

Simulation in models of health care quality

Karl Heiner, State University of New York at New Paltz
David Laws, University of Sheffield
David Walshaw, University of Newcastle

Abstract

The AIDS Institute of the New York State Department of Health monitors the quality of care delivered by hospitals, community health centers and drug treatment centers to individuals infected with HIV. A medical peer review organization visits these facilities each year and applies a number of protocols reflecting the standard of care to random samples of medical records. Bayesian techniques are used to model aspects of the quality of care. Both conjugate and non conjugate models are employed in the analysis of the data. For models that are not conjugate Gibbs sampling is used to simulate posterior distributions of parameters. In all cases, in order to make inferences from the models, simulation methods are applied and trellis graphics are used to display overall and among facility trends.

Keywords: healthcare quality, Bayesian models, simulation, MCMC, beta binomial model, BUGS.

1 Introduction

The purpose of this paper is to describe how simulation is used in the analysis of data monitoring the quality of medical care delivered to individuals infected with HIV residing in New York State (NYS). We will describe three situations in which we use simulation to produce information that managers at the AIDS Institute (AI) of the New York State Department of Health (NYS DOH) use to monitor health care quality. In each of these three situations we use Bayesian modeling techniques.

In the first situation, we describe a beta-binomial conjugate model that produces quality scores and report cards for each of approximately 130 medical care providers each year for a number of quality of care indicator random variables. Despite the conjugate nature of this model, the answers to certain questions of interest must be answered with simulation. In the second situation, we describe a model to analyze data that monitors compliance with a NYS law requiring HIV counseling to

all pregnant women. Here, the model is not a conjugate model and Gibbs sampling is employed to produce draws from the posterior distributions of parameters of interest. In the third situation, we revisit situation one and describe a more complicated model to incorporate information from facility self reporting.

2 Background

In NYS, certain quality of care measures are indicator variables that reflect the presence of (or lack of) evidence in an individual's medical record of important aspects of this care. The medical care in question is care delivered to individuals infected with human immunodeficiency virus (HIV). The data are collected by the AIDS Institute (AI) of the New York State (NYS) Department of Health (DOH). More than one fifth of the US population of AIDS patients resides in NYS. It is the mission of the AI to monitor the care provided to these individuals. They do so by contracting with a Peer Review Organization (PRO) to annually visit approximately 130 medical care providers with a team of Peer Review Agents (PRA's), mainly nurses who have been trained to collect the relevant information of individual medical records. The relevant information is defined by a series of fairly elaborate protocols that have been developed by the Medical Director of the AI and a group of physicians specializing in HIV cases who advise the AI.

3 Key HIV Quality of Care Indicators, a Conjugate Model

Of particular interest is a rather small subset of the information elicited by the protocols. These are called the key quality indicators and include measures that reflect the diligence of the provider in monitoring the progression of the disease and treatment effectiveness. These key quality indicators reflect monitoring of CD4 counts and viral loads, testing for *Pneumocystis carinii* pneumonia

(PCP) prophylaxis, PPD screening and pelvic examinations for women. Random samples of medical records from each facility are reviewed each year. Females are over sampled to provide sufficient records to evaluate the pelvic exam indicator. For most facilities, this process includes eight such annual reviews by the PRO. Historically, each facility received a score on each of the five key quality indicators. This score is the mean of the posterior distribution of a beta-binomial conjugate model, where the parameters of the prior distribution are obtained by successfully down weighting parameters from earlier posterior distributions. This is analogous to exponentially weighted moving averages.

We wish to produce a posterior distribution of the probability of a “yes” for each indicator in each facility for the current year. We would also like to produce a “report card” for each facility for the current year, i.e. an estimate of the proportion of each facility’s current case load that had a positive score on each indicator. Although we feel it is desirable to construct a hierarchical model where among facility distributions are included, the use of Bayesian models to report public health information in the United States is somewhat novel, therefore the AI is reluctant to publish quality scores for individual facilities that incorporate information about other facilities. We believe that in the future this will be possible, since, under a hierarchical model, facility scores will shrink slightly towards their mean. One would not expect objections from the stronger facilities, and weaker facilities’ scores will be slightly improved by this shrinkage. We will describe the conjugate model that has been used to produce quality scores for all 130 facilities for the past eight years which is based only on data collected by the PRO.

In a sample of size n , drawn from a sampling frame of HIV related medical records at a particular medical provider, we would like to estimate the proportion of the sampling frame possessing a particular attribute, S , e.g. the proper documentation of a thorough pelvic exam. The number of these successes in the sample follows the binomial form. Let X is the number of successes and θ be the probability of success on any one of n independent trials. When n and X are known and θ is not we have the likelihood

$$f(\theta|x) = \binom{n}{x} \theta^x (1 - \theta)^{n-x}.$$

The beta distribution is a continuous family of distributions on the interval $(0, 1)$. With $\alpha > 0$ and $\beta > 0$ its pdf is

$$f(y|\alpha, \beta) = \frac{\Gamma(\alpha, \beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1} (1 - y)^{\beta-1}$$

where $\Gamma(\alpha)$ denotes Euler’s gamma function.

We use the beta distribution as the form of the prior distribution for the binomial likelihood. It is the conjugate prior for the binomial likelihood, and the family can assume a variety of shapes which makes it convenient for subjective priors. Thus, following Bayes theorem

$$p(\theta|X) = \frac{p(\theta)p(X|\theta)}{\int p(\theta)p(X|\theta)d\theta}.$$

So,

$$p(\theta|X) = \frac{\Gamma(n + \alpha + \beta)}{\Gamma(x + \alpha)\Gamma(n - x + \beta)} \theta^{x+\alpha-1} (1 - \theta)^{n-x+\beta-1}$$

is the beta distribution with parameters $x + \alpha$ and $n - x + \beta$.

We may now use this posterior distribution to obtain estimates of the parameter of interest, here this is θ .

3.1 Predictive distribution and finite population sampling

Should we be interested in predicting the number of successes, y in the next m trials, the predictive distribution is the integral of the product of the likelihood and the posterior distribution.

$$\begin{aligned} p(y|n, x, \alpha, \beta) &= \int p((y|\theta)p(\theta|n, x, \alpha, \beta)d\theta \\ &= \binom{m}{y} \frac{\Gamma(n + \alpha + \beta)}{\Gamma(x + \alpha)\Gamma(n - x + \beta)} \\ &\times \frac{\Gamma(y + x + \alpha)\Gamma(m - y + n - x + \beta)}{\Gamma(m + n + \alpha + \beta)} \end{aligned}$$

a Polya probability mass function that may be used to make predictions and develop prediction intervals.

In order to accomplish this we need to calculate quantiles of the Polya distribution. Here we are working with a probability distribution that is not commonly found in statistical software, nor in the standard add-on packages for symbolic programming languages, e.g. Mathematica. We would like to define a discrete probability distribution and work with it as if it were defined in the Mathematica standard add-on packages. In particular, we would like to use the Quantile function on it. We know the probability mass function, pdf, as well as the recurrence formula. The pdf for the beta-binomial random variable, x , is a function of three parameters, n , a , and b . The parameter n is a positive integer and x takes integer values between 0 and n . In Mathematica, the pdf may be expressed as

```
betaBinomialPDF[x_, n_, a_, b_] :=
```

```
Binomial[n,x] (Beta[a+x,b+n-x])/(Beta[a, b])
```

We wish to define `betaBinomial[n, a, b]` so that `PDF[betaBinomial[n,a,b,x]` would return the above pdf or some variant there of, and so that `Quantile[betaBinomial[n,a,b],q]` would return the q^{th} quantile of this distribution.

```
betaBinomialCDF[x_, n_, a_, b_] :=
```

```
Sum[(Binomial[n,t] (Beta[a+t,b+n-t])/(Beta[a,b])),  
{t, 0, x}]
```

```
betaBinomialQuantile[n_, a_, b_, q_]
```

```
:= Module[{low, high, mid},  
If[q == 1, Return[n];  
If[N[betaBinomialPDF[0, n, a, b]] < q,  
low = 0; high = n;  
While[high - low > 1,  
mid = Floor[(high + low)/2];  
If[N[betaBinomialCDF[mid, n, a, b]] < q,  
low = mid,  
high = mid,  
high = Fail; Break[]  
]  
]; high,  
0,  
Fail  
]  
]
```

A hospitals performance may be viewed as a report card for the current review period and when the sampling fraction for any review period is not small, i.e. $f > 0.10$, issues of finite population sampling are relevant. To develop probability intervals for Θ , the proportion of charts documenting a key service, we utilize the beta-binomial predictive distribution to predict the number of successes, y , in the medical records not sampled, m . By developing a prediction interval for y , (y_L, y_U) , adding x and dividing the sum by $m + n$, a probability interval for Θ is obtained.

3.2 Discounting prior information

If for a prior on θ we adopt the beta distribution with $\alpha = 1$ and $\beta = 1$, a flat prior, then the posterior distribution of θ given x success and $n - x$ failures would be the `BetaDistribution[1+x, 1+n-x]`. If we use the posterior

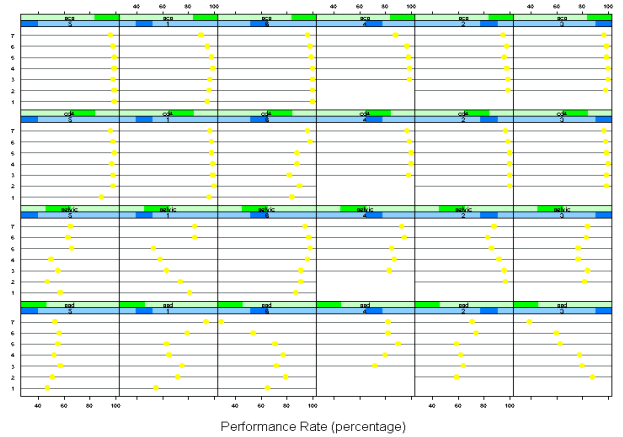


Figure 1: A trellis display of six individual hospital (trellis columns) performance trends on four key quality indicators (trellis rows).

from the first review period as the prior for the second, we merely add the number of successes to α and the number of failures to β to give the posterior distribution after the second review, which is `BetaDistribution[1 + $x_1 + x_2$, 1 + $(n_1 - x_1) + (n_2 - x_2)$]`. Consequently, after r review cycles we would be left with a posterior density on θ with $\alpha = 1 + \sum_{t=1}^r x_t$, and $\beta = 1 + \sum_{t=1}^r (n_t - x_t)$. Note that the information from each trial has equal weight. We would like the most recent review to carry more weight than the previous one, etc.

We specify a discount factor c so that at review $t = r$, the first parameter of the posterior beta distribution is $c^{r-1}\alpha_0 + \sum_{t=1}^r c^{r-t}x_t$ and the second parameter is $c^{r-1}\beta_0 + \sum_{t=1}^r c^{r-t}(n_t - x_t)$, where α_0 and β_0 are the beta distribution parameters for the initial prior distribution, e.g. $\alpha = 1$ and $\beta = 1$. By discounting both parameters by the same factor the means of the posterior distributions of θ form an exponentially weighted moving average. The maximum likelihood estimate of the discount factor c was 0.55 for the current data. This is consistent with the analogous value used in exponentially weighted moving average quality control charts used by production engineers (Box and Luceno, 1997). Figure 1 is a trellis diagram depicting the trends for one of the indicators for a set of providers.

3.3 Pairwise Comparisons

Of particular interest to certain members of the patient community are between facility comparisons. The form of linear combinations of beta distributed random variables is not a standard distribution. Thus in order to

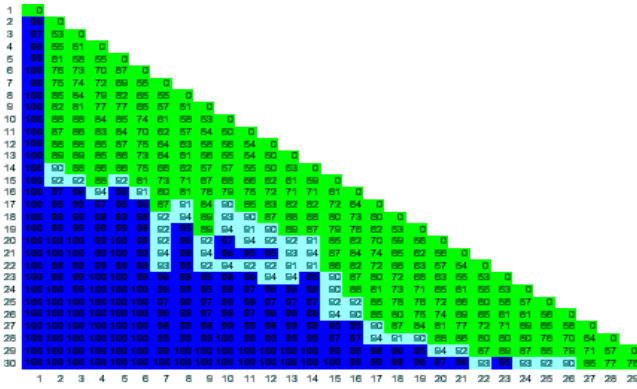


Figure 2: Pairwise comparisons between pairs of health care providers within a specific class of providers on one of the quality indicators.

obtain an estimate of the probability that one facility outperforms another, we simulate from their respective posterior distributions. For each of the 130 facilities, we made 1000 draws from the most recent posterior distribution of theta. If we think of these 130 vectors of 1000 draws as columns of a matrix, we may construct a 1000 by 130x129 matrix of pairwise cell differences within rows. By counting the numbers of values in each column of this matrix that exceed zero we have an indication of whether differences between pairs of facilities could have happened by chance.

Figure 2 is a display for forming pairwise comparisons between certain pairs of providers on one of the quality indicators. These providers belong to one of four classes of providers that are of particular interest. Provider row and column names have been deleted since these data are not yet released to the public. In this figure, providers are ordered by the mean of their posterior distribution for θ . The entries in the figure indicate the percent of draws from the posterior distributions where the first provider's θ exceeded that of the second. Percents of 95 or more have been colored red, those between 90 and 94 yellow, and the remainder are colored green.

4 HIV Counseling for Pregnant Women, Gibbs Sampling

NYS regulations require that health care providers serving pregnant women counsel them about HIV transmission, testing, therapy and potential infection of babies born to HIV infected women. Furthermore, the regulations require that these women be offered an HIV test (NYSDOH AI, 2000). This policy is consistent with recommendations from the Institute of Medicine, the Amer-

ican Academy of Pediatrics, and the American College of Obstetricians and Gynecologists (ACOAG,1999).

In an effort to monitor these requirements, the AIDS Institute (AI) of the New York State (NYS) Department of Health (DOH) reviews a sample of the medical records of pregnant women. The reviews are conducted by PRA's, usually nurses supervised by doctors, who are employed by a Peer Review Organization (PRO). In each review, various information is recorded, including answers to three key questions, did the pregnant women receive HIV counseling, was a counseled woman tested for HIV, and was the woman counseled after the HIV test.

Of the institutions in New York State, our data concerns a subgroup of 61 of them which can be classified as one of three types of centers:

1. CHCs - Community Health Centers
2. DACs - Designated AIDS Centers
3. NDAs - Hospitals Not Designated as AIDs centers

Information collected in the peer review visit includes sample data on the follow three indicators:

1. Whether pregnant women received HIV counseling.
2. Whether those pregnant women who were given counseling then had a test.
3. Whether pregnant women who were given an HIV test received post-test counseling.

Note that the first and third of these can be regarded as indicators of the quality of care of patients. The second indicator can not as it is for the patient to decide whether to have a test.

4.1 The Model

For each of the three key activities, we wish to estimate the probability that the activity would occur for a pregnant woman visiting a particular facility. In addition, we would like to compare facility types to determine whether there is a difference in quality of care between the different types. Thus we propose the following Bayesian model which requires us to use Monte Carlo methods, in this case Gibbs Sampling. We use the WinBUGS software to perform the analysis.

For each indicator the model is a logistic regression model allowing for over-dispersion:

$$r_i \sim bin(n_i, p_i)$$

$$\text{logit}(p_i) = \beta_i$$

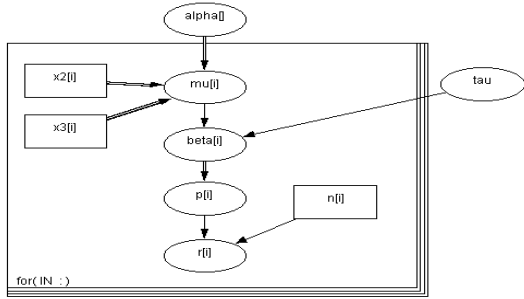


Figure 3: Graphical Representation of the model.

$$\beta_i \sim N\left(\mu_i, \frac{1}{\tau}\right)$$

$$\mu_i = \alpha_1 + \alpha_2 x_{2i} + \alpha_3 x_{3i}$$

where

- r_i is (e.g.) the number of pregnant women out of n_i counseled in institute i ,
- p_i is the probability of a pregnant women being counseled in location i ,
- μ_i is the linear predictor for institute i ,
- β_i varies around the linear predictor to allow for over-dispersion
- τ is the precision parameter for the distribution of β_i around μ_i .
- the parameters of the linear predictors are given by treatment contrasts:
 - α_1 is the mean of $\text{logit}(p_i)$ in CHC's
 - α_2 is the difference in the mean between DAC's and CHC's
 - α_3 is the difference in the mean between NDA's and CHC's

