

Implementation of multivariate control charts in a clinical setting

MARY WATERHOUSE^{1,2}, IAN SMITH³, HASSAN ASSAREH¹ AND KERRIE MENGERSEN¹

¹School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia, ²St Andrew's Medical Institute, Brisbane, Australia, and ³Manager of Physical Sciences, St Andrew's War Memorial Hospital, Brisbane, Australia

Address reprint requests to: Mary Waterhouse, School of Mathematical Sciences, