Using the Bayesian Logistic Regression Model to Determine the Relationship of Demographics and Hyperaldosteronism

A.A. Bartolucci[a], K.P. Singh[b] and S. Bae[b]

[a] Department of Biostatistics, University of Alabama at Birmingham, Birmingham, Alabama, 35294, USA.

[b] Division of Preventive Medicine and UAB Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, Alabama, 35294, USA.

Email:albartol@uab.edu

Abstract: Aldosterone excess (Hyperaldosteronism) has been reported to be a common cause of resistant hypertension. High aldosterone is also linked to sleep apnea. Investigators have reported that primary aldosteronism (PA) has a prevalence of approximately 20% in patients with resistant hypertension. These studies have been based largely on office blood pressure (BP) measurements. Thus, the degree to which hyperaldosteronism contributes to resistance of 24-h blood pressure control is not known. Thus, an ambulatory investigation appeared to be in order.

Two hundred fifty-one patients were referred to the University of Alabama at Birmingham (UAB) Hypertension Clinic for resistant hypertension and were enrolled over a 46-month period. Resistant hypertension was defined as uncontrolled hypertension (>140/90 mm Hg) as determined at 2 or more clinic visits, in spite of use of 3 or more antihypertensive medications at pharmacologically effective doses. Secondary causes of hypertension other than aldosteronism, such as renovascular hypertension, pheochromocytoma or Cushing's syndrome, had been excluded by laboratory analysis and/or radiological imaging as clinically indicated. Subjects with a history of congestive heart failure or chronic kidney disease (creatinine clearance <60 ml/min) were excluded from study participation. In the reporting of the results the investigators noted that there was no difference in mean office BP between aldosterone excess (H-Aldo) subjects and normal aldosterone status (N-Aldo) patients. Daytime, night-time, and 24-h systolic and diastolic BP were significantly higher in H-Aldo compared to N-Aldo males. Daytime, night-time, and 24-h systolic BP were significantly higher in H-Aldo compared to N-Aldo females. Multivariate analysis indicated a significant interaction between age and aldosterone status such that the effects of aldosterone on ambulatory BP levels were more pronounced with increasing age.

The purpose of the present research is to pursue the secondary finding of demographics and its relationship to aldosterone excess using a Bayesian logistic model. Applying a Monte Carlo Markov Chain approach, the authors focused on both age and gender as possible correlates of hyperaldosteronism. Using normal priors for the age and gender parameters we note that both younger age (posterior OR=1.04) and males (posterior OR= 4.33) are more likely to experience aldosterone excess. The 95% posterior credible intervals did not contain unity. Posterior density profiles also indicated marked departure from posterior expected unit odds ratios.

We demonstrate how the model performs under relevant clinical conditions. The conditions are all tested using a Bayesian statistical approach allowing for the robust testing of the model parameters under various stress conditions which we introduce into the model. The convergence of the parameters to stable values are seen in trace plots which follow the convergence patterns This allows for precise estimation for determining clinical conditions under which the logistic pattern will change. We give further numerical and graphical examples of our results.

Keywords: Aldosterone, Bayesian, demographics, logistic, hypertension quality

Bartolucci *et al*. Using the Bayesian logistic model to determine the relationship of demographics and Aldosteronism

1.   INTRODUCTION

Many studies have been performed on examining aldosterone and hypertension (e.g. Calhoun DA, Nishizaka, MK , Zaman, MA et al. 2002, Mattson, C. and Young, WF 2006, and Zacharieva , S., Orbetsova. M, Elenkova, A et al., 2006). Multiple studies point to regular ambulatory blood pressure (BP) monitoring (ABPM) as a relevant predictor of cardiovascular morbidity and mortality (e.g. White, WB, Schulmen, P., McCabe, EJ et al. 1989, Clement, DL., Buyzere, ML , De Bacquer, DA at al. 2003 and Hansen, TW, Jeppesen, J, Rasmussen, S. et al. 2006). This is in place of regular office BP checks of hypertension which may provide greater control of hypertension. Pimenta, E. et al. (2007) reported a study which prospectively compared the 24 hour ABPM levels in resistant hypertensive patients with or without hyperaldesteronism (aldosterone excess) to test the hypothesis that aldosterone excess is associated with sustained resistance to antihypertensive treatment. Aldosterone is a steroid hormone. It plays a central role in the regulation of blood pressure. The authors point out that in observational studies hypertensive patients with primary aldosterone independent of BP measurements are at greater risk of cardiovascular complications than primary hypertensive patients. Their study suggests that higher cardiovascular morbidity and mortality that is associated with hyperaldosteronism is in part BP related. The issue is somewhat confusing. However, their bottom line is that greater cardiovascular risk in patients with aldosterone excess is expected simply on the basis of higher sustained ambulatory BP evaluations. A secondary finding of the data indicated that perhaps aldosterone excess is gender and age related with   excess being more prevalent in the younger age group and in the male population. We re-examined their results in a similar data set data from a Bayesian perspective to determine the exact association between demographics and hyperaldosteronism. The descriptive statistics for our sample are given in Table 1.

One can see from the demographics that there is a higher proportion of aldosterone excess in the males (p< .001, Chi-square test) and a lower average age for the hyperaldosterone group as well. We decided to test the significance of these initial results using a Bayesian approach to the logistic regression procedure.

Table 1. Demographics of Aldosterone Sample

Gender	Aldosterone	N	% within Gender	Age (SD)	95% CI
Male	Low	67	58	54.8 (11.5)	(51.9, 57.6)
	High	49	42	51.2 (10.2)	(48.2, 54.1)
Female	Low	100	86	56.4 (12.0)	(54.0, 58.8)
	High	17	14	50.6 (10.4)	(45.2, 55.9)

2.   METHODS

As mentioned above, the original article by Pimenta et al. (2007) was quite detailed and the published results gave an indication of the association of hyperaldosteronism and BP. We are primarily interested in the association of the demographics and high or low aldosterone. For the purposes of our investigation we apply a logistic sampling model to the data. From the model we derive the logit of the response (high or low aldesterone) and compute the posterior odds ratio of response for each of the model variables. Once this is done then we derive the posterior predicted odds ratio for the group. The authors have utilized the MCMC (Markov Chain Monte Carlo) procedure for deriving the posterior parameters of the model which include the posterior odds and predicted posterior probabilities. The authors have determined that a Bayesian consideration of the results would afford a more coherent interpretation of the effect of the model assuming a random effect of the covariates of gender (sex) and age. Thus, using a Markov Chain Monte Carlo method of parameter estimation with non informative priors, one is able to obtain the posterior estimates and credible regions of estimates of these effects as well as their interaction on logit outcome. Graphical displays of convergence history and posterior densities affirm the stability of most results. We demonstrate how the model performs under relevant clinical conditions. The conditions are all tested using a Bayesian statistical approach allowing for the robust testing of the model parameters under various recursive partitioning conditions of the covariates and hyper parameters which we introduce into the model. The convergence of the parameters to stable values are seen in trace plots (not shown) which follow the convergence patterns This allows for precise estimation for determining clinical conditions under which the response pattern will change. We give a numerical example of our results. The major platform for the theoretical development follows the Bayesian methodology and the multiple parameter logistic model with random effects having carefully chosen hyper parameters. We have done the analysis using the WinBugs software (available at http://www.mrc-bsu.cam.ac.uk/software/bugs/the-bugs-project-winbugs/)   employing   the   MCMC methodology. The BUGS language allows a concise expression of the parametric model to denote stochastic (probabilistic) relationships and to denote deterministic (logical) relationships. We also are aware of some issues with this procedure determining the Monte Carlo convergence and will detail that later.