This dependence structure is given in the graph, figure 3. Standard vague priors for this model are as follows:

$$\tau \sim \text{gamma}(0.001, 0.001)$$

$$\alpha_i \sim N(0, 1/1E - 6)$$

where each α_i is assumed independent of the others. However, there are problems with this prior. Firstly it does not acknowledge that the parameters of the linear predictor will be correlated, and it gives different prior

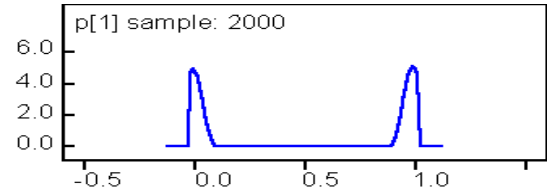


Figure 4: Plot of the simulated prior density for p_i .

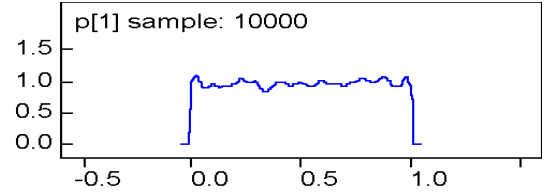


Figure 5: Plot of the simulated marginal prior density of p_i under alternative priors for $\underline{\tau}$ and $\underline{\alpha}$.

variances to $\text{logit}(p_i)$ for the three different types of institutions (CHCs have smaller prior variance than the others.). But, worst of all, it induces an unusual prior structure for the p_i 's, as shown in the plot of the simulated prior density for p_i below, figure 4.

Alternatively, we found the following prior to be sensible:

$$\tau \sim \text{gamma}(1, 1)$$

$$\underline{\alpha} \sim \text{MVN}_3(\underline{0}, \Sigma)$$

where $\underline{0}$ is a vector of zeroes and

$$\Sigma = \begin{pmatrix} 1 & -0.05 & -0.05 \\ -0.05 & 0.1 & 0 \\ -0.05 & 0 & 0.1 \end{pmatrix}$$

This prior is vague about the p_i 's both marginally, and in terms of their pair-wise joint distributions. Figure 5 indicates this for p_1 marginally: The above prior also gives the same prior variance to $\text{logit}(p_i)$ for each type of center.

4.2 Results

The following results are from Gibbs sampling using WinBUGS using the above model. 5000 burn-in samples were discarded and the plots and statistics below are obtained from a further 10000 samples. There was little evidence of auto-correlation and so we did not systematically sample the generated values, e.g., every 50th generated value. q_i denotes the probability of a success (e.g. a

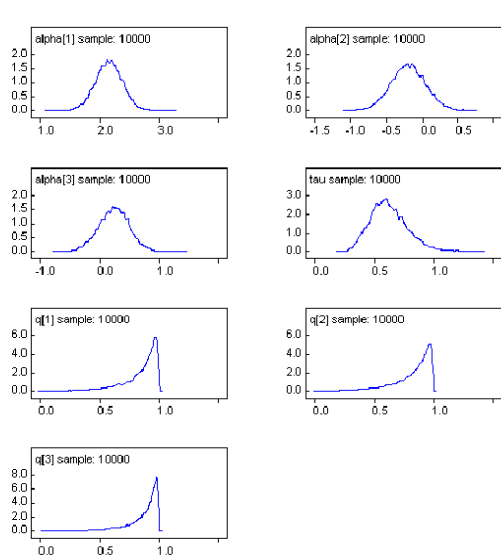


Figure 6: Density traces of similar values for model parameters for pregnancy counseling.

pregnant woman being counseled) in a new/unvisited facility of type i . Therefore, $pr(q_i > q_j)$, for $i \neq j$, denotes the probability that a new facility of type i has a higher proportion/probability of a woman being counseled than facility type j .

4.2.1 Pregnancy counseling

Figure 6, gives density traces for each of the model parameters for the pregnancy counseling indicator. These simulations appear quite reasonable and will allow us to answer questions about differences in facility type.

The table below reports statistics on the simulated parameters in our model for the pre-test counseling.

Parameter	Mean	sd	2.5%tile	Median	97.5%tile
α_1	2.151	0.23	1.707	2.151	2.597
α_2	-0.193	0.25	-0.673	-0.194	0.294
α_3	0.234	0.25	-0.263	0.238	0.735
q_1	0.835	0.16	0.369	0.893	0.992
q_2	0.819	0.17	0.340	0.877	0.989
q_3	0.864	0.15	0.441	0.917	0.994
τ	0.625	0.16	0.365	0.607	0.983

It seems that there is a difference but not a 'significant' one. The NDAs are better than the CHCs which are better than the DACs.

