# A Bayesian Control Chart for a Common Standardised Normal Mean

R. van Zyl[1], A.J. van der Merwe[2]

[1]Quintiles International, ruaanvz@gmail.com

[2]University of the Free State

UNIVERSITY OF THE FREE STATE
UNIVERSITEIT VAN DIE VRYSTAAT
YUNIVESITHI YA FREISTATA

UFS·UV
NATURAL AND AGRICULTURAL SCIENCES
NATUUR- EN LANDBOUWETENSKAPPE
MATHEMATICAL STATISTICS AND ACTUARIAL SCIENCE
WISKUNDIGE STATISTIEK EN AKTUARIËLE WETENSKAP

## Abstract

By using the medical data analysed by Kang, Lee, Seong, and Hawkins (2007) a Bayesian procedure is applied to obtain control limits for the standardised mean. Reference and probability matching priors are derived for a common standardised mean across the range of sample values. By simulating the posterior predictive density function of a future standardised mean it is shown that the inverse of the control limits for the standardised mean are effectively identical to those calculated by Kang et al. (2007) for the coefficient of variation. This article illustrates the flexibility and unique features of the Bayesian simulation method for obtaining the posterior predictive distribution of $\delta = \frac{\mu}{\sigma}$ (the population standardised mean) predictive interval and run lengths for the future sample standardised means. A simulation study shows that the 95% Bayesian confidence intervals for $\delta$ has the correct frequentist coverage.



**Keywords:** control charts, probability-matching prior, reference prior, standardised normal


# 1 Introduction

The monitoring of variability is a vital part of modern statistical process control (SPC). Shewart control charts are widely used SPC tools for detecting changes in the quality of a process. In most settings where the process is under control the process have readings that have a constant mean ($\mu$) and constant variance ($\sigma^2$). In such settings the $\bar{X}$ chart is usually used to monitor the mean, and the R and S control charts the variance of the process.

In practice there are some situations though where the mean is not a constant and the usual SPC reduces to the monitoring of the variability alone. As a further complication it sometimes happens that the variance of the process is a function of the mean. In these situations the usual R and S charts can also not be used.

The proposed remedy depends on the nature of the relationship between the mean and the variance of the process. One common relationship that we will look at is when the mean and standard deviation is directly proportional so that the standardised mean ($\delta = \frac{\mu}{\sigma}$) is a constant. According to Kang et al. (2007) this is often the case in medical research.

Scientists at the Clinical Research Organization, Quintiles, also confirmed that the standardised mean or coefficient of variation of drug concentrations is constant or approximately constant. By using frequentist methods, Kang et al. (2007), developed a Shewart control chart, equivalent to the S chart, for monitoring the coefficient of variation using rational groups of observations. The chart is a time-ordered plot of the coefficient of variation for successive samples. It contains three lines:

- A center line;
- The upper control limit (UCL);
- The lower control limit (LCL).

By using the posterior predictive distribution in this paper, a Bayesian procedure will be developed to obtain control limits for a future sample standardised mean. These limits will be compared to the classical limits obtained by Kang et al. (2007).

Bayarri and García-Donato (2005) give the following reasons for recommending a Bayesian analysis:

- Control charts are based on future observations and Bayesian methods are very natural for prediction.
- Uncertainty in the estimation of the unknown parameters is adequately handled.

- Implementation with complicated models and in a sequential scenario poses no methodological difficulty, the numerical difficulties are easily handled via Monte Carlo methods;
- Objective Bayesian analysis is possible without introduction of external information other than the model, but any kind of prior information can be incorporated into the analysis, if desired.

## 2 Frequentist Methods

Assume that $X_j$ $(j = 1, 2, \dots, n)$ are independently, identically normally distributed with mean $\mu$ and variance $\sigma^2$. $\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$ is the sample mean and $S^2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(X_j - \bar{X}\right)^2$ is the sample variance. The sample coefficient of variation is defined as

$$W = \frac{S}{\bar{X}}$$

and the sample standardised mean as

$$W^{-1} = \frac{\bar{X}}{S}.$$

Kang et al. (2007) suggested a control chart for the sample coefficient of variation, similar to that of the $\bar{X}$, R and S charts. By deriving a canonical form for the distribution of the coefficient of variation they obtained control limits for a selection of values of $n$ and $\gamma = \frac{\sigma}{\mu}$. The probability of exceeding these limits is $\frac{1}{740}$ on each side when the process is in control.

In this paper the emphasis will rather be on the inverse of the coefficient of variation, i.e., the standardised mean. From a statistical point of view it is easier to handle the standardised mean than the coefficient of variation.

It is well known that

$$T = \frac{\sqrt{n}\bar{X}}{S} = \sqrt{n}W^{-1}$$

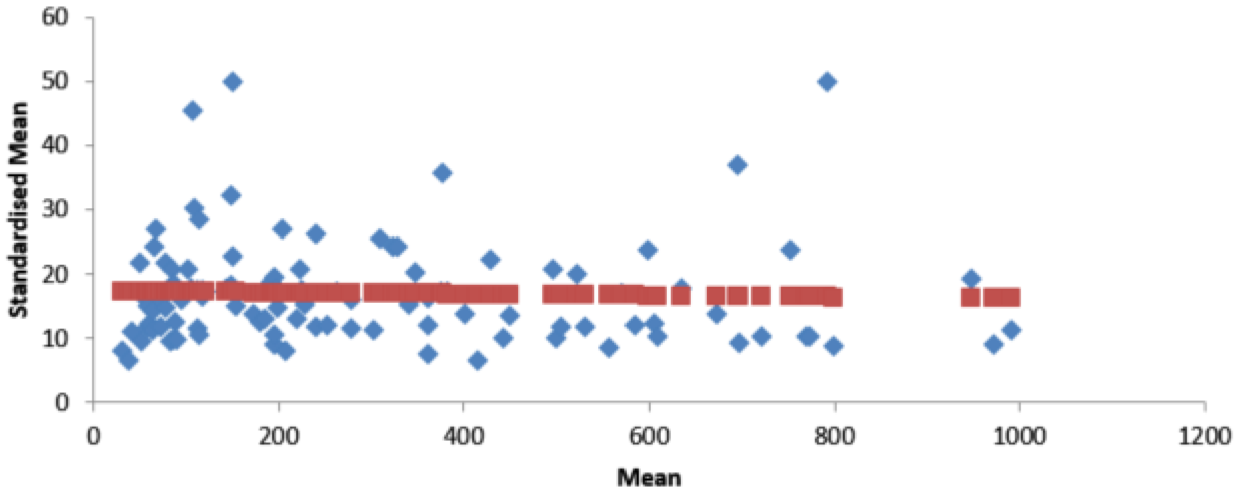
follows a non-central t distribution with $(n-1)$ degrees of freedom and non-centrality parameter $\sqrt{n}\delta$. Inferences about a future standardised mean can therefore be made if $\delta$ is known.

## The Data

The example used by Kang et al. (2007) was that of patients undergoing organ transplantation, for which Cyclosporine is administered. For patients undergoing immunosuppressive treatment, it is vital to control the amount of drug circulating in the body. For this reason frequent blood assays were taken to find the best drug stabilizing level for each patient. The dataset consist of $m = 105$ patients and the number of assays obtained for each patient is $n = 5$. By doing a regression test they confirmed that there is no evidence that the coefficient of variation depends on the mean which means that the assumption of a constant coefficient of variation is appropriate. They used the weighted root mean square estimator $\hat{\gamma} = \sqrt{\frac{1}{m}\sum_{i=1}^{m} w_i^2} = \sqrt{\frac{0.593515}{105}} = 0.0752$ to pool the samples for estimating $\gamma$. $w_i = \frac{s_i}{\bar{x}_i}$ is the sample coefficient of variation for the $i$th patient. $\bar{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij}$ and $s_i^2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(x_{ij} - \bar{x}_i\right)^2$ where $x_{ij}$ is the $j$th blood assay for the $i$th patient. By substituting $\hat{\gamma}$ as an estimate for the unknown $\gamma$ in the distribution of $W$ and by calculating the lower and upper $\frac{1}{740}$ percentage points, they obtained a LCL=0.01218 and UCL=0.15957. The chart was then applied to a fresh data set of 35 observations from a different laboratory. The data used by Kang et al. (2007) is given in Appendix A.

