hierarchical model [8, 9]. Since γ_{K+1} represents the chemo effect in patients for whom no elements of their chemo combination were known, it is not included on the hierarchical structure and has its own prior. Similar considerations apply to $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_R)$ and α_{R+1}. Let $\text{Ga}(a,b)$ denote the gamma distribution with mean a/b and variance a/b^2. For each distribution of T considered, we assume the following Bayesian structure, which includes usual (Level 1) priors and hierarchical (Level 2) priors.

Level 1: Priors

$$\begin{aligned}
\beta_1,\ldots,\beta_q &\sim \text{i.i.d. } \text{N}(0,\tau_\beta),\\
\xi \text{ or } \xi_1,\ \xi_2 &\sim \text{i.i.d. } \text{Ga}(a_\xi,b_\xi),\\
\alpha_{R+1} &\sim \text{N}(0,\tau_{\alpha,\text{mis}}),\\
\gamma_{K+1} &\sim \text{N}(0,\tau_{\gamma,\text{mis}}),\\
\alpha_1,\ldots,\alpha_R\,|\,\mu_\alpha,\tau_\alpha &\sim \text{i.i.d. } \text{N}(\mu_\alpha,\tau_\alpha),\\
\gamma_1,\ldots,\gamma_K\,|\,\mu_\gamma,\tau_\gamma &\sim \text{i.i.d. } \text{N}(\mu_\gamma,\tau_\gamma),
\end{aligned}$$

Level 2: Hyperpriors

$$\begin{aligned}
\mu_\alpha &\sim \text{N}(0,\tilde{\tau}_\alpha),\\
\tau_\alpha &\sim \text{Ga}(\tilde{a}_\alpha,\tilde{b}_\alpha),\\
\mu_\gamma &\sim \text{N}(0,\tilde{\tau}_\gamma),\\
\tau_\gamma &\sim \text{Ga}(\tilde{a}_\gamma,\tilde{b}_\gamma).
\end{aligned}$$

To express the model more compactly, we partition the parameter vector as $\boldsymbol{\theta}=(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2)$, where $\boldsymbol{\theta}_1=(\boldsymbol{\beta},\alpha_{R+1},\gamma_{K+1},\xi)$ is the vector of parameters having only Level 1 priors and no hierarchical structure, and $\boldsymbol{\theta}_2=(\boldsymbol{\gamma},\boldsymbol{\alpha})$ is the vector of country category and chemo combination effects, which have hierarchical structures. Denote the vector of fixed parameters characterizing the Level 1 priors of $\boldsymbol{\theta}_1$ by $\boldsymbol{\phi}_1=(\tau_\beta,\tau_{\alpha,\text{mis}},\tau_{\gamma,\text{mis}},a_\xi,b_\xi)$, the vector of random parameters characterizing the Level 1 priors on $\boldsymbol{\theta}_2$ by $\boldsymbol{\phi}_2=(\mu_\alpha,\tau_\alpha,\mu_\gamma,\tau_\gamma)$, and the vector of fixed parameters characterizing the Level 2 priors on $\boldsymbol{\phi}_2$

by $\tilde{\boldsymbol{\phi}}_2=(\tilde{\tau}_\alpha, \tilde{a}_\alpha, \tilde{b}_\alpha, \tilde{\tau}_\gamma, \tilde{a}_\gamma, \tilde{b}_\gamma)$. The Level 1 priors now may be denoted by

$$p(\boldsymbol{\theta}_1 \mid \boldsymbol{\phi}_1) \quad \text{and} \quad p(\boldsymbol{\theta}_2 \mid \boldsymbol{\phi}_2)=p(\boldsymbol{\alpha} \mid \mu_\alpha, \tau_\alpha)\ p(\boldsymbol{\gamma} \mid \mu_\gamma, \tau_\gamma),$$

and the Level 2 priors (hyperpriors) may be denoted by

$$p(\boldsymbol{\phi}_2 \mid \tilde{\boldsymbol{\phi}}_2)=p(\mu_\alpha, \tau_\alpha \mid \tilde{\tau}_\alpha, \tilde{a}_\alpha, \tilde{b}_\alpha)\ p(\mu_\gamma, \tau_\gamma \mid \tilde{\tau}_\gamma, \tilde{a}_\gamma, \tilde{b}_\gamma).$$

With this notation, the posterior may be expressed as

$$p(\boldsymbol{\theta}|\boldsymbol{T}^{\mathrm{o}}, \boldsymbol{\delta}, \boldsymbol{Z}, \boldsymbol{X}, \boldsymbol{W}, \boldsymbol{\phi}_1, \tilde{\boldsymbol{\phi}}_2) \propto \mathscr{L}(\boldsymbol{T}^{\mathrm{o}}, \boldsymbol{\delta} \mid \boldsymbol{Z}, \boldsymbol{X}, \boldsymbol{W}, \boldsymbol{\theta})\ p(\boldsymbol{\theta}_1 \mid \boldsymbol{\phi}_1)\ p(\boldsymbol{\theta}_2 \mid \boldsymbol{\phi}_2)\ p(\boldsymbol{\phi}_2 \mid \tilde{\boldsymbol{\phi}}_2). \tag{3}$$

To complete the model, numerical values of the fixed Level 1 prior parameters $\boldsymbol{\phi}_1$ and the fixed Level 2 prior parameters $\tilde{\boldsymbol{\phi}}_2$ must be specified. This will be done in the context of the data analyses in Section 4 and sensitivity analyses in Section 5.

An important question is whether the hierarchical structure assumed on $\gamma_1, \ldots, \gamma_K$ contributes substantively to the posterior estimates and ranks of single agents and agent pairs. To assess this, we also will consider the model obtained by dropping the Level 2 priors on μ_γ and τ_γ, and instead assuming that μ_γ=0 and τ_γ takes on a fixed positive value. This model may be considered a limiting case of the hierarchical model in which the precision parameter $\tilde{\tau}_\gamma \to \infty$, and $\tilde{a}_\gamma, \tilde{b}_\gamma \to \infty$ subject to $\tilde{a}_\gamma/\tilde{b}_\gamma=\tau_\gamma$. In this model, we retain the Level 2 priors on the country category effects.

3. Defining effects of individual agents and agent pairs

In this section, we present methods for addressing the primary goal of the analyses, which is to define and rank the individual chemo agents in terms of their effects on OS. We do this in terms of the chemo combination effects appearing in the regression model for OS, which also accounts for the effects of country categories, other baseline covariates, and the two non-chemo treatments, radiation and surgery. To avoid technical difficulties, in the definition we use only the effects $\gamma_1, \ldots, \gamma_{23}$ of the 23 chemo combinations for which all agents are known (Table I). To keep track of both individual chemo agents and chemo combinations, let $Y_{c,a}$ denote the indicator that combination c includes agent a, for $c=1, \ldots, 23$ and $a=1, \ldots, J$, where here $J=7$. Denote the number of combinations among the first 23 that contain agent a by $N_a=\sum_{c=1}^{23} Y_{c,a}$.

Our definition of an individual agent effect is

$$\zeta_a=\frac{\sum_{c=1}^{23} Y_{c,a}\,\gamma_c}{\sum_{c=1}^{23} Y_{c,a}}=\sum_{c=1}^{23} w_{c,a}\,\gamma_c \quad \text{for } a=1, \ldots, J, \tag{4}$$

where $w_{c,a}=Y_{c,a}/N_a$. This says that the effect of agent a is the equally weighted average of the effects of all combinations, for which all agents are known, that contain agent a. We make no attempt to define the quantitative contribution of agent a to any γ_c for which $Y_{c,a}$=1, since this cannot be determined from the available data. In any case, it is difficult to determine one agent's contribution to a chemo combination due to unmodeled interactions between agents. We do not include between-agent interaction terms in the model because the data cannot provide a reliable basis for their estimation, and because it is a common oncology practice to adjust the doses of each agent in a combination to control the patient's overall amount of chemotherapy.

It is very important to bear in mind that ζ_a is not the effect of agent a when administered alone. Rather, it is the average of the effects of the combinations, for which all agents were identified, that included a. This set of combinations may or may not include a given as a single agent. Moreover, the definition of ζ_a is data-dependent, since it relies on the particular chemo combinations that include a in the data set at hand. However, given the facts that CPC is a rare disease, no large-scale comparative trials of CPC therapies have been conducted, and the data were obtained by a painstaking literature review, it is highly unlikely that another CPC data set having similar structure will become available in the near future.

A natural question is how the individual agents may behave when given together. While it is not practical to address this in general based on the available data, some progress can be made by considering the pairs of agents in terms of $\boldsymbol{\gamma}$. To do this, we first define $Y_{c,\{a,b\}}$ to be the indicator that both agent

a and agent b are included in combination c, and we define agent pair effects analogously to the way that we defined single agent effects, as follows. Denote the number of combinations, again for which all agents are known, that include both a and b by $N_{\{a,b\}}=\sum_{c=1}^{23} Y_{c,\{a,b\}}$. While there are $7\times\frac{6}{2}=21$ possible pairs, in the CPC data set no combination contained both IFOS and PRC; hence, only 20 pairs were included in at least one combination given to at least one patient. Our definition of the effect of agents a and b when given together in some combination, either with or without other agents, is

$$\zeta_{\{a,b\}}=\frac{\sum_{c=1}^{23} Y_{c,\{a,b\}}\,\gamma_c}{\sum_{c=1}^{23} Y_{c,\{a,b\}}}=\sum_{c=1}^{23} w_{c,\{a,b\}}\,\gamma_c, \tag{5}$$

where $w_{c,\{a,b\}}=Y_{c,\{a,b\}}/N_{\{a,b\}}$. Thus, $\zeta_{\{a,b\}}$ is the equally weighted average of the effects on OS of all combinations containing both a and b. As with the single-effect definition, $\zeta_{\{a,b\}}$ does not pertain to how the chemo combination pair $\{a,b\}$ alone would affect OS.

4. Ranks of individual agent effects and agent pair effects

Our primary goal is to rank the single agents with regard to their effects on OS. To augment these analyses, we also will rank the effects of the 20 agent pairs that were included together in at least one combination. There is an extensive literature on Bayesian and empirical Bayes ranking [10–13]. A review is given by Carlin and Louis [14, Chapter 7]. As pointed out by Laird and Louis [10], in a Bayesian or empirical Bayes setting, ranking a set of parameters corresponding to a set of observational units using the parameters' posterior means suffers from the problem that unequal variances may produce misleading conclusions. This issue is important in the present setting due to the fact that the numbers of chemo combinations including each agent, $N_1,\ldots,N_7$, vary from 5 to 17 (Table I). This implies that, under our data-based definition (4), the individual agent effects $\zeta_1,\ldots,\zeta_7$ may have different variances due to sample size differences, aside from intrinsic differences among the posterior variances of the chemo combination effects themselves. Furthermore, it has been shown that ranking based on either the maximum likelihood estimates (MLEs) of the parameters or Z-scores of frequentist test statistics also perform poorly [11]. Consequently, rather than ranking the individual agents based on the posterior means of the ζ_a's, we will follow the recommendation of Laird and Louis [10] by estimating the posterior ranks of the ζ_a's, defined by

$$R_a=\sum_{k=1}^{J} I(\zeta_a\geqslant\zeta_k),\quad a=1,\ldots,J. \tag{6}$$

We denote the posterior mean of R_a by $\bar{R}_a$. Since J=7 in the CPC date set, each $\bar{R}_a$ takes on values in the domain $[1,7]$, and in general this statistic is not integer-valued [11]. Among possible estimators for R_a, the posterior mean $\bar{R}_a$ is optimal under squared error loss [13]. We denote the integer rank of $\bar{R}_a$ among $\{\bar{R}_1,\ldots,\bar{R}_7\}$ by $\hat{R}_a$.

Extending the above structure to agent pairs, we denote the rank of $\zeta_{\{a,b\}}$ among $\{\zeta_{1,2},\ldots,\zeta_{6,7}\}$ by

$$R_{\{a,b\}}=\sum_{1\leqslant r<s\leqslant 7} I(\zeta_{\{a,b\}}\geqslant\zeta_{\{r,s\}}). \tag{7}$$

We denote the posterior mean of $R_{\{a,b\}}$ by $\bar{R}_{\{a,b\}}$. Since there are 20 agent pairs included in at least one chemo combination in the brain tumor data set, each $\bar{R}_{\{a,b\}}$ takes on a value in the domain $[1,20]$. We denote the integer rank of $\bar{R}_{\{a,b\}}$ among the $\bar{R}_{\{a,b\}}$'s by $\hat{R}_{\{a,b\}}$.

Since each ζ_a and $\zeta_{\{a,b\}}$ is a linear combination of the chemo combination effects, $\boldsymbol{\gamma}$, computing the posteriors of these parameters and of the ranks R_a and $R_{\{a,b\}}$ requires the posterior of $\boldsymbol{\gamma}$, which may be obtained as the marginal

$$p(\boldsymbol{\gamma}|\boldsymbol{T}^{\mathrm{o}},\boldsymbol{\delta},\boldsymbol{Z},\boldsymbol{X},\boldsymbol{W},\boldsymbol{\phi}_1,\tilde{\boldsymbol{\phi}}_2)=\int_{(\boldsymbol{\theta}_1,\boldsymbol{\alpha})} p(\boldsymbol{\theta}|\boldsymbol{T}^{\mathrm{o}},\boldsymbol{\delta},\boldsymbol{Z},\boldsymbol{X},\boldsymbol{W},\boldsymbol{\phi}_1,\tilde{\boldsymbol{\phi}}_2)\,\mathrm{d}\boldsymbol{\theta}_1\,\mathrm{d}\boldsymbol{\alpha}.$$

5. Analysis of the CPC data

5.1. Covariates and priors

The non-chemo treatment covariates used in the survival time regression models are type of surgery (CR=complete resection, PR=partial resection, BN=biopsy or no surgery), and radiation (Yes or No), which we record using the indicators $Z_1 = I(\text{CR})$, $Z_2 = I(\text{PR})$, and $Z_3 = I(\text{radiation})$. The two baseline covariates are Z_4=age in years and the indicator $Z_5 = I(\text{gender=male})$. Thus, in the linear term $\boldsymbol{\beta Z} = \beta_1 Z_1 + \cdots \beta_5 Z_5$, the parameters β_1 and β_2 are the respective effects of surgery resulting in CR or PR compared with the baseline group BN, β_3 is the radiation effect, β_4 is the per-year age effect, and β_5 is the effect of being male. Because there are 13 known country categories and an 'unknown country category,' $R = 13$ and $\dim(\boldsymbol{X}_i) = 14$. Similarly, since there are $K = 27$ chemo combinations with at least one agent known while 32 patients received chemo with no agents known, $\dim(\boldsymbol{W}_i) = 28$.

Two additional baseline covariates in the data set are tumor location (INF=infratentorial versus SUP=supratentorial) and publication year. Because tumor location is highly associated with type of surgery, and publication year is highly associated with both radiation and whether a patient received chemo, to avoid collinearity we do not include either tumor location or publication year in any regression model for OS. These two covariates are utilized in the regression models used to impute missing values, described in Section 5.2.

To complete the Bayesian model, numerical values must be specified for the four fixed parameters in $\boldsymbol{\phi}_1$ and the six fixed parameters in $\tilde{\boldsymbol{\phi}}_2$. We assume vague priors, with μ_α, μ_γ, α_{R+1}, γ_{K+1}, and the elements of $\boldsymbol{\beta}$ i.i.d. N(0, 0.01), and τ_α, τ_γ and each ξ_j i.i.d. Ga(0.10, 0.10), which has mean 1 and variance 10. It is important to note, however, that a Ga(ε, ε) prior with small ε is not necessarily 'non-informative' since, if small values of the parameter are possible, then posterior inferences may be sensitive to ε. A discussion and illustration of this point is given by Gelman [15]. In summary, the fixed parameters determining the Level 1 priors are $(\tau_\beta, \tau_{\text{mis}}, a_\xi, b_\xi), = (0.01, 0.01, 0.10, 0.10)$ and the fixed parameters determining the Level 2 priors are $(\tilde{\tau}_\alpha, \tilde{a}_\alpha, \tilde{b}_\alpha, \tilde{\tau}_\gamma, \tilde{a}_\gamma, \tilde{b}_\gamma) = (0.01, 0.10, 0.10, 0.01, 0.10, 0.10)$. The small normal precision parameters $\tilde{\tau}_\alpha = \tilde{\tau}_\gamma = 0.01$ and the large gamma variances $\tilde{a}_\alpha/\tilde{b}_\alpha^2 = \tilde{a}_\gamma/\tilde{b}_\gamma^2 = 10$ quantify the lack of prior knowledge about the parent populations for the country category effects and chemo combination effects.

5.2. Imputing missing values

Some patients in the data set had partially missing covariates, including age (2 missing), gender (10 missing), surgery (9 missing), tumor location (13 missing), and radiation (1 missing). In addition, Table I shows that there were four chemo combinations ($c = 24, \ldots, 27$) in which the elements of the agent inclusion indicator vector $\boldsymbol{Y}_c = (Y_{1,c}, \ldots, Y_{J,c})$ were only partially known. These combinations were administered to a total of seven patients. In order to include the data from patients who had some missing entries of $\boldsymbol{Z}$ or received combinations for which some $Y_{c,a}$'s are missing, we employ Bayesian model-based data augmentation [16–19]. The data augmentation is carried out, as in the conventional Bayesian parametric model-based regression analysis, by treating the missing values like additional parameters when applying the Markov chain Monte Carlo (MCMC) algorithm to compute the posterior [19].