Estimates of the probability as to which are better are (obtained by counting the number of times $q_i > q_j$ in our simulation, and then dividing by 10000:

$$pr(q_1 > q_2) = 0.5362,$$

$$pr(q_1 > q_3) = 0.4358,$$

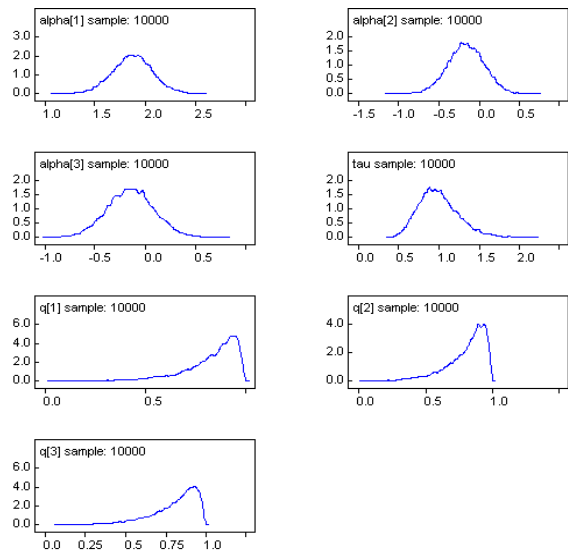


Figure 7: Density traces of similar values for model parameters for HIV testing.

$$pr(q_2 > q_3) = 0.4079.$$

Thus the apparent differences in facility type are not significant.

4.2.2 HIV testing

Figure 7, gives density traces for each of the model parameters for the HIV test indicator.

As above, the table below reports statistics on the simulated parameters in our model for the HIV testing.

Parameter	Mean	sd	2.5%tile	Median	97.5%tile
α_1	1.871	0.1969	1.488	1.871	2.267
α_2	-0.1605	0.2282	-0.6071	-0.161	0.2823
α_3	-0.1491	0.2356	-0.6088	-0.15	0.3173
q_1	0.8277	0.1382	0.4603	0.8682	0.9809
q_2	0.8031	0.1521	0.4022	0.8462	0.9785
q_3	0.8056	0.1506	0.4156	0.8475	0.9788
τ	1.0	0.2481	0.5917	0.9731	1.565

The CHCs have the highest probability of a counseled woman being tested followed by the DACs and then the NDAs.

The probabilities below give an indication of whether there is a real difference:

$$pr(q_1 > q_2) = 0.5533,$$

$$pr(q_1 > q_3) = 0.5526,$$

$$pr(q_2 > q_3) = 0.4951.$$

Thus it is apparent that differences in facility type are not significant.

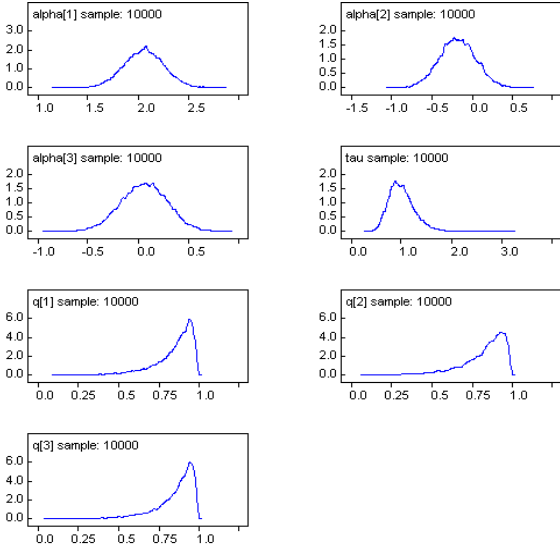


Figure 8: Density traces of similar values for model parameters for post-test counseling.

4.2.3 Post-test counseling

Figure 8, gives density traces for each of the model parameters for the post-test counseling indicator.

The table below reports statistics on the simulated parameters in our model for the post-test counseling testing.