As mentioned, in this paper the emphasis will be on the standardised mean $\delta = \frac{\mu}{\sigma}$. By using the predictive distribution, a Bayesian procedure will be developed to obtain control limits for a future sample standardised mean. Assuming that the process remains stable, the predictive distribution can be used to derive the distribution of the "run-length" and "average run-length". In the last section of this paper a simulation study will be conducted to evaluate the accuracy of our Bayesian procedure. The data that will be used are those of Kang et al. (2007). A plot of the sample standardised means against the sample means are given in Figure 1.

Figure 1: Scatter Plot of Sample Standardised Mean Versus Sample Mean




From Figure 1 and the least squares regression line it is clear that a common standardised mean assumption is appropriate for the Phase I Cyclosporine data. The analysis of variance test in Table 1 confirms that there is no evidence that the standardised mean depends on the mean. A common standardised mean control chart is therefore justified for ongoing control.

Table 1: Regression Test of Dependence of Standardised Mean on Mean

| Source | SS | df | MSS | F | P-value |
|---|---|---|---|---|---|
| Regression | 8.7093 | 1 | 8.7093 | 0.1272 | 0.722 |
| Error | 7052.7 | 103 | 68.4731 | | |
| Total | 7061.4 | 104 | 67.8984 | | |

Note that the p-value of 0.722 is larger than the p-value of 0.245 calculated by Kang et al. (2007) for the coefficient of variation which is an indication that the standardised mean is more appropriate to use than a common coefficient of variation.

## 3 Bayesian Procedure

The non-central t distribution can be used to make inferences about a future standardised mean if $\delta$ is known. In practice $\delta$ is usually unknown.

By assigning a prior distribution to the unknown parameters the uncertainty in the estimation of the unknown parameters can adequately be handled. The information contained in the prior is combined with the likelihood to obtain the posterior distribution of $\delta$. By using the posterior distribution the predictive distribution of a future standardised mean can be obtained. The predictive distribution on the other hand can be used to obtain control limits and to determine the distribution of the "run length". Determination of reasonable non-informative priors is however not an easy task. Therefore, in the next section, reference and probability matching priors will be derived for a common standardised mean across the range of sample values.

# 4 Reference and Probability-Matching Priors for a Common Standardised Mean

As mentioned the Bayesian paradigm emerges as attractive in many types of statistical problems, also in the case of the standardised mean.

Prior distributions are needed to complete the Bayesian specification of the model. Determination of reasonable non-informative priors in multi-parameter problems is not easy; common non-informative priors, such as the Jeffreys' prior can have features that have an unexpectedly dramatic effect on the posterior.

Reference and probability-matching priors often lead to procedures with good frequency properties while returning to the Bayesian flavor. The fact that the resulting Bayesian posterior intervals of the level $1-\alpha$ are also good frequentist intervals at the same level is a very desirable situation.

See also Bayarri and Berger (2004) and Severine, Mukerjee, and Ghosh (2002) for a general discussion.

## 4.1 The Reference Prior

In this section the reference prior of Berger and Bernardo (1992) will be derived for a common standardised mean, $\delta$, across the range of sample values. In general, the derivation depends on the ordering of the parameters and how the parameter vector is divided into sub-vectors. As mentioned by Pearn and Wu (2005) the reference prior maximizes the difference in information (entropy) about the parameter provided by the prior and posterior. In other words, the reference prior is derived in such a way that it provides as little information possible about the parameter of interest. The reference prior algorithm is relatively complicated and, as mentioned, the solution depends on the ordering of the parameters and how the parameter vector is partitioned into sub-vectors. In spite of these difficulties, there is growing evidence, mainly through examples that reference priors provide "sensible" answers from a Bayesian point of view and that frequentist properties of inference from reference posteriors are asymptotically "good". As in the case of the Jeffreys' prior, the reference prior is obtained from the Fisher information matrix. In the case of a scalar parameter, the reference prior is the Jeffreys' prior.

Bernardo (1998) derived the reference prior for the standardised mean in the case of a single sample. From the medical example given in Kang et al. (2007) it is clear that the standard deviation of measurements is approximately proportional to the mean; that is, the standardised mean is constant across the range of means, which is an indication that the a reference prior for a common standardised mean should be derived.

**Theorem 1.** *Let $X_{ij} \sim N\left(\mu_i, \sigma_i^2\right)$ where $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$ and the standardised mean is $\delta = \frac{\mu_1}{\sigma_1} = \frac{\mu_2}{\sigma_2} = \cdots \frac{\mu_m}{\sigma_m}$. The reference prior for the ordering $\left\{\delta, \left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2\right)\right\}$ is given by*

$$p_R\left(\delta, \sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2\right) \propto \left(1 + \frac{1}{2}\delta^2\right)^{-\frac{1}{2}} \prod_{i=1}^{m} \sigma_i^{-2} \tag{1}$$

*Proof.* The proof is given in Appendix B. □

*Note*: The ordering $\left\{\delta, \left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2\right)\right\}$ means that the standardised mean is the most important parameter while the $m$ variance components are of equal importance, but not as important as $\delta$. Also if $m = 1$, Equation 1 simplifies to the reference prior obtained by Bernardo (1998).

## 4.2 Probability-Matching Priors

The reference prior algorithm is but one way to obtain a useful non-informative prior. Another type of non-informative prior is the probability-matching prior. This prior has good frequentist properties. Two reasons for using probability-matching priors are that they provide a method for constructing accurate frequentist intervals, and that they could be potentially useful for comparative purposes in a Bayesian analysis.

There are two methods for generating probability-matching priors due to Tibshirani (1989) and Datta and Ghosh (1995).

Tibshirani (1989) generated probability-matching priors by transforming the model parameters so that the parameter of interest is orthogonal to the other parameters. The prior distribution is then taken to be proportional to the square root of the upper left element of the information matrix in the new parametrization.

Datta and Ghosh (1995) provided a different solution to the problem of finding probability-matching priors. They derived the differential equation that a prior must satisfy if the posterior probability of a one-sided credibility interval for a parametric function and its frequentist probability agree up to $O\left(n^{-1}\right)$ where $n$ is the sample size.

According to Datta and Ghosh (1995) $p\left(\underline{\theta}\right)$ is a probability-matching prior for $\underline{\theta}=\left[\delta,\sigma_1^2,\sigma_2^2,\ldots,\sigma_m^2\right]'$ the vector of unknown parameters, if the following differential equation is satisfied:

$$\sum_{\alpha=1}^{m+1}\frac{\partial}{\partial\theta_\alpha}\left\{\Upsilon_\alpha\left(\underline{\theta}\right)p\left(\underline{\theta}\right)\right\}=0$$

where

$$\Upsilon\left(\underline{\theta}\right)=\frac{F^{-1}\left(\underline{\theta}\right)\nabla_t\left(\underline{\theta}\right)}{\sqrt{\nabla_t'\left(\underline{\theta}\right)F^{-1}\left(\underline{\theta}\right)\nabla_t\left(\underline{\theta}\right)}}=\left[\begin{array}{cccc}\Upsilon_1\left(\underline{\theta}\right) & \Upsilon_2\left(\underline{\theta}\right) & \cdots & \Upsilon_{m+1}\left(\underline{\theta}\right)\end{array}\right]'$$

and

$$\nabla_t\left(\underline{\theta}\right)=\left[\begin{array}{cccc}\frac{\partial}{\partial\theta_1}t\left(\underline{\theta}\right) & \frac{\partial}{\partial\theta_2}t\left(\underline{\theta}\right) & \cdots & \frac{\partial}{\partial\theta_{m+1}}t\left(\underline{\theta}\right)\end{array}\right]'.$$

$t\left(\underline{\theta}\right)$ is a function of $\underline{\theta}$ and $F^{-1}\left(\underline{\theta}\right)$ is the inverse of the Fisher information matrix.