In the fit of each regression model considered, imputation entails embedding a fit of a regression model for each missing value as a function of other variables and $\log(T)$ in each iteration of the MCMC algorithm, using the fitted model to simulate the missing value, and treating the simulated value as the missing value's latest update. To establish notation for missing covariates and models used in the imputation, if the covariate Z_j was missing for patient i, let $\boldsymbol{Z}_{i,-Z_j}$ denote a vector including $\log(T_i)$ and a selected subvector of that patient's covariates other than Z_j. If some covariates other than Z_j also are missing for that patient, then those other covariates are represented in $\boldsymbol{Z}_{i,-Z_j}$ initially as sample means and subsequently as the imputed value from the previous MCMC step. The imputation model used for age, which is the only non-categorical-valued covariate, is the linear regression model $\log(Z_{i,\text{age}}) = \boldsymbol{\omega Z}_{i,-\text{age}} + \varepsilon_i$, with $\varepsilon_i \sim \text{N}(0, \tau_\varepsilon)$ and $\boldsymbol{\omega}$ a parameter vector including an intercept. The fitted $\text{N}(\hat{\boldsymbol{\omega}}\boldsymbol{Z}_{i,-\text{age}}, \hat{\tau}_\varepsilon)$ distribution is used to simulate $Z_{i,\text{age}}$, with this value included in the next MCMC iteration. For a missing binary covariate, $Z_{i,j}$, a logistic regression model with linear term $\boldsymbol{\omega Z}_{i,-Z_j}$ is used with $Z_{i,j}$ simulated from the estimated probability $\exp(\hat{\boldsymbol{\omega}}\boldsymbol{Z}_{i,-Z_j})/\{1+\exp(\hat{\boldsymbol{\omega}}\boldsymbol{Z}_{i,-Z_j})\}$. Only $\log(T_i)$ is included in $\boldsymbol{Z}_{i,-\text{age}}$, $\boldsymbol{Z}_{i,-\text{gender}}$, and $\boldsymbol{Z}_{i,-(\text{tumor location})}$ because survival time is by far the strongest

predictor of these three covariates. The imputation model for missing radiation values also includes publication period as the three-category variable ($\leqslant$1979, 1980–1999, $\geqslant$2000) in $\boldsymbol{Z}_{i,-\text{radiation}}$ because this was highly predictive of whether a patient received radiation. A three-category generalized logistic model is used for missing surgery, and $\boldsymbol{Z}_{i,-(CR,PR)}$ consists of tumor location and $\log(T_i)$.

To impute missing chemo combination membership indicators, we proceeded similarly. If $Y_{i,a,c}$ was missing for patient i, let $\boldsymbol{Y}_{i,-a,c}$ denote the vector obtained from the $J-1$ dimensional subvector of $\boldsymbol{Y}_{i,c}$ without the entry $Y_{i,a,c}$, also including $\log(T_i)$, country category, and publication period ($\leqslant$ 1994 versus $\geqslant$ 1995). The missing $Y_{i,a,c}$ value is simulated using a fitted logistic model for $\Pr(Y_{i,a,c}=1)$ having linear term $\boldsymbol{\omega}\boldsymbol{Y}_{i,-a,c}$.

5.3. *Goodness of fit for survival time distributions*

To allow the data to determine a form for the OS time distribution, initially we consider several possible distributions for $[T|\boldsymbol{Z},\boldsymbol{X},\boldsymbol{W},\boldsymbol{\theta}]$. Under each distribution, we assess the regression model's fit to the data using Johnson's Bayesian χ^2 test [20], a modified version of Schwarz's Bayesian information criterion (BIC, [21]) and the deviance information criterion (DIC, Spiegelhalter *et al.* [22]). We compute the BIC as

$$\text{BIC}=-2\log(\mathscr{L}(\hat{\boldsymbol{\theta}}^{\text{mle}}))+p\log(d),$$

where p is the number of model parameters, d the number of deaths, and $\hat{\boldsymbol{\theta}}^{\text{mle}}$ denotes the frequentist MLE. Following the recommendations of Volinsky and Raftery [23] and Ibrahim *et al.* [7], we use d rather than the sample size in the penalty term $p\log(d)$ since this provides a better approximation to the Bayes Factor when using the BIC to compare models for right-censored event time data. To compute the BIC for the full hierarchical model, the frequentist version of the model has parameter vector $\boldsymbol{\theta}=(\boldsymbol{\beta},\mu_\gamma,\tau_\gamma,\gamma_{K+1},\mu_\alpha,\tau_\alpha,\alpha_{R+1},\xi)$, where ξ is the lognormal precision parameter; hence, $p=\dim(\beta)+7$. If the hierarchical structure for the γ_c's is dropped, so that there are no Level 2 priors on (μ_γ,τ_γ), then these become fixed prior parameters and in the frequentist version of the model (μ_γ,τ_γ) are replaced by $\gamma_1,\ldots,\gamma_{23}$; hence, p becomes much larger. For the data at hand, however, this is a moot point for this model due to the fact that there were 10 chemo combinations for which there were 0 deaths; hence, the MLE does not exist and the BIC cannot be computed. Since Johnson's Bayesian χ^2 test can be computed only under the null model with no covariates, in Table III(a) we also computed the BIC and DIC under this null model. The BIC computations were carried out by obtaining $\mathscr{L}(\hat{\boldsymbol{\theta}}^{\text{mle}})$ using SAS PROC Lifereg, which parameterizes survival time regression models in the accelerated failure form $\log(T)=\eta+\xi\log(T_0)$, where T_0 follows a standard distributional form. DIC values were computed using WinBUGS version 1.4.3. Because the BIC, DIC, and p-value bounds of Johnson's test all indicate that the best fit is provided by the log normal distribution, we chose to use this distribution for all of our analyses.

To assess the effect of the assumed hierarchical structure for the γ_c's on the log normal model fit when including treatment and covariate variables in η, we computed the DIC both with (Full Model) and without (NH Chemo model) this hierarchical structure (Table III(b)). The DIC values indicate that, when including treatments and covariates, the full model provides a better fit to the data than any version of the NH Chemo model with precision parameter varying from 1.0 to 0.01. The DIC values also indicate that, under the NH Chemo model, assuming a highly imprecise prior with $\tau_\gamma\leqslant 0.1$, equivalently a prior variance of each γ_c greater than 10, gives a substantially worse fit to the data.

5.4. *Fitted models and effect estimates*

The fitted log normal model for OS is summarized in Table IV. The posterior probability $\Pr(\mu_\gamma>0|\text{data})>0.99$ indicates that receiving some form of chemo was greatly beneficial compared with receiving no chemo. For the covariates and treatment variables, the posterior probabilities $\Pr(\beta_j>0|\text{data})$ indicate that achieving either a complete or partial resection at surgery was greatly beneficial compared with either a biopsy or no surgery, and receiving radiation also was greatly beneficial. It appears that older age was moderately advantageous and that males had moderately worse survival than females, but neither Age nor Gender was strongly associated with OS. We thus fit a reduced version of the full model excluding Age and Gender, which gave BIC=598.2 and DIC=1400.6. While this BIC is slightly larger than the value 595.7 of the full model, indicating that dropping Age and Gender gives a slightly worse fit, the DIC is dramatically smaller than the value 2831.9 of the full model including

Table IV. Fitted Bayesian log normal regression model for overall survival time, including hierarchical structures on the country categories and on the chemo combinations.

		Posterior quantities		
Parameter	Effect description	Mean	std	$\Pr(\beta > 0 \mid \text{data})^*$
Level 1 parameters				
β_1	Surgery=CR (vs biopsy/no surgery)	2.43	0.45	1.00
β_2	Surgery=PR (vs biopsy/no surgery)	0.70	0.43	0.95
β_3	Radiation=yes	1.28	0.29	1.00
β_4	Age	0.013	0.012	0.87
β_5	Gender=male	−0.28	0.27	0.15
ξ	Lognormal precision	0.37	0.06	—
α_{14}	Missing country category	−2.02	0.68	—
γ_{28}	Missing chemo combination	1.42	0.44	—
Level 2 parameters				
μ_α	Country category mean†	−1.40	0.49	—
τ_α	Country category precision	5.87	5.37	—
μ_γ	Chemo combination mean‡	1.21	0.35	>0.99
τ_γ	Chemo combination precision	5.35	5.55	—

*A value close to 1 corresponds to a significantly large beneficial effect.
†Mean overall country category effect.
‡Mean overall chemo combination effect, compared with no chemo.