Parameter	Mean	sd	2.5%tile	Median	97.5%tile
α_1	2.052	0.1992	1.663	2.053	2.443
α_2	-0.1795	0.2299	-0.6273	-0.1817	0.2749
α_3	0.05454	0.2343	-0.4072	0.05712	0.5146
q_1	0.847	0.1296	0.4843	0.8858	0.9831
q_2	0.8262	0.1399	0.4516	0.8652	0.9822
q_3	0.8541	0.1264	0.5032	0.8924	0.9859
τ	1.004	0.2543	0.5949	0.9738	1.589

This time the NDAs are best, followed by the CHCs, and then last come the DACs.

The final set of estimated probabilities for comparisons:

$$\begin{aligned} \text{pr}(q_1 > q_2) &= 0.5475, \\ \text{pr}(q_1 > q_3) &= 0.4853, \\ \text{pr}(q_2 > q_3) &= 0.4253. \end{aligned}$$

Again it is clear that there is little evidence that one facility type is “better” than any other.

5 Facility Self Reporting, Markov Chain Monte Carlo

We now develop the non-conjugate model which will be used for providing inferences to the AI from now on.

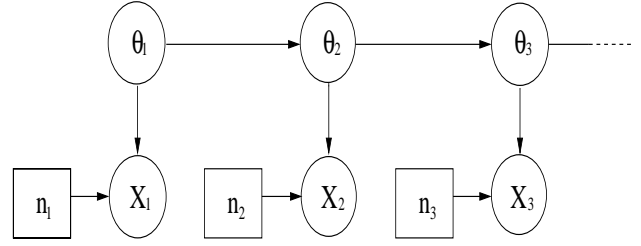


Figure 9: Directed graph of model for PRO outcomes.

5.1 A model for annual peer-reviewed success probabilities

Considering the assessments made by the PRO as definitive, we develop a model for the observations on a single indicator for a single facility. For year $i; i = 1, 2, \dots, t$, we suppose that independent and identically distributed binary outcomes (0 or 1) have been obtained for each of n_i individuals. Denoting the sum of these outcomes by X_i , we have

$$X_i \sim B[\theta_i, n_i] \quad (1)$$

where $\theta_i; i = 1, \dots, t$ is the success probability in year i .

It is reasonable to suppose that the values of $\theta_i; i = 1, 2, \dots, t$ are not independent, and we incorporate this belief into the prior specification for θ_i by conditioning on the value of θ_{i-1} . We use a Beta prior for θ , and set $E[\theta_i] = \theta_{i-1}$. It follows that

$$\theta_i | \theta_{i-1} \sim \text{Be}[a\theta_{i-1}, a(1 - \theta_{i-1})] \quad (2)$$

for $i = 1, \dots, t$, where $a > 0$ determines $\text{Var}[\theta_i]$, and is chosen to represent our beliefs about how similar θ_{i-1} and θ_i will be. We use a value $a = 10$, which corresponds to an amount of information equivalent to that provided in a sample of size 10, and is consistent with discount factors we have found to be appropriate in the conjugate model. For θ_1 , the success probability in the first year for which data are obtained by the PRO, we use a flat $\text{Be}(1, 1)$ prior. Figure 9 shows a representation, as a directed graph, of the model as it applies to the first few years.

The Beta distribution is chosen arbitrarily as a commonly used prior for the binomial distribution. The model structure specified in (2) determines that the conjugacy of this choice only applies to the posterior distribution for θ_1 , while the posteriors for $\theta_2, \theta_3, \dots$ are not tractable.

5.2 A model for the reliability of HIVQUAL

The single year for which HIVQUAL data has been collected so far is the most recent year of assessment, year

t in the notation of equation (1). From a facility population of N patients, a sample of n has been assessed using HIVQUAL, and a subsample r of these has had their assessments verified by the PRO. We regard the PRO assessments as representing the true outcomes, and model the HIVQUAL outcomes by assuming conditional independence of the H_k given the P_k , and:

$$\begin{aligned} (a) \quad \eta &= \Pr(H_k = 1 | P_k = 0); \\ (b) \quad \lambda &= \Pr(H_k = 1 | P_k = 1); \end{aligned} \quad (3)$$

where H_k and P_k are respectively the HIVQUAL and PRO assessments for patient $k = 1, \dots, n$, the PRO outcomes being unobserved for the $n - r$ individuals in the HIVQUAL sample who have not had their assessments verified by the PRO.