**Theorem 2.** *The probability-matching prior for the standardised mean, $\delta$, and the variance components is given by*

$$p_M\left(\delta,\sigma_1^2,\sigma_2^2,\ldots,\sigma_m^2\right)\propto\left(1+\frac{1}{2}\delta^2\right)^{-\frac{1}{2}}\prod_{i=1}^{m}\sigma_i^{-2}$$

*Proof.* The proof is provided in Appendix C. □

From Theorem 1 and Theorem 2 it can be seen that the reference and probability-matching priors are equal and the Bayesian analysis using either of these priors will yield exactly the same results.

Note that the reference (probability-matching) prior in terms of $\delta$ and the standard deviations, $\sigma_1,\sigma_2,\ldots,\sigma_m$ is

$$p\left(\delta,\sigma_1,\sigma_2,\ldots,\sigma_m\right)=\left(1+\frac{1}{2}\delta^2\right)^{-\frac{1}{2}}\prod_{i=1}^{m}\sigma_i^{-1}.$$

### 4.3 The Joint Posterior Distribution

By combining the prior distribution with the likelihood function the joint posterior distribution of $\delta, \sigma_1, \sigma_2, \ldots, \sigma_m$ can be obtained:

$$p\left(\delta, \sigma_1, \sigma_2, \ldots, \sigma_m | data\right) \propto \left(1 + \frac{1}{2}\delta^2\right)^{-\frac{1}{2}} \prod_{i=1}^{m} \sigma_i^{-(n+1)} \exp\left\{-\frac{1}{2\sigma_i^2}\left[n\left(\bar{x}_i - \sigma_i\delta\right)^2 + (n-1)\, s_i^2\right]\right\} \tag{2}$$

where $\bar{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij}$ and $(n-1)\, s_i^2 = \sum_{j=1}^{n}\left(x_{ij} - \bar{x}_i\right)^2$.

In Theorem 3 it will be proved that the joint posterior distribution is proper and can be used for inferences.

**Theorem 3.** *The posterior distribution is $p\left(\delta, \sigma_1, \sigma_2, \ldots, \sigma_m | data\right)$ is a proper posterior distribution.*

*Proof.* The proof is given in Appendix D. □

The conditional posterior distributions follow easily from the joint posterior distribution:

$$p\left(\delta | \sigma_1, \sigma_2 \ldots, \sigma_m, data\right) \propto \left(1 + \frac{1}{2}\delta^2\right)^{-\frac{1}{2}} \exp\left\{-\frac{n}{2}\sum_{i=1}^{m}\frac{1}{\sigma_1^2}\left(\bar{x}_i - \sigma_i\delta\right)^2\right\} \tag{3}$$

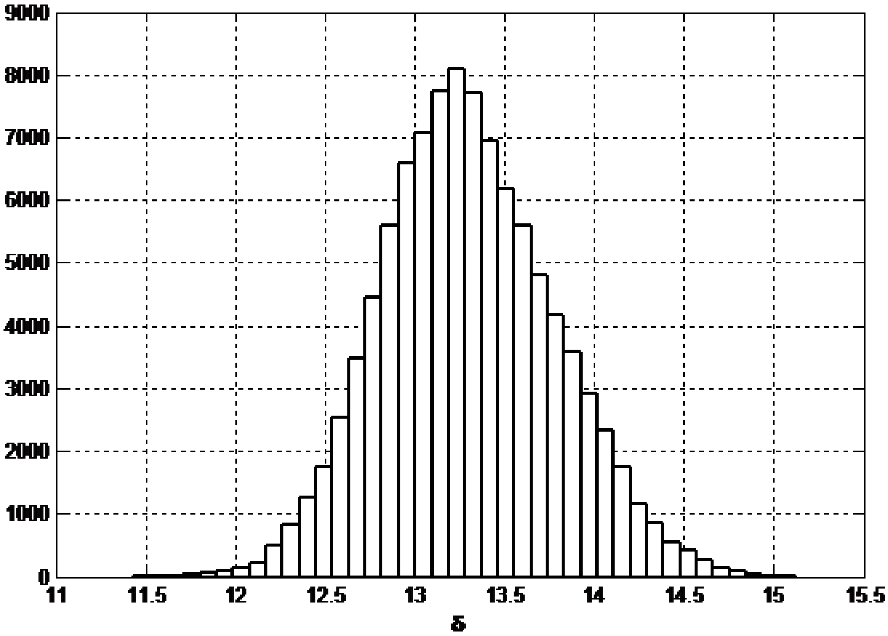
and

$$p\left(\sigma_1, \sigma_2, \ldots, \sigma_m | \delta, data\right) \propto \prod_{i=1}^{m}\left(\sigma_i^{-(n+1)} \exp\left\{-\frac{1}{2\sigma_i^2}\left[n\left(\bar{x}_i - \sigma_i\delta\right)^2 + (n-1)\, s_i^2\right]\right\}\right). \tag{4}$$

By using the conditional posterior distributions (Equation 3 and Equation 4) and Gibbs sampling the unconditional posterior distributions can be obtained.

In Figure 2 the unconditional posterior distribution of $\delta$ (the standardised mean), $p\left(\delta | data\right)$ from the medical data is illustrated ($m = 105$ and $n = 5$).

Figure 2: Histogram of the Posterior Distribution of $\delta = \frac{\mu}{\sigma}$



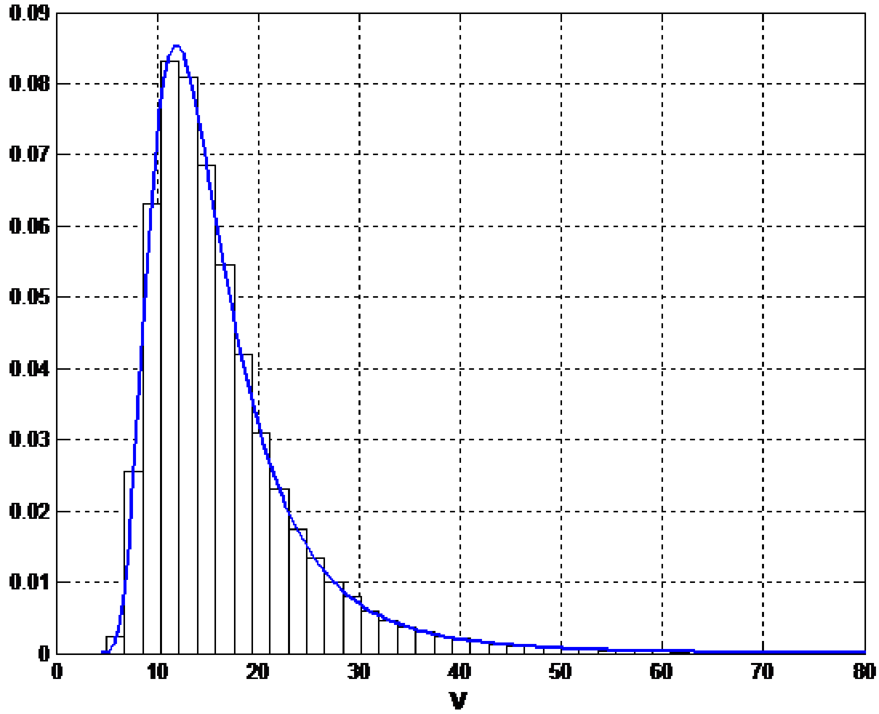

$$mean\,(\delta) = 13.2984$$

$\{mean\,(\delta)\}^{-1} = (13.2984)^{-1} = 0.075197$ can therefore be used as an estimate for the common coefficient of variation, $\gamma$. As mentioned in Section 2, Kang et al. (2007) used the weighted root mean square estimator $\hat{\gamma} = \sqrt{\frac{1}{m}\sum_{i=1}^{m} w_i^2} = \sqrt{\frac{0.593515}{105}} = 0.0752$ to pool the samples for estimating $\gamma$. $w_i = \frac{s_i}{\bar{x}_i}$ is the sample coefficient of variation. It is interesting to note that $\{mean\,(\delta)\}^{-1}$ is for all practical purposes the same as $\hat{\gamma}$.