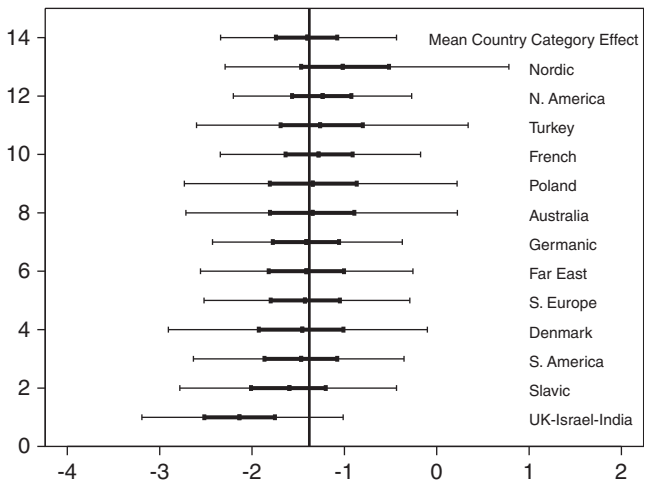

Figure 1. Summaries of the posterior distributions of the overall mean country category effect μ_α and of the individual country category effects $\alpha_1, \ldots, \alpha_{13}$ on overall survival time. For each effect distribution, the thin line runs from the 2.5 to 97.5th percentiles, the thick line runs from the 25 to 75th percentiles, and the midpoint is the median. The vertical line is located at the posterior mean of μ_α.

these two covariates. This is due to the fact that some values of both Age and Gender were missing and thus were imputed, which required additional model fits embedded in the MCMC. It thus appears that the large decrease in the DIC value was an artifact of the fact that there was no imputation when Age and Gender were dropped from the model.

The posterior country category effects are summarized in Figure 1. The shorter median survival time of 0.4 years for the UK–Israel–India country category, and the longer survival but small sample of 2 patients in the Nordic country category (Table II), both are illustrated graphically in Figure 1. All differences between posterior country category effects appear to be small, however. The posteriors of the chemo combination effects and overall mean chemo combination effect are given in Figure 2, which shows a clear survival advantage for patients receiving some form of chemo. The combinations VP16+CARBO+CYC, VP16+VCR+CDDP+CYC, and VP16+CARBO+IFOS are each associated with slightly longer survival time, but there is substantial overlap among all the posteriors of the γ_c's. We next perform a more formal ranking of individual agents and agent pairs.

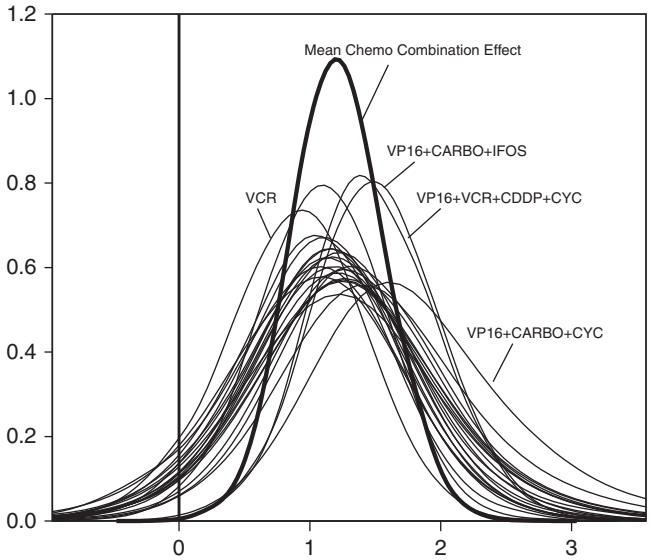

Figure 2. Posterior distributions of the overall mean chemo combination effect μ_γ and of all chemo combination effects $\gamma_1, \ldots, \gamma_{23}$ on the overall survival time.

6. Ranking chemotherapy agents and agent pairs

6.1. Posteriors of effects and ranks

For the full model fit both with and without Age and Gender included in η, Table V summarizes the posteriors of each ζ_a and its rank R_a, and gives the estimates $\bar{R}_a$ and the corresponding integer rank $\hat{R}_a$, for each $a = 1, \ldots, 7$. There is very little difference among the posteriors of the ζ_a's. In terms of the posterior mean or integer ranks, bearing in mind that in general some form of chemo is beneficial, a message of Table V is that chemo combinations containing CARBO and CYC have the largest and second largest beneficial effects on survival, while combinations containing VCR and CDDP have the two smallest beneficial effects. The posterior ranks, in terms of both $\bar{R}_a$ and $\hat{R}_a$, were insensitive to whether Age and Gender were included in the model, with only the ranks 1 and 2 of VCR and CDDP reversed by dropping these covariates. The statistics $\tilde{R}_a^{\max}$ and $\tilde{R}_a^{\min}$ given in Table V will be defined and discussed below.

The histograms of the posteriors of the ranks $R_1, \ldots, R_7$, given in Figure 3, are much more revealing than the summary statistics in Table V. While the right-skewed posterior of the rank for CARBO and the left-skewed posterior of the rank for CDDP both agree with what was seen in terms of the $\bar{R}_a$'s and $\hat{R}_a$'s, two striking results are that the posteriors of the ranks for both IFOS and PRC are clearly bimodal, with highest peaks at the two extreme ranks 1 and 7. The fact that the effect ζ_a for IFOS was often ranked either the highest (by $\tilde{R}_a^{\max}$) or lowest (by $\tilde{R}_a^{\min}$) says that chemo combinations containing IFOS were associated with either longer or shorter survival times, compared with combinations not containing IFOS. The same considerations apply to PRC. Thus, if the shapes of the posteriors were ignored, the fact that both of the summary statistics $\bar{R}_a$ and $\hat{R}_a$ for IFOS and PRC were in the middle of their rank orderings would be very misleading. These results suggest that there may have been interactive effects between each of IFOS and PRC and other agents. Another possible explanation is that these agents were used frequently in combinations with other agents that, comparatively, were either better or worse, as discussed below.

Table VI summarizes the posterior of each $\zeta_{\{a,b\}}$ and its rank $R_{\{a,b\}}$, and gives the estimates $\bar{R}_{\{a,b\}}$ and $\hat{R}_{\{a,b\}}$ for each distinct pair $\{a, b\}$. As for the single agent effects, these were computed under the full model both with and without Age and Gender. Based on the summary statistics of the posterior ranks, chemo combinations containing the pair {CARBO, CYC} had the largest beneficial effects on survival, combinations containing the pair {CARBO, IFOS} had the second largest beneficial effects, whereas combinations containing the pair {VP16, PRC} had the smallest beneficial effects. These results are consistent with those seen for the individual agent ranks. Figure 4 shows that, as with the

Table V. For each chemo agent, the posterior mean of the effect ζ_a, the posterior mean $\bar{R}_a$ of the rank R_a of ζ_a, and the integer-valued rank $\hat{R}_a$ of $\bar{R}_a$.

	Posterior quantities							
	Full model					Full model-age–gender		
Chemo agent	Mean (ζ_a)	$\bar{R}_a$	$\hat{R}_a$	$\tilde{R}_a^{\max}$	$\tilde{R}_a^{\min}$	Mean (ζ_a)	$\bar{R}_a$	$\hat{R}_a$
CARBO	$1.32_{0.36}$	$4.73_{1.86}$	7	5	2	$1.25_{0.35}$	$4.75_{1.86}$	7
CYC	$1.30_{0.36}$	$4.57_{1.79}$	6	4	1	$1.23_{0.35}$	$4.62_{1.78}$	6
VP16	$1.27_{0.34}$	$4.06_{1.74}$	5	3	3	$1.19_{0.33}$	$4.06_{1.75}$	5
IFOS	$1.27_{0.39}$	$4.06_{2.31}$	4	6	6	$1.19_{0.38}$	$4.03_{2.31}$	4
PRC	$1.26_{0.42}$	$3.88_{2.43}$	3	7	7	$1.18_{0.41}$	$3.87_{2.41}$	3
VCR	$1.23_{0.35}$	$3.36_{1.45}$	2	1	4	$1.15_{0.34}$	$3.31_{1.45}$	1
CDDP	$1.22_{0.35}$	$3.32_{1.81}$	1	2	5	$1.14_{0.34}$	$3.36_{1.78}$	2

$\tilde{R}_a^{\max}$ and $\tilde{R}_a^{\min}$ are the estimated ranks of the agents with the largest posterior probabilities of being ranked the largest or smallest, respectively. Standard deviations are given as subscripts.

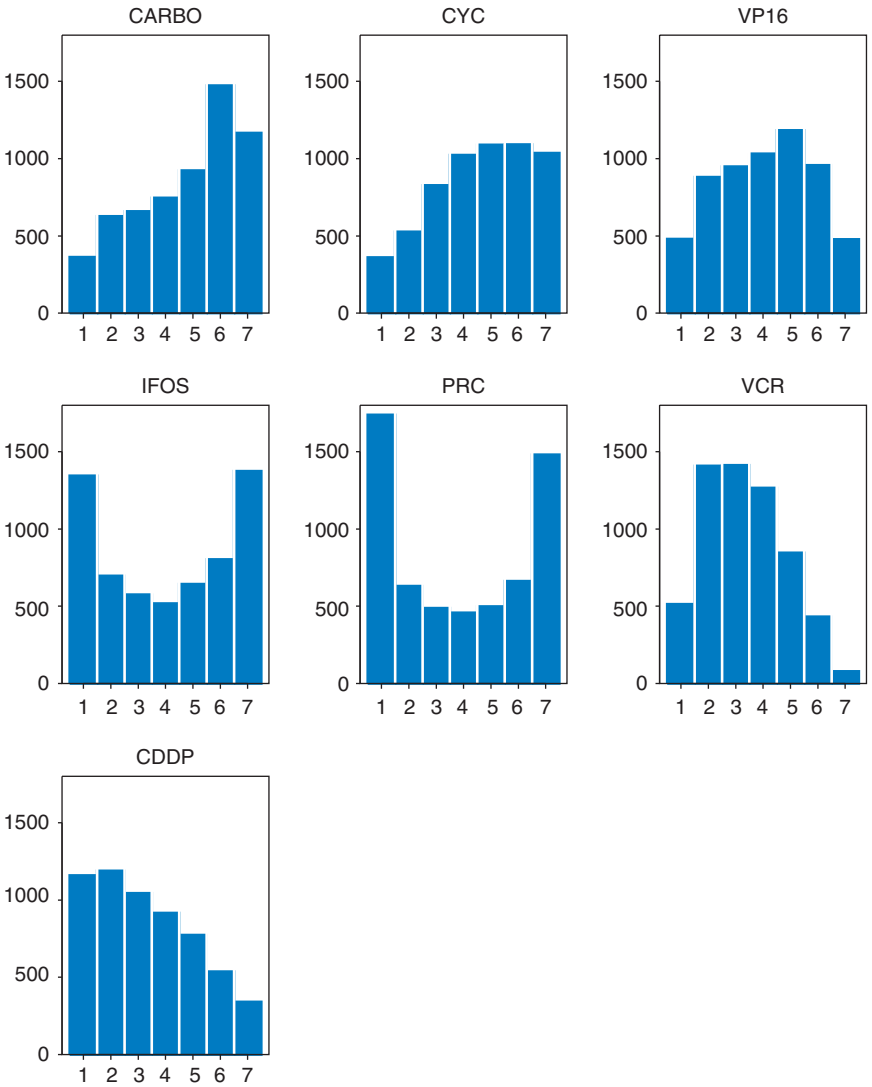

Figure 3. Posterior distributions of the individual chemo agent effect ranks $R_1, \ldots, R_7$, where R_a = rank of ζ_a among $\{\zeta_1, \ldots, \zeta_7\}$.

Table VI. For each pair $\{a, b\}$ of chemo agents appearing together in at least one combination, the posterior mean $E(\zeta_{\{a,b\}})$ of the effect $\zeta_{\{a,b\}}$, the posterior mean $\bar{R}_{\{a,b\}}$ of the rank $R_{a,b}$ of $\zeta_{\{a,b\}}$, and the rank $\hat{R}_{\{a,b\}}$ of $\bar{R}_{\{a,b\}}$.

	Posterior quantities					
	Full model			Full model-age–gender		
Chemo agent pair	$E(\zeta_{\{a,b\}})$	$\bar{R}_{\{a,b\}}$	$\hat{R}_{\{a,b\}}$	$\mathrm{E}(\zeta_{\{a,b\}})$	$\bar{R}_{\{a,b\}}$	$\hat{R}_{\{a,b\}}$
CARBO–CYC	$1.39_{0.43}$	$13.78_{05.22}$	20	$1.32_{0.42}$	$13.74_{5.22}$	20
CARBO–IFOS	$1.36_{0.44}$	$12.65_{06.22}$	19	$1.28_{0.42}$	$12.56_{6.28}$	19
CYC–VP16	$1.32_{0.38}$	$12.51_{04.50}$	18	$1.24_{0.37}$	$12.35_{4.51}$	17
CARBO–VP16	$1.32_{0.37}$	$12.31_{04.75}$	17	$1.25_{0.36}$	$12.36_{4.81}$	18
VP16–IFOS	$1.29_{0.41}$	$11.68_{05.41}$	16	$1.22_{0.40}$	$11.54_{5.39}$	16
CYC–CDDP	$1.28_{0.37}$	$11.42_{03.93}$	15	$1.21_{0.36}$	$11.43_{3.95}$	15
CARBO–VCR	$1.28_{0.41}$	$11.14_{04.94}$	14	$1.21_{0.40}$	$11.09_{5.01}$	14
IFOS–VCR	$1.28_{0.54}$	$10.96_{06.73}$	13	$1.21_{0.54}$	$11.01_{6.71}$	13
PRC–VCR	$1.26_{0.42}$	$10.76_{05.69}$	12	$1.18_{0.41}$	$10.65_{5.69}$	12
VP16–VCR	$1.24_{0.38}$	$10.38_{04.27}$	11	$1.17_{0.37}$	$10.30_{4.25}$	9
CYC–VCR	$1.24_{0.36}$	$10.29_{04.13}$	10	$1.17_{0.35}$	$10.31_{4.10}$	10
VCR–CDDP	$1.24_{0.36}$	$10.24_{04.11}$	9	$1.17_{0.35}$	$10.21_{4.10}$	8
CARBO–CDDP	$1.24_{0.43}$	$10.12_{05.78}$	8	$1.18_{0.42}$	$10.31_{5.85}$	11
CYC–IFOS	$1.17_{0.66}$	$9.53_{07.57}$	7	$1.10_{0.66}$	$9.57_{7.57}$	7
CARBO–PRC	$1.19_{0.48}$	$9.48_{06.51}$	6	$1.11_{0.47}$	$9.34_{6.50}$	4
PRC–CDDP	$1.16_{0.44}$	$9.29_{05.96}$	4.5	$1.09_{0.43}$	$9.46_{5.98}$	5.5
CYC–PRC	$1.16_{0.44}$	$9.29_{05.96}$	4.5	$1.09_{0.43}$	$9.46_{5.98}$	5.5
IFOS–CDDP	$1.13_{0.50}$	$8.66_{06.54}$	3	$1.06_{0.49}$	$8.64_{6.52}$	3
VP16–CDDP	$1.17_{0.36}$	$8.58_{04.39}$	2	$1.10_{0.35}$	$8.58_{4.39}$	2
VP16–PRC	$1.09_{0.49}$	$7.96_{06.61}$	1	$1.03_{0.47}$	$8.12_{6.66}$	1

Standard deviations are given as subscripts.

single agent effect ranks, several pairs had ranks $R_{\{a,b\}}$ with bimodal posteriors, and this was most pronounced for {CYC, IFOS}.

Some of these results may be explained, in part, by examining the posteriors of the γ_c's (Figure 2) and the distributions of patients among the combinations (Table I). The comparatively superior performance of CARBO and CYC in terms of the posterior ranks may be explained by noting that these agents were included in the three combinations having γ_c's with largest posterior means (Figure 2). The bimodal posterior of the rank of ζ_a for PRC may be explained by the fact that, examining Table I, PRC was included in five combinations given to a total of 20 patients, and the combination among these given to 13 patients also included both CYC and CARBO. Moreover, Table VI shows that four of the six pairs having the lowest posterior means $\bar{R}_{\{a,b\}}$ included PRC. Similarly, the bimodal posterior of R_a for IFOS may be explained by the fact that it was included in five combinations given to a total of 24 patients, and the combination among these given to 14 patients also included CARBO.