Bartolucci *et al*. Using the Bayesian logistic model to determine the relationship of demographics and Aldosteronism

For our purposes we use the simple model,

$$\text{logit}(p[i]) = \beta_0 + \beta_1 \text{ sex}[i] + \beta_2 \text{ age}[i] \qquad , i=1,\dots,233. \qquad (1).$$

with the prior input,

$$\beta_0 \quad \sim \text{normal } (0.0, 1.0E\text{-}4) \qquad \text{Prior for intercept} \qquad (2)$$
$$\beta_1 \quad \sim \text{normal } (0.0, 1.0E\text{-}4) \qquad \text{Prior for slope of sex}$$
$$\beta_2 \quad \sim \text{normal } (0.0, 1.0E\text{-}4) \qquad \text{Prior for slope of age}.$$

We will investigate varying coherent prior input for tests of robustness as well.

3. RESULTS

Proceeding with this model, Table 2 gives the posterior results of the model noted in equation (1). Examining Table 2, one is able to interpret the model. The posterior estimate of $\beta_1 = 1.467$ indicates an increase in aldosterone for the males and a negative coefficient estimate of $\beta_2 = -0.03504$ indicates that an increasing age yields a decreasing aldosterone value. Note in both cases that the 95% posterior credible intervals do not contain the value 0 indicating a possibly true influence of these

Table 2. Posterior Parameter Estimates

Variable (parameter)	Mean	SD	Median	95% Credible Interval
β_0	0.070	1.471	0.106	(-1.44, 1.66)
β_1	1.467	0.464	1.453	(0.82, 2.12)
β_2	-0.035	0.022	-0.035	(-0.06, -0.01)
OR Sex	4.34	3.365	2.268	(2.27, 8.35)
OR Age	0.966	0.027	0.965	(0.94, 0.99)
Pred.Sex	0.802	0.117	0.825	(0.52, 0.96)
Pred.Age	0.516	0.170	0.518	(0.19, 0.83)

values on the aldosterone level. We've modified the program to include the posterior odds ratios (OR) and predicted values of the variables of interest. Note that the OR for sex is 4.34 indicating that the odds of high vs. low aldosterone is 4.34 times more likely for males than females. Likewise, the OR for age is 0.966 indicating that the odds of high vs. low aldosterone is 1/0.966 = 1.04 times more likely for younger age than older age individuals. Examining the Table 2 further one notes that the posterior prediction values are given.

The posterior predicted value for sex is 0.802 which is interpreted as a predicted 0.802 chance for a higher aldosterone value for males. The predicted value for age is 0.516 which implies a 0.52 chance of aldosterone excess for the younger age group. The discriminatory cutoff for younger age in this context is less than 54 years of age. Let's examine the density plots for the parameter estimates of the age and sex. On Figures 1 and 2 β_1 and β_2 are labeled as b.sex and b.age respectively as this is how they were coded. We note that for the posterior density of the coefficient for sex that it presents a fairly normal density. The Monte Carlo (MC) error for this parameter did not yield a small error or uncertainty that would be ideal for convergence. Ideally the MC error should be 0.005 or less. In the case of the sex variable parameter the error was about 0.04. This demonstrates the instability of the convergence criteria with the software that the authors have encountered However, the MC error for the other parameters of Table 2 were quite stable with an upper bound of 0.001 for most cases. The plot for the age parameter density was in fact around 0.00007 which is ideal. The number of iterations for the procedure was about 40,000 as can be seen on Figure 1. The posterior density plot for the age parameter is seen in Figure 2. One can see from this plot that the density is fairly normal as well.
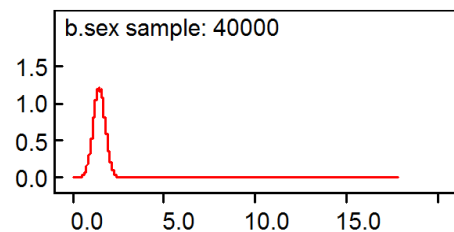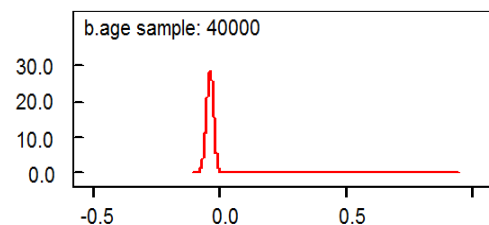

Figure 1. Posterior density for β_1

Figure 2. Posterior density for β_2

Thus it appears that the aldosterone based on the Bayesian investigation is in fact of further excess in males as well as in the younger age group. However, the posterior probability of males having excess aldosterone is around 0.802 which is fairly strong evidence as compared to the posterior probability of age which yields only about a 50/50 chance of excess aldosterone in younger individuals.