Since $T = \sqrt{n}W^{-1} = \sqrt{n}V$ follows a non-central t distribution with $(n-1)$ degrees of freedom and a non-centrality parameter $\sqrt{n}\delta$, $f\,(V|\delta)$ can be obtained. By substituting each of the simulated $\delta$ values of the posterior distribution of $\delta$ in $f\,(V|\delta)$ and using the Rao-Blackwell procedure (averaging the conditional distributions), the unconditional posterior predictive density of $f\,(V|data)$ (a future standardised mean) can be obtained. This is illustrated by the smooth curve in Figure 3.

The histogram in Figure 3 is obtained in the following way. Define a future sample mean as $\bar{X}_f$ and a future sample standard deviation as $S_f$. Since $V = \frac{\bar{X}_f}{S_f} = \frac{\tilde{Z}}{\sqrt{\frac{\chi^2_{n-1}}{n-1}}}$ where $\tilde{Z} \sim N\left(\delta, \frac{1}{n}\right)$, the histogram of the distribution of $V$ is obtained by simulating $\delta$ from the its posterior distribution and then $\tilde{Z} \sim N\left(\delta, \frac{1}{n}\right)$. Simulate now a $\chi^2_{n-1}$ random variable and calculate $V$ and repeat the process a large number of times. It is clear that the two distributions are the same.

Figure 3: Predictive Density $f(V|data)$ for $n = 5$




$$mean(V) = 16.6671$$
$$99.73\% \text{ Equal-tail Interval} = (6.212; 83.365)$$
$$99.73\% \text{ HPD Interval} = (5.176; 67.777)$$

According to this the 99.73% equal-tail interval for a future sample coefficient of variation is $\left[(83.365)^{-1}; (6.212)^{-1}\right] = [0.011995; 0.1609787]$. For a 99.73% equal-tail control chart for the coefficient of variation, Kang et al. (2007) calculated the lower control limit as 0.1218, the upper control limit as 0.15957 and as central line they used the root-mean square value $\hat{\gamma} = 0.075$. The frequentist limits calculated by them are for all practical purposes the same as our Bayesian control limits.

Kang et al. (2007) then applied their control chart to a new dataset of 35 patients from a different laboratory. Eight of the patients' coefficient of variation (based on five observations) lie outside the control limits. Since the 99.73% equal-tail prediction interval is effectively identical to the control limits of Kang et al. (2007) our conclusions are the same.

As mentioned the rejection region of size $\alpha$ ($\alpha = 0.0027$) for the predictive distribution is

$$\alpha = \int_{R(\alpha)} f(V|data)\, dV.$$

In the case of the equal-tail interval, $R(\alpha)$ represents those values of $V$ that are smaller than 6.212 or larger than 83.365.

It is therefore clear that statistical process control is actually implemented in two phases. In Phase I the primary interest is to assess process stability. The practitioner must therefore be sure that the process is in statistical control before control limits can be determined for online monitoring in Phase II.

Assuming that the process remains stable, the predictive distribution can be used to derive the distribution of the "run length" and "average run length". The "run length" is defined as the number of future standardised means, $r$ until the control chart signals for the first time (Note that $r$ does not include that standard mean when the control chart signals). Given $\delta$ and a stable Phase I process, the distribution of the run length $r$ is geometric with parameter

$$\Psi(\delta) = \int_{R(\alpha)} f(V|\delta)\, dV$$

where $f(V|\delta)$ is the distribution of a future standardised mean ($\sqrt{n}V$ is a non-central t distribution with $(n-1)$ degrees of freedom and a non-centrality parameter $\sqrt{n}\delta$). The value of $\delta$ is of course unknown and its uncertainty is described by its posterior distribution. The predictive distribution of the "run length" or the "average run length" can therefore easily be obtained.

The mean and variance of $r$ given $\delta$ are given by

$$E(r|\delta) = \frac{1-\psi(\delta)}{\psi(\delta)}$$

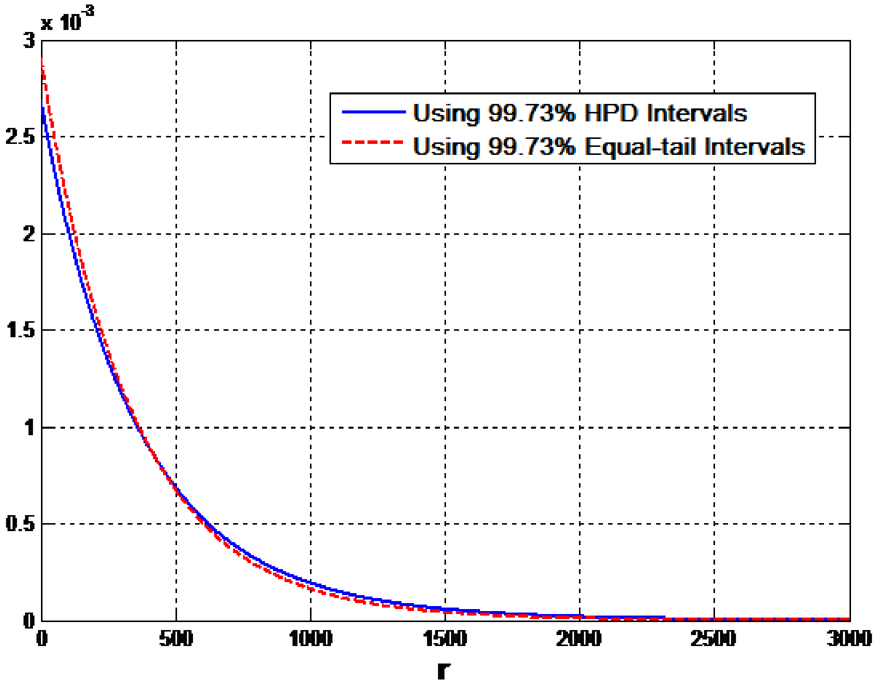
and

$$Var(r|\delta) = \frac{1-\psi(\delta)}{\psi^2(\delta)}.$$

The unconditional moments $E(r|data)$, $E(r^2|data)$ and $Var(r|data)$ can therefore easily be obtained by simulation or numerical integration. For further details see Menzefricke (2002, 2007, 2010a,b).

In Figure 4 the predictive distributions of the "run-length" is displayed for the 99.73% equal tail interval as well as for the 99.73% HPD interval. As mentioned for given $\delta$, the "run-length", $r$, is geometric with parameter $\psi(\delta)$. The unconditional "run-length" given in Figure 4 are therefore obtained by the Rao-Blackwell method, i.e., the average of a large number of conditional "run-lengths".

Figure 4: Predictive Density of Run Length $f(r|data)$with $n = 5$

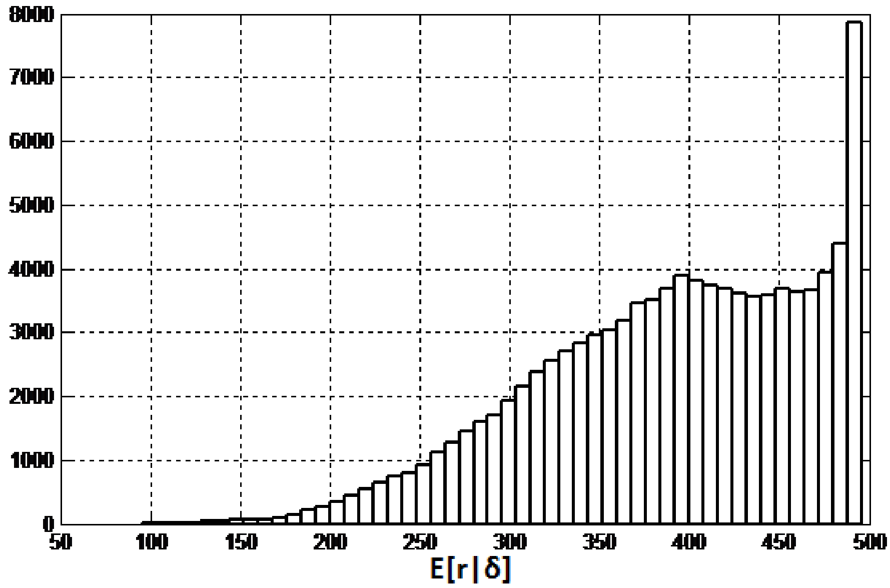



99.73% Equal-tail Interval:
$E(r|data) = 385.943$, $Median(r|data) = 261.27$
99.73% HPD Interval:
$E(r|data) = 347.625$, $Median(r|data) = 238.50$

From the figure it can be seen that the expected "run-length" $E(r|data) = 385.943$ for the 99.73% equal-tail interval is somewhat larger than the ARL of 370 given by Kang et al. (2007). The median "run-length", $Median(r|data) = 261.27$ on the other hand is smaller than the mean "run-length". This is clear from the skewness of the distribution. From Figure 4 it is also clear that the "run-length" do not differ much for equal-tail and HPD intervals.