The bimodal posteriors of the ranks of the ζ_a's for IFOS and PRC also illustrate potential difficulties in identifying a 'best' or 'worst' agent based on combination chemo data of the form considered here. For $0<\lambda<1$, identifying the 'best' agent, formally, identifying the index a such that $\Pr(R_a=K|\text{data})$ is largest, is equivalent to identifying the upper 100λ per cent of the agents for $\lambda>(K-1)/K$, since in this case $\pi_a(\lambda)=\Pr(R_a>[\lambda K]|\text{data})=\Pr(R_a=K|\text{data})$. In general, Lin *et al.* [13] showed that the rank of $\pi_a(\lambda)$ among $\{\pi_1(\lambda),\ldots,\pi_K(\lambda)\}$ is the optimal estimator under $0-1$ loss, defined as the number of misclassifications of the R_a's either above or below λK. For $\lambda>(K-1)/K$, we denote this estimator by $\tilde{R}_a^{\max}$. We define $\tilde{R}_a^{\min}$ similarly for the 'worst' agent, formally the index a with largest $\Pr(R_a=1|\text{data})$. These statistics are given in Table V under the full model. A large value of $\tilde{R}_a^{\max}$ in the domain $[1,7]$ corresponds to $\Pr(R_a=7|\text{data})$ often being largest, i.e. ζ_a often being ranked largest among $\{\zeta_1,\ldots,\zeta_7\}$. Similarly, a large value of $\tilde{R}_a^{\min}$ in the domain $[1,7]$ corresponds to $\Pr(R_a=1|\text{data})$ often being ranked largest, i.e. ζ_a often being ranked smallest among $\{\zeta_1,\ldots,\zeta_7\}$. The results $\tilde{R}_a^{\max}=\tilde{R}_a^{\min}=7$ for PRC and $\tilde{R}_a^{\max}=\tilde{R}_a^{\min}=6$ for IFOS in Table V appear to be incongruous, since they imply that $\Pr(R_a=7|\text{data})$ and $\Pr(R_a=1|\text{data})$ for PRC were both ranked the highest most often, and that this same effect was seen for IFOS. These apparent anomalies are explained quite easily, however, by simply examining the bimodal posteriors of the ranks for PRC and IFOS given in Figure 3. This illustrates the more general points that summary statistics, such as means or ranks, may be very misleading for bimodal

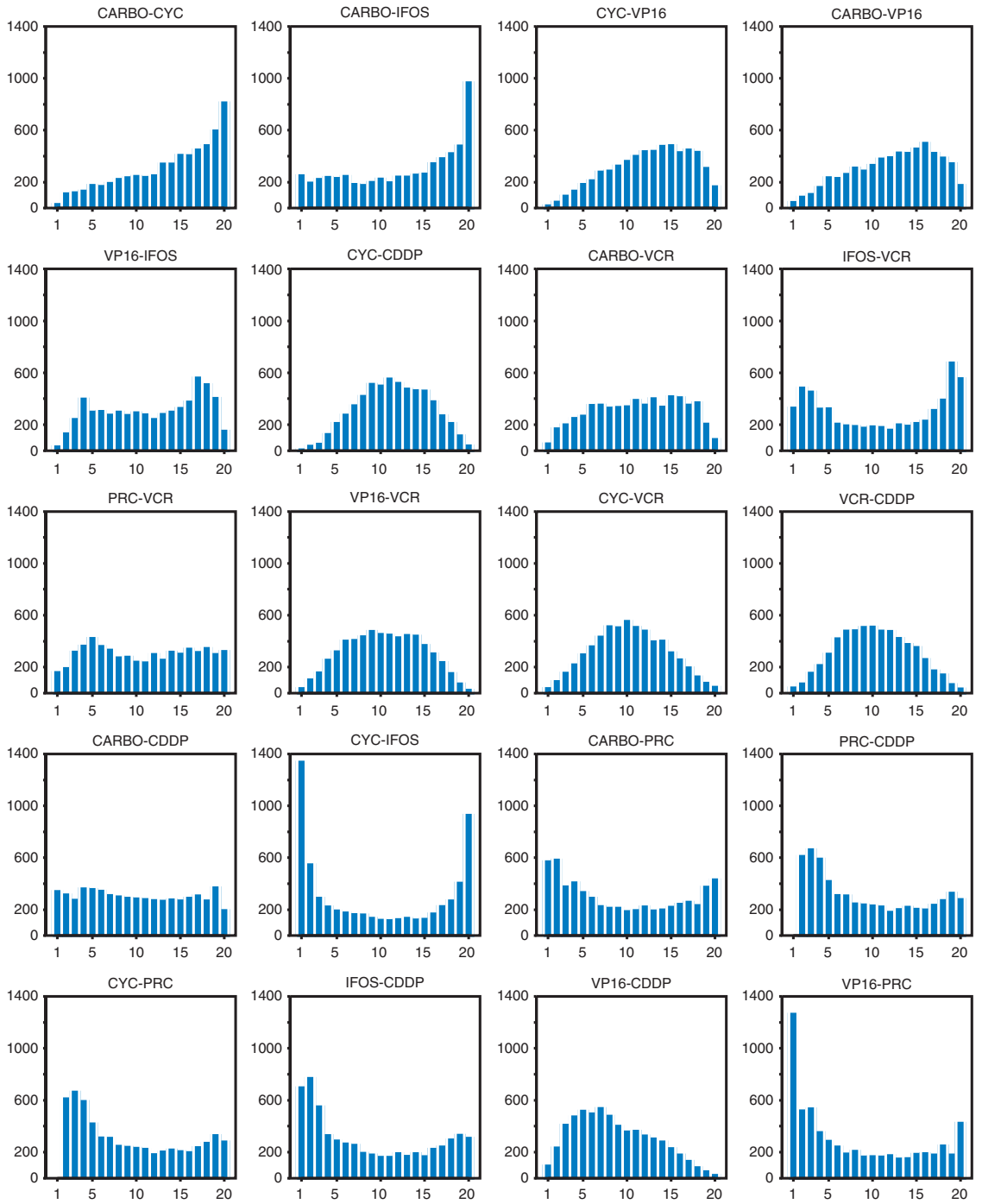

Figure 4. Posterior distributions of the individual chemo agent pair effect ranks $R_{a,b}$ for $1 \leqslant a < b \leqslant 7$, where $R_{\{a,b\}} =$ rank of $\zeta_{\{a,b\}}$ among the 20 effects of agent pairs given together in at least one chemo combination in the CPC data set.

or multimodal distributions, and that it often is important in Bayesian inference to examine the entire posterior distribution.

6.2. *Sensitivity analyses*

An important issue is the sensitivity of the posterior estimates and rankings to the model assumptions. These include the fixed Level 1 and Level 2 prior parameter values under the assumed full model and, more generally, the assumed hierarchical structure on the γ_c's. To address this, we first carry

out a sensitivity analysis of the posterior rank estimates to the numerical values of the fixed Level 1 and Level 2 prior parameters under our full model. These include (i) the precision parameter τ_β characterizing the Level 1 normal priors on the β_j's; (ii) the parameters (a_ξ, b_ξ) characterizing the Level 1 gamma prior on the precision parameter ξ of the lognormal distribution; (iii) for the Level 2 hyperpriors of the country category effects $\alpha_1, \ldots, \alpha_{13}$, the precision parameter $\tilde{\tau}_\alpha$ of the mean μ_α and the gamma parameters $(\tilde{a}_\alpha, \tilde{b}_\alpha)$ of the precision τ_α; and (iv) for the Level 2 hyperpriors of the chemo combination effects $\gamma_1, \ldots, \gamma_{27}$, the precision parameter $\tilde{\tau}_\gamma$ of the mean μ_γ and the gamma parameters $(\tilde{a}_\gamma, \tilde{b}_\gamma)$ of the precision τ_γ. Numerically, each of the precision parameters τ_β, $\tilde{\tau}_\alpha$, and $\tilde{\tau}_\gamma$ was varied from the prior value 0.01 to the larger values 0.02 and 0.10, equivalently, variances ranging from 100 to 10. Each of the gamma prior pairs (a_ξ, b_ξ), $(\tilde{a}_\alpha, \tilde{b}_\alpha)$, and $(\tilde{a}_\gamma, \tilde{b}_\gamma)$ was varied from the prior pair (0.10, 0.10) to the larger values (0.50,0.50) and (1.0, 1.0), equivalently, mean identically 1 and variance ranging from 10 to 1.