From the arguments leading to (1), we have

$$\begin{aligned} P_k &\sim \text{Bern}[\theta_t]; \\ H_k | P_k &\sim \text{Bern}[(1 - P_k)\eta + P_k\lambda] \end{aligned} \quad (4)$$

for $k = 1, \dots, n$.

We assign logistic normal prior distributions for η and λ :

$$\begin{aligned} \log(\eta/(1 - \eta)) &\sim N[\mu_0, \sigma_0^2]; \\ \log(\lambda/(1 - \lambda)) &\sim N[\mu_1, \sigma_1^2], \end{aligned} \quad (5)$$

where (μ_0, σ_0^2) and (μ_1, σ_1^2) allow the specification of our prior beliefs about η and λ , the probabilities that HIVQUAL agrees with a negative and positive PRO assessment respectively. In consultation with an expert we elicited values of $\mu_0 = -1, \mu_1 = 1$ and $\sigma_0^2 = \sigma_1^2 = 4$. These values give a prior modal probability of about 0.75 for a HIVQUAL agreement with each of a negative and a positive assessment by the PRO. Figure 10 shows a directed graph representation of the model as it applies to the most recent two years, where the most recent (HIVQUAL) year has a rectangular ‘plate’ to handle each of the individuals measures separately.

6 Summary

We have illustrated three aspects of healthcare quality monitoring for which we have made use of simulation methods in addressing the research questions of interest using Bayesian methods. The methods used range from sampling schemes to make draws and comparisons based on known distributions, to the construction of non-conjugate models in which the posterior distributions are not analytically tractable. In all cases, it is the simulation which enables us to make the inferences required by the users in evaluating the quality of care.

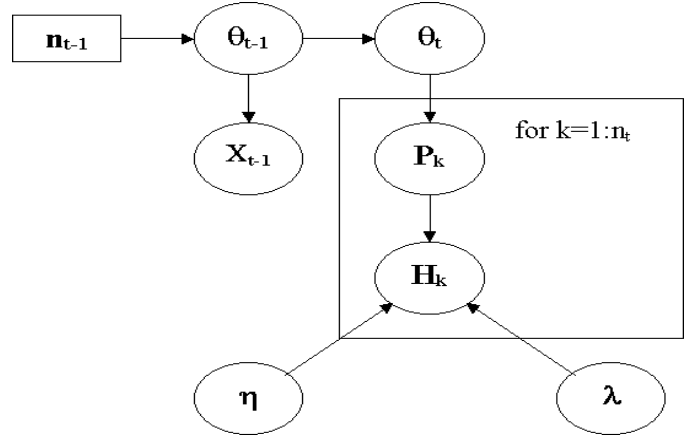


Figure 10: Directed graph of model for HIVQUAL.

Acknowledgements

We wish to thank the NYS DOH AIDS Institute, and in particular it’s Medical Director Dr. Bruce Agins for motivation and access to data, Tony O’Hagan for his extremely helpful suggestions concerning the model specification and to our colleague Arilee Bagley for helpful suggestions and preparing our TEX document.

References

- ACOG (American College of Obstetricians and Gynecologists) (1999). Joint Statement of the AAP and ACOG. *Pediatrics*, **104**, 128-9.
- Box, G. and Luceno, A. (1997). *Statistical Control by Monitoring and Feedback Adjustment*. New York: John Wiley & Sons, Inc.
- Cleveland, W. (1993). *Visualizing Data*. Summit, NJ: Hobart Press.
- Hunter, J. S. (1986). The exponentially weighed moving average. *Journal of Quality Technology*, **18**, 203-210.
- NYSDOH AI (New York State Department of Health AIDS Institute) (2000). *Prevention of Perinatal HIV Transmission*.
- Pole, A., West, M. and Harrison, J. (1994). *Applies Bayesian Forecasting and Time Series Analysis*, New York: Chapman & Hall.
- Roberts, S. W. (1950). Control chart test based on geometric moving averages. *Technometrics*, **1**, 239-250.

Smith, M. A., Atherly, A. J., Kane, R. L., and Pacala, J. T. (1997). Peer Review of the Quality of Care. *J. Am. Med. Assn.*, **278**, 1573-1578.

West, M. and Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*. New York: Springer.

Wolfram, S. (1996). *The Mathematica Book*. Champaign, IL: Wolfram Research, Inc.