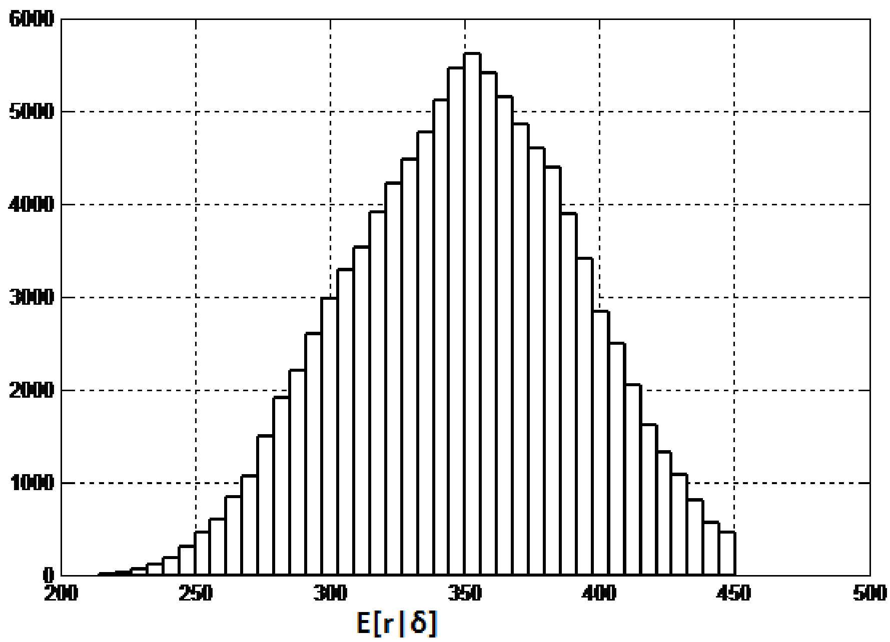
Figure 5 illustrates the distribution of $E(r|data)$ for each simulated value of $\delta$, i.e., the distribution of the expected "run-length" in the case of the 99.73% equal-tail interval while Figure 6 display the distribution of the expected "run-length" for the 99.73% HPD interval.

Figure 5: Distribution of the Expected Run-Length, Equal-tail Interval, $n = 5$




$Mean = 389.63$ and $Median = 398.21$

Figure 6: Distribution of the Expected Run-Length, HPD Interval, $n = 5$




$Mean = 349.505$ and $Median = 350.711$

As it should be the corresponding mean "run-lengths" in Figure 4 are for all practical purposes the same as the mean "run-lengths" in Figure 5 and Figure 6.

Further descriptive statistics for Figure 2 to Figure 6 are given in Appendix E.

# 5 Simulation Study

In this section a simulation study will be conducted to observe if the 95% Bayesian confidence intervals for $\delta$ have the correct frequentist coverage.

For the simulation study the following combinations of parameters will be used:

| $\mu_i$ | 10 | 20 | 30 | 40 | 50 | 60 | $\cdots$ | 1000 | 1010 | 1020 | $\cdots$ | 1050 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma_i$ | 0.75 | 1.5 | 2.25 | 3.0 | 3.75 | | $\cdots$ | 75 | | | $\cdots$ | 78.15 |

which means that $\delta = \frac{\mu_i}{\sigma_i} = 13.3333$, $\gamma = 0.075$, $i = 1, 2, \ldots, m$ and $m = 105$. These parameter combinations are representative of the parameter values of the medical dataset on patients undergoing organ transplantation analysed by Kang et al. (2007). As mentioned, the dataset consist of $m = 105$ patients and the number assays obtained for each patient is $n = 5$. As a common estimate for $\gamma$, the weighted mean square $\hat{\gamma} = 0.075$ was used.

For the above given parameter combination a dataset can be simulated consisting of $m$ samples and $n = 5$ observations per sample. However since we are only interested in the sufficient statistics $\bar{X}_i$ and $S_i$ these can be simulated directly, namely $\bar{X}_i \sim N\left(\mu_i, \frac{\sigma_i^2}{n}\right)$ and $S_i^2 \sim \frac{\sigma_i^2 \chi_{n-1}^2}{n-1}$.

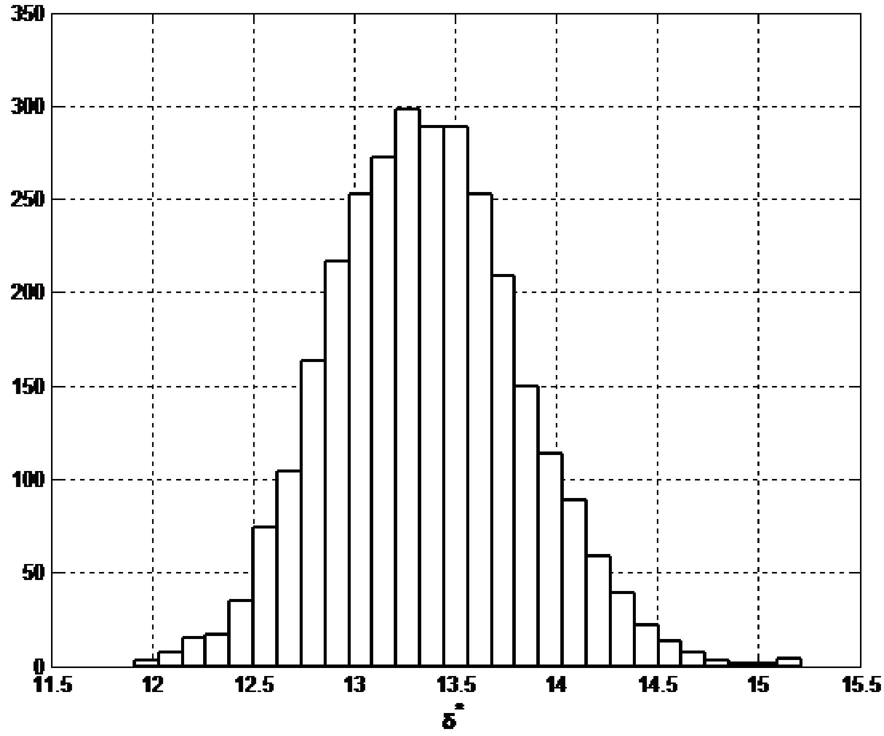
The simulated $\bar{X}_i$ and $S_i^2$ $(i = 1, 2, \ldots, m)$ values are then substituted int he conditional posterior distributions given in Equation 3 and Equation 4. By using the conditional posterior distributions and Gibbs sampling the unconditional posterior distribution $p(\delta|data)$ can be obtained. A confidence interval for $\delta$ will be calculated as follows: Simulate $l = 10,000$ values of $\delta$ and sort them in ascending order $\tilde{\delta}_{(1)} \leq \tilde{\delta}_{(2)} \leq \cdots \leq \tilde{\delta}_{(l)}$.

Let $K_1 = \left[\frac{\alpha}{2} l\right]$ and $K_2 = \left[\left(1 - \frac{\alpha}{2}\right) l\right]$ where $[a]$ denotes the largest integer not greater than $a$. $\left\{\tilde{\delta}_{(K_1)}, \tilde{\delta}_{(K_2)}\right\}$ is then a $100(1-\alpha)\%$ Bayesian confidence interval for $\delta$. By repeating the procedure for $R = 3,000$ datasets it is found that the 3,000, 95% Bayesian confidence intervals ($\alpha = 0.05$) cover the true parameter value $\delta = 13.333$ in 2,841 cases.