Figures 1 and 2 are fairly standard output from the WinBugs program that we have been using. We decided to test for the interaction term of sex and age. The convergence to the posterior value of the coefficient was 0.0154 with an MC error of about 0.002. The posterior 95% credible interval, (-0.05, 0.08), contained the value 0 giving evidence of no interaction effect. Note Figure 3. However, it proved interesting to examine the surface profiler plot of the effect of age and sex on the aldosterone outcome, high or low. Here we get into design space analysis which is very prominent in the pharmaceutical industry with statistical applications in response surface methodology. This involves the multidimensional combination and interaction of input variables (e.g., material attributes) and process parameters that have been demonstrated to provide assurance of quality. In the clinical setting it comprises input variables such as patient demographics, lab values and clinical parameters and possibly genetic markers to determine the optimal patient outcome in terms of maximal response, minimal side effect or marker level such as aldosterone. For the purpose of this presentation and appropriate interpretation of the surface plot, we should share the coding scheme of the variables. It is very simple. For gender the males and, females are nominal and left as such, age is raw age and aldosterone is coded as excessive/high =1 and 0 otherwise.
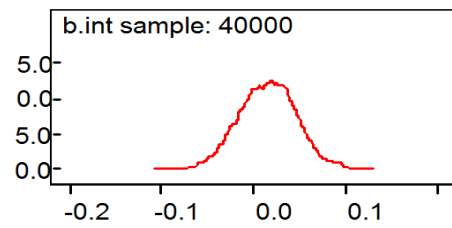

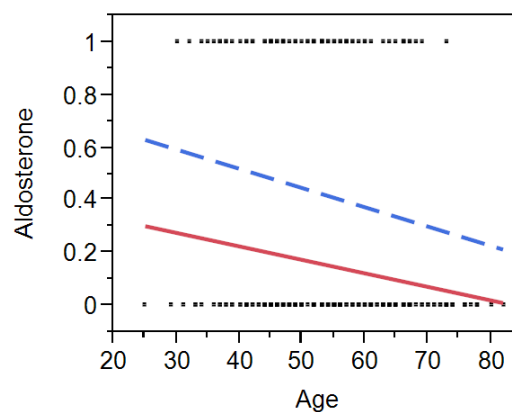**Figure 3**. Posterior density interaction coefficient, b.int

Figure 4. Linear relationship of aldosterone to age and gender

Figure 4 is a linear plot of the relationship of age and sex to the aldosterone level. Clearly, when examined separately, the lines are decreasing with increasing age demonstrating that the aldosterone decreases with increasing age.

Also note that the females (solid line) generally have lower aldosterone levels than the males (dashed line). When one examines the profiler plot or the three dimensional space involving the interaction term in Figure 5 we note that when considering the simultaneous role of sex or gender and age in the model the effect of aldosterone level becomes even more apparent. Clearly as the age increases from 20 to 90 years the surface is on a downward trend. The male gender is to the left of the graph and one sees the increase or the higher level in the surface. The fold in the plain surface indicates the interaction term's role. This allows one to see the drop in aldosterone on the female side. The response surface methodology works quite well in this particular example. We included quadratic terms in the model to determine if there was any possible curvature to the response. However, such was not the case.
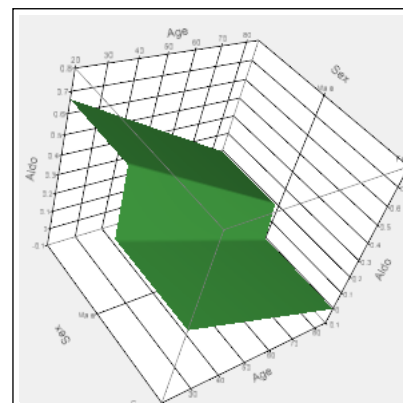