The posterior means of $\beta_1, \ldots, \beta_5$, α_{14}, γ_{28}, μ_α, and μ_γ changed by at most 12.5 per cent (0.64 vs 0.72) and at least 0.8 per cent (1.28 vs 1.29). The posterior means of τ_α and τ_γ changed by at most 8.5 per cent (5.50 vs 5.97) and 9.4 per cent (5.24 vs 5.73) due to the variation of τ_β, $\tilde{\tau}_\alpha$, $\tilde{\tau}_\gamma$, and (a_ξ, b_ξ). In contrast, these posterior means changed by over 60 per cent (5.97 vs 2.16) and 67.5 per cent (5.73 to 1.86) as functions of $(\tilde{a}_\alpha, \tilde{b}_\alpha)$ and $(\tilde{a}_\gamma, \tilde{b}_\gamma)$ ranging from (0.10, 0.10) to (1.0, 1.0). Thus, the posteriors of the precision parameters for both the country category effect and the chemo combination effect distributions were quite sensitive to the Level 2 gamma hyperpriors on their precision parameters. As shown by the upper portion of Table VII, however, the sensitivity of τ_γ to its Level 2 prior parameters $(\tilde{a}_\gamma, \tilde{b}_\gamma)$ under the full model had almost no effect on the posterior ranks. For the 10 values of $(\tilde{a}_\gamma, \tilde{b}_\gamma)$ studied, the mean $\tilde{a}_\gamma/\tilde{b}_\gamma$ was either 0.5 or 1 while the variance $\tilde{a}_\gamma/\tilde{b}_\gamma^2$ varied from 0.1 to 10, with both $\bar{R}_a$ and $\hat{R}_a$ robust to these hyperprior parameters despite the fact that the posterior of τ_γ was sensitive. We also considered $(\tilde{a}_\gamma, \tilde{b}_\gamma)$ pairs with variance 100, but the MCMC computations did not converge for these hyperpriors. This is in agreement with the discussion given by Gelman [15] of Level 2 inverse gamma priors on the variance; in our notation, $\sigma_\gamma^2 = 1/\tau_\gamma$ of the γ_c's in hierarchical models, equivalent to the Ga(ε, ε) priors on τ_γ assumed here. Gelman concluded that very small values of ε, corresponding to very high prior variance of τ_γ, are non-truly 'non-informative' and are problematic in that inferences are very sensitive to ε, which is what we saw with our data. Following Gelman's

Table VII. Sensitivity analysis of the Level 2 priors for τ_γ under the full hierarchical model for chemo combination effects, and of fixed values of τ_γ under the NH Chemo model.

	Posterior ranks													
	CARBO		CYC		VP16		IFOS		PRC		VCR		CDDP	
Hierarchical models with $Ga(\tilde{a}_\gamma, \tilde{b}_\gamma)$ *hyperprior on* τ_γ														
$\tilde{a}_\gamma, \tilde{b}_\gamma$	$\bar{R}_a$	$\hat{R}_a$	$\bar{R}_a$	$\hat{R}_a$	$\bar{R}_a$	$\hat{R}_a$	$\bar{R}_a$	$\hat{R}_a$	$\bar{R}_a$	$\hat{R}_a$	$\bar{R}_a$	$\hat{R}_a$	$\bar{R}_a$	$\hat{R}_a$
0.1,0.1	$4.73_{1.86}$	7	$4.57_{1.79}$	6	$4.06_{1.74}$	5	$4.06_{2.31}$	4	$3.88_{2.43}$	3	$3.36_{1.45}$	2	$3.32_{1.81}$	1
10,10	$5.05_{1.76}$	7	$4.78_{1.73}$	6	$3.90_{1.72}$	3	$3.96_{2.25}$	4	$4.01_{2.41}$	5	$3.26_{1.39}$	2	$3.05_{1.73}$	1
2,2	$4.96_{1.80}$	7	$4.79_{1.72}$	6	$3.92_{1.72}$	3	$3.99_{2.27}$	5	$3.96_{2.41}$	4	$3.27_{1.40}$	2	$3.10_{1.74}$	1
1,1	$4.93_{1.80}$	7	$4.73_{1.75}$	6	$3.96_{1.72}$	4	$4.03_{2.28}$	5	$3.93_{2.41}$	3	$3.26_{1.42}$	2	$3.16_{1.75}$	1
0.2,0.2	$4.77_{1.84}$	7	$4.66_{1.77}$	6	$4.01_{1.74}$	4	$4.12_{2.30}$	5	$3.87_{2.42}$	3	$3.31_{1.44}$	2	$3.26_{1.79}$	1
2.5,5	$5.13_{1.74}$	7	$4.83_{1.71}$	6	$3.81_{1.69}$	3	$3.87_{2.23}$	4	$4.13_{2.40}$	5	$3.28_{1.40}$	2	$2.96_{1.71}$	1
0.5,1	$4.97_{1.79}$	7	$4.76_{1.73}$	6	$3.95_{1.70}$	4	$4.02_{2.28}$	5	$3.92_{2.41}$	3	$3.26_{1.41}$	2	$3.11_{1.75}$	1
0.25,0.5	$4.91_{1.79}$	7	$4.73_{1.77}$	6	$3.98_{1.73}$	4	$4.02_{2.28}$	5	$3.96_{2.40}$	3	$3.26_{1.41}$	2	$3.15_{1.76}$	1
0.05,0.1	$4.76_{1.87}$	7	$4.65_{1.78}$	6	$4.00_{1.71}$	4	$4.09_{2.32}$	5	$3.87_{2.41}$	3	$3.35_{1.45}$	2	$3.28_{1.78}$	1
0.025,0.05	$4.66_{1.88}$	7	$4.55_{1.81}$	6	$4.08_{1.73}$	4	$4.11_{2.32}$	5	$3.85_{2.42}$	3	$3.40_{1.46}$	2	$3.36_{1.81}$	1
Hierarchical model with Uniform(0, 1000) *hyperprior on* $\sigma_\gamma = \tau_\gamma^{-1/2}$														
	$4.68_{1.86}$	7	$4.60_{1.79}$	6	$4.02_{1.75}$	4	$4.07_{2.31}$	5	$3.87_{2.43}$	3	$3.40_{1.48}$	2	$3.37_{1.80}$	1
Non-hierarchical models with i.i.d. $N(0, \tau_\gamma)$ *priors on* $\gamma_1, \ldots, \gamma_{23}$														
$\tau_\gamma = 1$	$5.06_{1.74}$	7	$4.62_{1.73}$	6	$4.38_{1.65}$	5	$4.12_{2.25}$	4	$3.95_{2.38}$	3	$2.80_{1.28}$	1	$3.06_{1.71}$	2
$\tau_\gamma = 0.5$	$5.32_{1.62}$	7	$4.70_{1.71}$	6	$4.12_{1.67}$	5	$3.91_{2.20}$	3	$4.04_{2.39}$	4	$2.86_{1.27}$	1	$3.05_{1.73}$	2
$\tau_\gamma = 0.1$	$5.69_{1.44}$	7	$4.77_{1.70}$	6	$3.33_{1.56}$	2	$3.45_{2.08}$	4	$4.58_{2.33}$	5	$3.45_{1.29}$	3	$2.73_{1.66}$	1
$\tau_\gamma = 0.01$	$5.96_{1.21}$	7	$4.76_{1.55}$	5	$2.64_{1.26}$	3	$2.45_{1.63}$	1	$5.07_{2.14}$	6	$4.49_{1.06}$	4	$2.63_{1.50}$	2

Standard deviations are given as subscripts.

suggestion to use a uniform prior on σ_γ instead, we also fit the full hierarchical model assuming a uniform prior on the domain (0, 1000) for σ_γ, also given in Table VII. The resulting posterior integer ranks $\hat{R}_a$ match those obtained with the last four Ga($\tilde{a}_\gamma, \tilde{b}_\gamma$) hyperpriors, and the posteriors of the $\bar{R}_a$'s were nearly identical to those obtained under the Ga(0.025, 0.05) hyperprior, which has mean 0.5 and variance 10.

Our second set of sensitivity analyses focuses on effects of the fixed variance τ_γ of the chemo combination effects $\gamma_1, \ldots, \gamma_{23}$ under the NH Chemo model without a hierarchical structure on these effects, instead assuming that they are iid with N(0, τ_γ) prior. The results are given in the lower portion of Table VII. Recall that, as shown in Table III(b), the DIC values increased substantially as τ_γ was decreased from 1.0 to 0.01. The lower portion of Table VII shows that the posterior mean ranks $\bar{R}_a$ varied slightly, and the integer ranks $\hat{R}_a$ were quite stable for τ_γ=1 or 0.5, with only the values 3 and 4 of $\hat{R}_a$ switching for IFOS and PRC. There was a substantial effect on the ranks as τ_γ was set to the smaller values 0.1 and 0.01, corresponding to var(γ_c)=10 or 100. It thus appears that the posterior ranks are robust to changes in the Level 2 hyperprior parameters for the γ_c's under the full model; for a range of values that give convergence of the MCMC, but for the NH Chemo model the posterior ranks become unstable if a very disperse prior on the γ_c's is assumed. It then appears that, under either the full hierarchical model or NH Chemo model, the fixed prior parameters should be chosen so that prior uncertainty is reasonably represented but is not too large.