An estimate of the frequentist probability of coverage is therefore $P\left\{\tilde{\delta}_{(K_1)} \leq \delta \leq \tilde{\delta}_{(K_2)}\right\} = 0.9470$. Also, $P\left\{\delta \leq \tilde{\delta}_{(K_1)}\right\} = 0.0313$ and $P\left\{\delta \geq \tilde{\delta}_{(K_2)}\right\} = 0.0217$.

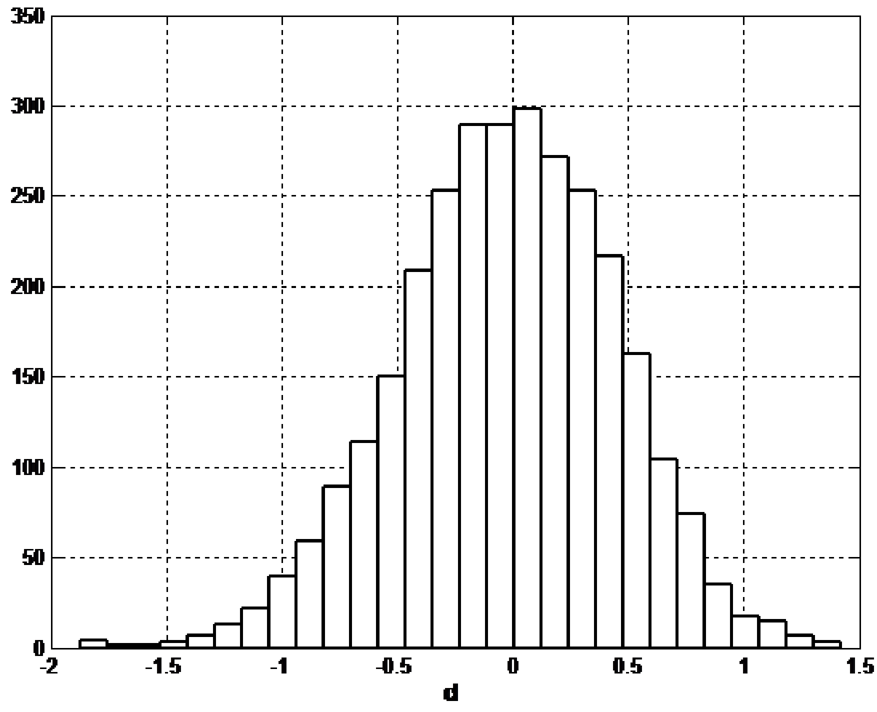
For each dataset the posterior mean, $\delta^*$ of the $l = 10,000$ simulated $\delta$ values is calculated as well as $d = 13.3333 - \delta^*$, the difference between the posterior mean and the true parameter value. The histograms of the $R = 3,000$ $\delta^*$ and $d$ values are displayed in Figure 7 and Figure 8. The histogram of the posterior means $\delta^*$ (Figure 7) is for all practical purposes symmetrical. For the $d$ values (Figure 8), the histogram is slightly skew to the left.

Figure 7: Histogram of Posterior Means $\delta^*$




$Mean\,(\delta^*) = 13.356$; $Median\,(\delta^*) = 13.3439$; $Var\,(\delta^*) = 0.2177$; 95% Interval $= (12.492; 14.298)$

Figure 8: Histogram of $d = \delta - \delta^*$




$Mean\,(d) = -0.0227$; $Median\,(d) = -0.0106$; $Var\,(d) = 0.2177$; 95% Interval $= (-0.968; 0.838)$

# 6 Conclusion

This paper develops a Bayesian control chart for monitoring a common standardised mean across a range of sample values. In the Bayesian approach prior knowledge about the unknown parameters is formally incorporated into the process of inference by assigning a prior distribution to the parameters. The information contained in the prior is combined with the likelihood function to obtain the posterior distribution. By using the posterior distribution the predictive distribution of a future standardised mean can be obtained.

Determination of reasonable non-informative priors in multi-parameter problems is not an easy task. The Jeffreys' prior for example can have a bad effect on the posterior distribution. Reference and probability matching priors are therefore derived for a constant standardised mean across a range of sample values. The theory and results are applied to a real problem of patients undergoing organ transplantation for which Cyclosporine is administered. This problem is discussed in detail by Kang et al. (2007). The 99.73% equal tail prediction interval of a future coefficient of variation (inverse of the standardised mean) is effectively identical to the lower and upper control chart limits calculated by Kang et al. (2007). A simulation study shows that the 95% Bayesian confidence intervals for $\delta$ has the correct frequentist coverage.

The example illustrates the flexibility and unique features of the Bayesian simulation method for obtaining posterior distributions, prediction intervals and run lengths.

# References


Bayarri, M., Berger, J., 2004. The interplay of bayesian and frequentist analysis. Statistical Science 19 (1), 58–80.

Bayarri, M., García-Donato, G., 2005. A bayesian sequential look at u-control charts. Technometrics 47(2), 142–151.

Berger, J., Bernardo, J., 1992. On the development of reference priors. Bayesian Statistics 4, 35–60.

Berger, J. O., Liseo, B., Wolpert, R. L., 1999. Integrated likelihood methods for eliminating nuisance parameters. Statistical Science 14, 1–28.

Bernardo, J., 1998. Bayesian Reference Analysis - A Postgraduate Tutorial Course. Departament d'Estadistica i l. O, Facultat de Matematiques 46100-Burjasson, Valenicia, Spain.

Datta, G., Ghosh, J., 1995. On priors providing frequentist validity for bayesian inference. Biometrika 82 (1), 37–45.

Kang, C. W., Lee, M. S., Seong, Y. J., Hawkins, D. M., 2007. A control chart for the coefficient of variation. Journal of Quality Technology 39 (2), 151–158.

Menzefricke, U., 2002. On the evaluation of control chart limits based on predictive distributions. Communications in Statistics - Theory and Methods 31(8), 1423–1440.

Menzefricke, U., 2007. Control chart for the generalized variance based on its predictive distribution. Communications in Statistics - Theory and Methods 36(5), 1031–1038.

Menzefricke, U., 2010a. Control chart for the variance and the coefficient of variation based on their predictive distribution. Communications in Statistics - Theory and Methods 39(16), 2930–2941.

Menzefricke, U., 2010b. Multivariate exponentially weighted moving average chart for a mean based on its predictive distribution. Communications in Statistics - Theory and Methods 39(16), 2942–2960.

Pearn, W., Wu, C., 2005. A bayesian approach for assessing process precision based on multiple samples. European Journal of Operational Research 165 (3), 685–695.

Severine, T., Mukerjee, R., Ghosh, M., 2002. On an exact probability matching property of right-invariant priors. Biometrika 89 (4), 952–957.

Tibshirani, R., 1989. Noninformative priors for one parameter of many. Biometrika 76 (3), 604–608.