Figure 5. Surface profiler of aldosterone response

We mentioned above that we checked for the robustness of the models under certain conditions. This was done merely by varying the prior input of the means in equation (1) which varied from 0 to 1 leaving the prior variance hyper parameters fairly well dispersed. The same was done for the interaction model which resulted in reasonable contour plots and ultimately the response surface profiler seen in Figure 5. There were issues with convergence for the coefficient for the sex parameter, β_1, as seen by the high Monte Carlo error of

Bartolucci *et al*. Using the Bayesian logistic model to determine the relationship of demographics and Aldosteronism

about 0.04. Added simulations or iterations did not yield any better results. However, given the clinically meaningful result of the gender relationship to aldosterone in our analysis, suffice to it say we are content at this point to acknowledge this one minor concern in our results.

4.   DISCUSSION AND CONCLUSIONS

As mentioned above, the original article by Pimenta et al. (2007) was quite detailed and the published results gave an indication of the association of hyperaldosteronism and BP. We were primarily interested in the association of the demographics and high or low aldosterone. For the purposes of our straight forward investigation the task is to apply a logistic sampling model to the data. From the model we computed the posterior odds ratio of response for each of the model variables. Once this is done then we derived the posterior probabilities as well. The authors have utilized the MCMC (Markov Chain Monte Carlo) procedure for deriving the posterior parameters of the model which include the posterior odds and posterior predicted probabilities of the results for sex and age performance in aldosteronism which was consistent with higher posterior rates of excessive aldosterone in males and the younger age group. A novel approach involved the use of the surface profiler to demonstrate the simultaneous performance of age and gender influence on levels of aldosterone.

The authors have demonstrated that a Bayesian consideration of the results would afford a reasonably coherent interpretation of the logistic model assuming a random effect of the covariates of gender (sex) and age. Thus, using a Markov Chain Monte Carlo method of parameter estimation with non informative priors, one is able to obtain the posterior estimates and credible regions of estimates of these effects as well as their interaction on logit outcome. The models used were robust and convergence for the most part was tractable for the parameter estimates.

REFERENCES

Calhoun, D.A., Nishizaka, M.K., Zaman, M.A., Thakker, R.B. and Weissmann, P. (2002). Hyperaldosteronism among black and white subjects with resistant hypertension. *Hypertension.* 40, 892-896.

Clement, D.L., DeBuyzere, M.L.,DeBacquer, D.A., deLeeuw, P.W., Duprez, D.A., Fagard, et.al. (2003). Prognostic value of ambulatory blood pressure recordings in patients with treated hypertension. *N. Engl J Med,* 348, 2407-2415.

Hansen, T.W., Jeppesen, J., Rasmussen, S., Ibsen, H. and Torp-Pedersen, C. (2006). Ambulatory blood pressure monitoring and risk of cardiovascular disease: a population based study. *Am. J. Hypertens*, 19, 243-250.

Mattson, C. and Young, W.F. Jr. (2006). Primary aldosteronism: diagnostic and treatment strategies. *Nat Clin Pract Nephrol,* 2, 198-208.

Pimenta , E., Gaddam, K..K., Pratt-Ubunama, M.N., Nishizaka, M.K., Cofield, S.S., Oparil, S. and Calhoun, D.A. (2007). Aldosterone excess and resistance to 24-h blood pressure control. *J Hypertens,*, 25, 2131-2137.

White, W.B., Schulmen, P.,McCabe, E.J.and Dey, H.M., (1989). Average daily blood pressure, not office pressure, determines cardiac function in patients with hypertension. *JAMA*, 261, 873-877.

Zacharieva,S., Orbetzova, M., Elenkova, A., Stoynev, A., Yaneva, M., Schigarminova, R. et al.   (2006). Diurnal blood pressure patterns in patients with primary aldosteronism. *J. Endocrinol Invest*, 29, 26-31.