A final question is how the proposed methodology works when the values of the γ_c's and thus the ζ_a's are known. To investigate this, we assumed the same 23 chemo combinations as in the CPC data but, in order to focus on ranking, covariates were not included. We assumed that $\log(T) \sim$ N($\gamma_c^*, \tau_\gamma^*$) for a patient given chemo combination c, for fixed values $\gamma_1^*, \ldots, \gamma_{23}^*$ that gave fixed individual agent values $(\zeta_1^*, \ldots, \zeta_7^*) = (2.15, 1.50, 1.48, 1.32, 0.92, 0.85, 0.74)$ corresponding to (CARBO, VP16, CYC, IFOS, VCR, PRC, CDDP), and fixed τ_γ^*. A data set was simulated under this model using each of three different administrative censoring patterns. In each data set, each combination was given to a hypothetical sample of 100 patients, with an additional 100 patients receiving no chemo, for a total sample size of 2400. The data sets were generated assuming administrative censoring at 15 months (Scenario 1), at 25 months (Scenario 2), and no censoring (Scenario 3). The precision parameter τ_γ^*=1 was assumed for Scenarios 1 and 2, but τ_γ^*=4 was assumed for scenario 3 in order to avoid unreasonably long survival times. Each data set was analyzed under each of 14 different assumed models, 10 with hierarchical structure on the γ_c's and 4 without Level 2 priors, corresponding to the 14 models considered in Table VII. Specifically, the first 10 models assumed that $\gamma_1, \ldots, \gamma_K \,|\, \mu_\gamma, \tau_\gamma \sim$ i.i.d. N(μ_γ, τ_γ) with $\mu_\gamma \sim$ N(0, $\tilde{\tau}_\gamma$) and $\tau_\gamma \sim$ Ga($\tilde{a}_\gamma, \tilde{b}_\gamma$), using the 10 ($\tilde{a}_\gamma, \tilde{b}_\gamma$) pairs in Table VII. The last four models assumed that the γ_c's were i.i.d. N(0, τ_γ), with $\tau_\gamma = 1$, 0.5, 0.1, or 0.01.

Under Scenario 1, for all 10 hierarchical models, the posteriors of the γ_c's were quite stable. Consequently, the posteriors of the ζ_a's and in turn the R_a's were very stable, with the $\bar{R}_a$'s and $\hat{R}_a$'s reflecting the actual ordering of $(\zeta_1^*, \ldots, \zeta_7^*)$ perfectly. For the NH models, τ_γ=1 or 0.5 gave posteriors for the R_a's very similar to those given by the hierarchical models, and consequently the same estimated ranks. While prior precision $\tau_\gamma = 0.1$ produced posteriors on the γ_c's with larger standard deviations compared with those for τ_γ=1 or 0.5, or obtained under any of the hierarchical models, the effects on the $\bar{R}_a$'s and $\hat{R}_a$'s were small. Only the ranks of VP16 and CYC were reversed, which is of little practical consequence since their true effects, ζ_2^*=1.5 and ζ_3^*=1.48, were nearly identical by design. The MCMC did not converge for τ_γ=0.01, which is in agreement with the discussion of Gelman [15].

Under both Scenarios 2 and 3, all posteriors were very stable and thus all rank estimates corresponded perfectly to the actual ζ_a^*'s. This may have been due to the fact that, because the sample sizes were large and follow up was sufficiently long in these two cases, all combinations had some deaths and therefore all γ_c's could be estimated reliably either with or without the hierarchical structure on the γ_c's. A general conclusion regarding the full hierarchical model seems to be that, in addition to giving a better fit to the CPC data than the NH Chemo model (Table III(b)), in the type of setting addressed here it helps most when there are small subsample sizes without deaths for some chemo combinations, since posterior estimation of their γ_c's borrows strength from the estimated posterior effects of the other chemo combinations.

7. Discussion

We have defined and estimated the effects and ranks of individual chemo agents, and of agent pairs, based on survival time data from pediatric brain tumor patients treated with a wide variety of different regimes including chemotherapy, radiation and surgery. While our underlying regression model accounts for a variety of possible covariate effects on OS and we apply goodness-of-fit analyses to choose a distributional form, our results rely on a particular assumed structure for the linear component (1), as well as particular prior distribution forms. Our assumptions and analytic methods are motivated by the desire to account for substantial heterogeneity, as well as missing values, in this observational data set. We define individual chemo agent effects and agent pair effects as linear combinations of the chemo combination effects.

It is somewhat encouraging that nearly all of the models and methods produced the same two highest ranked agents, CARBO and CYC, and the same two lowest ranked agents, VCR and CDDP. The messages for single agent effects of IFOS and PRC were less clear in that the posterior distribution of the rank of each was bimodal. For agent pair effects, {CARBO, CYC} and {CARBO, IFOS} were ranked highest and second highest, while {VP16, PRC} and {VP16, CDDP} had the two lowest ranks. Based on these results and preliminary data from a prospective international choroid plexus tumor study (CPT-SIOP-2000), the choroid plexus tumor group decided that the standard arm of the planned future randomized trial, CPT-SIOP-2009, will include CARBO, CYC, and VP16.

An important question is whether severity of disease may have affected the choice of chemo combination, since an association between severity and chemo combination would cause confounding. In practice, however, CPC patients with more advanced disease are given more cycles of chemotherapy, regardless of the particular combination used.

Acknowledgements

The authors thank two referees for their helpful and constructive comments on an earlier version of the paper. This research was partially supported by NCI research grant 2RO1 CA083932.

References

1. Wolff JE, Sajedi M, Brant R, Coppes MJ, Egeler RM. Choroid plexus tumours. *British Journal of Cancer* 2002; **87**:1086–1091.
2. Jeibmann A, Hasselblatt M, Gerss J, Wrede B, Egensperger R, Beschorner R, Hans VH, Rickert CH, Wolff JE, Paulus W. Prognostic implications of atypical histologic features in choroid plexus papilloma. *Journal of Neuropathology and Experimental Neurology* 2006; **65**:1069–1073.
3. Wrede B, Hasselblatt M, Peters O, Thall PF, Kutluk T, Moghrabi A, Mahajan A, Rutkowski S, Diez B, Wang X, Pietsch T, Kortmann RD, Paulus W, Jeibmann A, Wolff JE. Atypical choroid plexus papilloma: clinical experience in the CPT-SIOP-2000 study. *Journal of Neuro-Oncology* 2009; **95**:383–392.
4. Wolff JE, Sajedi M, Coppes MJ, Anderson RA, Egeler RM. Radiation therapy and survival in choroid plexus carcinoma. *Lancet* 1999; **353**:2126.
5. Wrede B, Liu P, Ater J, Wolff JE. Second surgery and the prognosis of choroid plexus carcinoma—results of a meta-analysis of individual cases. *Anticancer Research* 2005; **25**(6C):4429–4433.
6. Wrede B, Liu P, Wolff JE. Chemotherapy improves the survival of patients with choroid plexus carcinoma: a meta-analysis of individual cases with choroid plexus tumors. *Journal of Neuro-Oncology* 2007; **85**:345–351.
7. Ibrahim JG, Chen M-H, Sinha D. *Bayesian Survival Analysis*. Springer: New York, 2001.
8. Kass RE, Steffey D. Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models). *Journal of the American Statistical Association* 1989; **84**:717–726.
9. Gelman A, Hill J. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press: New York, 2006.
10. Laird NM, Louis TA. Empirical Bayes ranking methods. *Journal of Educational Statistics* 1989; **14**:29–46.
11. Shen W, Louis TA. Triple goal estimates in two-stage hierarchical models. *Journal of the Royal Statistical Society, Series B* 1998; **60**:455–471.
12. Shen W, Louis TA. Empirical Bayes estimation via the smoothing and roughening approach. *Journal of Computational and Graphical Statistics* 1999; **8**:800–823.
13. Lin R, Louis TA, Paddock SM, Ridgeway G. Loss function based ranking in two-stage, hierarchical models. *Bayesian Analysis* 2006; **1**:915–946.
14. Carlin B, Louis TA. *Bayesian Methods for Data Analysis* (3rd edn). CRC Texts in Statistical Science. Chapman & Hall: London, 2000.
15. Gelman A. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis* 2006; **1**:515–533.

16. Tanner MA, Wong WH. The calculation of posterior distributions by data augmentation (with discussion). *Journal of the American Statistical Association* 1987; **82**:528–550.
17. Schafer JL. *Analysis of Incomplete Multivariate Data*. Chapman & Hall/CRC: New York, 1997.
18. Little RJA, Rubin DB. *Statistical Analysis with Missing Data* (2nd edn). Wiley: Hoboken, 2002.
19. Gilks WR, Richardson S, Spiegelhalter DJ. *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC: Boca Raton, 1996.
20. Johnson V. A Bayesian χ^2 test for goodness-of-fit. *Annals of Statistics* 2004; **32**(6):2361–2384.
21. Schwarz GE. Estimating the dimension of a model. *Annals of Statistics* 1978; **6**:461–464.
22. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society*, *Series B* (*Statistical Methodology*) 2002; **64**(4):583–639.
23. Volinsky CT, Raftery AE. Bayesian information criterion for censored survival models. *Biometrics* 2000; **56**:256–262.