# Appendices

## A Data for Medical Example

| $m$ | $\overline{X}$ | $w_i \times 100$ | $m$ | $\overline{X}$ | $w_i \times 100$ | $m$ | $\overline{X}$ | $w_i \times 100$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 31.7 | 12.4 | 36 | 120.3 | 5.8 | 71 | 361.4 | 8.3 |
| 2 | 37.7 | 15.3 | 37 | 143.7 | 5.6 | 72 | 361.5 | 13.4 |
| 3 | 40.6 | 9.1 | 38 | 148.6 | 5.5 | 73 | 361.8 | 6.1 |
| 4 | 50.5 | 4.6 | 39 | 149.1 | 3.1 | 74 | 374.6 | 5.8 |
| 5 | 52 | 10.5 | 40 | 149.9 | 2 | 75 | 376.3 | 2.8 |
| 6 | 57.6 | 6.2 | 41 | 151 | 4.4 | 76 | 382.3 | 5.8 |
| 7 | 58.3 | 6.6 | 42 | 153.6 | 6.6 | 77 | 401.7 | 7.3 |
| 8 | 58.9 | 8.4 | 43 | 172.2 | 7.2 | 78 | 415.2 | 15.1 |
| 9 | 61.2 | 8.1 | 44 | 179.8 | 7.9 | 79 | 428.8 | 4.5 |
| 10 | 64.3 | 7 | 45 | 185.3 | 7.6 | 80 | 442.1 | 9.9 |
| 11 | 64.5 | 8.8 | 46 | 192.1 | 5.3 | 81 | 450.1 | 7.4 |
| 12 | 65.6 | 4.1 | 47 | 193.8 | 5.9 | 82 | 496.5 | 4.8 |
| 13 | 68 | 3.7 | 48 | 195.1 | 11 | 83 | 499.7 | 10 |
| 14 | 71.8 | 6.2 | 49 | 195.2 | 5.1 | 84 | 504.6 | 8.4 |
| 15 | 72.1 | 8.4 | 50 | 195.4 | 9.4 | 85 | 523.1 | 5 |
| 16 | 78.4 | 6.8 | 51 | 196.4 | 5.6 | 86 | 531.7 | 8.5 |
| 17 | 78.4 | 4.6 | 52 | 199.6 | 6.8 | 87 | 556.4 | 11.8 |
| 18 | 79.5 | 5.7 | 53 | 204.4 | 3.7 | 88 | 571.4 | 5.9 |
| 19 | 83.2 | 10.5 | 54 | 207.8 | 12.4 | 89 | 584.1 | 8.3 |
| 20 | 85.1 | 4.8 | 55 | 219 | 7.6 | 90 | 597.6 | 4.2 |
| 21 | 85.6 | 5.4 | 56 | 222.9 | 4.8 | 91 | 606.2 | 8.2 |
| 22 | 86 | 10.1 | 57 | 225.1 | 5.7 | 92 | 609 | 9.7 |
| 23 | 87.3 | 7.9 | 58 | 227.6 | 6.5 | 93 | 635.4 | 5.6 |
| 24 | 89.1 | 10.3 | 59 | 240.5 | 3.8 | 94 | 672.2 | 7.2 |
| 25 | 95.4 | 6.2 | 60 | 241.1 | 8.4 | 95 | 695.9 | 2.7 |
| 26 | 101.9 | 4.8 | 61 | 252.2 | 8.3 | 96 | 696.4 | 10.6 |
| 27 | 105.4 | 5.6 | 62 | 262.2 | 5.8 | 97 | 721.3 | 9.8 |
| 28 | 107.2 | 2.2 | 63 | 277.9 | 8.7 | 98 | 752 | 4.2 |
| 29 | 108.2 | 3.3 | 64 | 278.3 | 6.2 | 99 | 769.5 | 9.7 |
| 30 | 112 | 8.7 | 65 | 303.4 | 8.8 | 100 | 772.7 | 9.6 |
| 31 | 112.3 | 5.7 | 66 | 309.7 | 3.9 | 101 | 791.6 | 2 |
| 32 | 113.5 | 9.4 | 67 | 323.9 | 4.1 | 102 | 799.9 | 11.4 |
| 33 | 114.3 | 3.5 | 68 | 328.7 | 4.1 | 103 | 948.6 | 5.2 |
| 34 | 116.8 | 6 | 69 | 341.2 | 6.5 | 104 | 971.8 | 11.1 |
| 35 | 117.8 | 5.7 | 70 | 347.3 | 4.9 | 105 | 991.2 | 8.8 |

## B Proof of Theorem 1

*Proof.* The likelihood function is given by

$$L\left(\delta, \sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2 | data\right) \propto \prod_{i=1}^{m} \left(\sigma_i^2\right)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma_i^2}\left[n\left(\bar{x}_i - \delta\sigma_i\right)^2 + (n-1)\, S_i^2\right]\right\}$$

where $\bar{x}_i = \frac{1}{n}\sum_{i=1}^{n} x_{ij}$ and $(n-1)\, S_i^2 = \sum_{j=1}^{n}\left(x_{ij} - \bar{x}_i\right)^2$.

By differentiating the log likelihood function $\tilde{l}$ twice with respect to the unknown parameters and taking expected values, the Fisher Information matrix can be obtained.

$$\tilde{l} = \log L\left(\delta, \sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2 | data\right) = -\frac{n}{2}\sum_{i=1}^{m} \log \sigma_i^2 - \frac{1}{2}\sum_{i=1}^{m} \frac{1}{\sigma_i^2}\left[n\left(\bar{x}_i - \delta\sigma_i\right)^2 + (n-1)\, S_i^2\right]$$

and

$$\frac{\partial^2 \tilde{l}}{\left(\partial\sigma_i^2\right)^2} = \frac{n}{2}\left(\frac{1}{\sigma_i^2}\right)^2 - \frac{2n\bar{x}_i^2}{2\left(\sigma_i^2\right)^3} + \frac{3n\bar{x}_i\delta}{4\sigma_i^5} - \frac{(n-1)\, S_i^2}{\left(\sigma_i^2\right)^3}.$$

Therefore

$$-E\left[\frac{\partial^2 \tilde{l}}{\left(\partial\sigma_i^2\right)^2}\right] = \frac{n}{2}\left(\frac{1}{\sigma_i^2}\right)^2 \left\{1 + \frac{1}{2}\delta^2\right\} \text{ where } i = 1, 2, \ldots m.$$

Also

$$-E\left(\frac{\partial^2 \tilde{l}}{\partial\sigma_i^2 \partial\sigma_l^2}\right) = 0.$$

Further

$$\frac{\partial^2 \tilde{l}}{\left(\partial\delta\right)^2} = -\frac{1}{2}\sum_{i=1}^{m} \frac{1}{\sigma_i^2}\left(2n\sigma_i^2\right) = -nm$$

$$\therefore -E\left(\frac{\partial^2 \tilde{l}}{\left(\partial\delta\right)^2}\right) = nm.$$

If we differentiate $\tilde{l}$ with respect to $\delta$ and $\sigma_i^2$ we get

$$\frac{\partial^2 \tilde{l}}{\partial\delta\partial\sigma_i^2} = \frac{-n\bar{x}_i}{2\sigma_i^3}$$

and

$$-E\left(\frac{\partial^2 \tilde{l}}{\partial\delta\partial\sigma_i^2}\right) = \frac{n\delta}{2\sigma_i^2} \;\; i = 1, 2, \ldots, m.$$

The Fisher Information matrix now follows as

$$F\left(\underline{\theta}\right) = F\left(\delta, \sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2\right) = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}$$

where

$$F_{11} = nm,\; F_{12} = F_{21}^{'} = \begin{bmatrix} \frac{n\delta}{2\sigma_1^2} & \frac{n\delta}{2\sigma_2^2} & \cdots & \frac{n\delta}{2\sigma_m^2} \end{bmatrix}$$

and

$$F_{22} = \begin{bmatrix} \frac{n}{2}\left(\frac{1}{\sigma_1^2}\right)^2\left\{1 + \frac{1}{2}\delta^2\right\} & 0 & \cdots & 0 \\ 0 & \frac{n}{2}\left(\frac{1}{\sigma_2^2}\right)^2\left\{1 + \frac{1}{2}\delta^2\right\} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{n}{2}\left(\frac{1}{\sigma_m^2}\right)^2\left\{1 + \frac{1}{2}\delta^2\right\} \end{bmatrix}.$$

To calculate the reference prior for the ordering $\{\delta, (\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2)\}$, $F_{11\cdot 2}$ must first be calculated and then $|F_{22}|$. Now

$$F_{11\cdot 2} = F_{11} - F_{12}F_{22}^{-1}F_{21} = nm - \frac{1}{2}nm\delta^2\left(1+\frac{1}{2}\delta^2\right)^{-1} = \frac{nm}{\left(1+\frac{1}{2}\delta^2\right)} = h_1$$

and

$$p(\delta) \propto h_1^{\frac{1}{2}} \propto \left(1+\frac{1}{2}\delta^2\right)^{-\frac{1}{2}}.$$

Also

$$|F_{22}| = \left\{\frac{n}{2}\left(1+\frac{1}{2}\delta^2\right)\right\}^m \prod_{i=1}^m \left(\frac{1}{\sigma_i^2}\right)^2 = h_2.$$

This means that

$$p\left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2|\delta\right) \propto h_2^{\frac{1}{2}} = \prod_{i=1}^m \left(\frac{1}{\sigma_i^2}\right).$$

Therefore the reference prior for the ordering $\{\delta, (\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2)\}$ is

$$p_R\left(\delta, \sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2\right) = p(\delta)\, p\left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2|data\right) \propto \left(1+\frac{1}{2}\delta^2\right)^{-\frac{1}{2}} \prod_{i=1}^m \sigma_i^{-2}.$$

$\square$

## C Proof of Theorem 2

*Proof.* The inverse of the Fisher Information matrix is given by

$$F^{-1}\left(\underline{\theta}\right)=F^{-1}\left(\delta,\sigma_1^2,\sigma_2^2,\ldots,\sigma_m^2\right)=\begin{bmatrix} F^{11} & F^{12} & F^{13} & \cdots & F^{1,m+1} \\ F^{21} & F^{22} & F^{23} & \cdots & F^{2,m+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ F^{m+1,1} & F^{m+1,2} & F^{m+1,3} & \cdots & F^{m+1,m+1} \end{bmatrix}.$$

Let

$$t\left(\underline{\theta}\right)=t\left(\delta,\sigma_1^2,\sigma_2^2,\ldots,\sigma_m^2\right)=\delta.$$

Since

$$\nabla_t^{'}\left(\underline{\theta}\right)=\begin{bmatrix}\frac{\partial}{\partial\delta}t\left(\underline{\theta}\right) & \frac{\partial}{\partial\sigma_1^2}t\left(\underline{\theta}\right) & \cdots & \frac{\partial}{\partial\sigma_m^2}t\left(\underline{\theta}\right)\end{bmatrix}=\begin{bmatrix}1 & 0 & \cdots & 0\end{bmatrix}$$

we have that

$$\nabla_t^{'}\left(\underline{\theta}\right)F^{-1}\left(\underline{\theta}\right)=\begin{bmatrix}F^{11} & F^{12} & \cdots & F^{1,m+1}\end{bmatrix}=\begin{bmatrix}\frac{2+\delta^2}{2mn} & \frac{-\delta\sigma_1^2}{mn} & \frac{-\delta\sigma_2^2}{mn} & \cdots & \frac{-\delta\sigma_m^2}{mn}\end{bmatrix}$$

and

$$\sqrt{\nabla_t^{'}\left(\underline{\theta}\right)F^{-1}\left(\underline{\theta}\right)\nabla_t\left(\underline{\theta}\right)}=\left(\frac{2+\delta^2}{2mn}\right)^{\frac{1}{2}}.$$

Further

$$\Upsilon^{'}\left(\underline{\theta}\right)=\frac{\nabla_t^{'}\left(\underline{\theta}\right)F^{-1}\left(\underline{\theta}\right)}{\sqrt{\nabla_t^{'}\left(\underline{\theta}\right)F^{-1}\left(\underline{\theta}\right)\nabla_t\left(\underline{\theta}\right)}}=\begin{bmatrix}\Upsilon_1\left(\underline{\theta}\right) & \Upsilon_2\left(\underline{\theta}\right) & \cdots & \Upsilon_m\left(\underline{\theta}\right)\end{bmatrix}$$

where

$$\Upsilon_1\left(\underline{\theta}\right)=\left(\frac{2+\delta^2}{2mn}\right)^{\frac{1}{2}},$$

$$\Upsilon_2\left(\underline{\theta}\right)=\frac{-\sqrt{2}\delta\sigma_1^2}{(mn)^{\frac{1}{2}}\left(2+\delta^2\right)^{\frac{1}{2}}},$$

$$\Upsilon_3\left(\underline{\theta}\right)=\frac{-\sqrt{2}\delta\sigma_2^2}{(mn)^{\frac{1}{2}}\left(2+\delta^2\right)^{\frac{1}{2}}}$$

and

$$\Upsilon_{m+1}\left(\underline{\theta}\right)=\frac{-\sqrt{2}\delta\sigma_m^2}{(mn)^{\frac{1}{2}}\left(2+\delta^2\right)^{\frac{1}{2}}}.$$

The prior

$$p_M\left(\underline{\theta}\right)=p_M\left(\delta,\sigma_1^2,\sigma_2^2,\ldots,\sigma_m^2\right)\propto\frac{1}{\left(2+\delta^2\right)^{\frac{1}{2}}}\prod_{i=1}^{m}\sigma_i^{-2}$$

is therefore a probability matching prior since

$$\frac{\partial}{\partial\delta}\left\{\Upsilon_1\left(\underline{\theta}\right)p_M\left(\underline{\theta}\right)\right\}+\frac{\partial}{\partial\sigma_1^2}\left\{\Upsilon_2\left(\underline{\theta}\right)p_M\left(\underline{\theta}\right)\right\}+\cdots+\frac{\partial}{\partial\sigma_m^2}\left\{\Upsilon_{m+1}\left(\underline{\theta}\right)p_M\left(\underline{\theta}\right)\right\}=0.$$

□

## D Proof of Theorem 3

*Proof.* The joint posterior distribution given in Equation 2 can be written as

$$p\left(\delta,\sigma_1,\sigma_2,\ldots,\sigma_m|data\right)\propto\left(1+\frac{1}{2}\delta^2\right)^{-\frac{1}{2}}\prod_{i=1}^{m}\exp\left[-\frac{n\delta^2}{2}\left(1-\frac{\bar{x}_i^2}{D_i^2}\right)\right]\times\left(\frac{1}{\sigma_i}\right)^{n+1}\exp\left[-\frac{n}{2}D_i^2\left(\frac{1}{\sigma_i}-\frac{\bar{x}_i\delta}{D_i^2}\right)\right]$$

where $\bar{x}_i=\frac{1}{n}\sum_{j=1}^{n}x_{ij}$ and $D_i^2=\frac{1}{n}\sum_{j=1}^{n}x_{ij}^2$.

Therefore

$$p\left(\delta|data\right)=\int_0^\infty\cdots\int_0^\infty p\left(\delta,\sigma_1,\sigma_2,\ldots,\sigma_m|data\right)d\sigma_1 d\sigma_2\ldots d\sigma_m$$

$$\propto\left(1+\frac{1}{2}\delta^2\right)^{-\frac{1}{2}}\exp\left\{-\frac{n\delta^2}{2}\sum_{i=1}^{m}\left(1-\frac{\bar{x}_i^2}{D_i^2}\right)\right\}\prod_{i=1}^{m}\left\{\left(\frac{1}{\sigma_i}\right)^{n+1}\exp\left[-\frac{n}{2}D_i^2\left(\frac{1}{\sigma_i}-\frac{\bar{x}_i\delta}{D_i^2}\right)^2\right]\right\}$$

which is proper. As mentioned by Berger et al. (1999) it is usually the case that the reference and Jeffrey's priors will yield proper posterior distributions. □

# E Descriptive Statistics for Medical Data

| | *Figure* | | | | | |
|---|---|---|---|---|---|---|
| Descriptive Statistics | 1<br>$p(\delta \vert data)$ | 2<br>$f(V \vert data)$ | 3<br>$f(r \vert data)$<br>Equal Tail | 4<br>Exp Run Length<br>Equal Tail | 5<br>$f(r \vert data)$<br>HPD Limits | 6<br>Exp Run Length<br>HPD Limits |
| *Mean* | 13.2984 | 16.6671 | 385.943 | 389.630 | 347.625 | 349.505 |
| *Median* | 13.2733 | 14.276 | 261.27 | 398.210 | 238.500 | 350.711 |
| *Mode* | 13.230 | 11.910 | - | - | - | - |
| *Variance* | 0.2296 | 75.0494 | $1.5538e^5$ | $5.6568e^3$ | $1.2292e^5$ | $1.7614e^3$ |
| *95% Equal Tail* | $(12.4056; 14.2805)$ | $(7.661; 37.540)$ | $(8.41; 1469.42)$ | $(226.052; 494.735)$ | $(7.70; 1298.75)$ | $(267.230; 428.919)$ |
| *95% HPD* | $(12.3796; 14.2500)$ | $(6.830; 32.703)$ | $(0; 1183.63)$ | $(252.514; 495.817)$ | $(0; 1050.56)$ | $(268.621; 430.002)$ |
| *99.73% Equal Tail* | - | $(6.212; 83.365)$ | - | - | - | - |
| *99.73% HPD* | - | $(5.179; 67.777)$ | - | - | - | - |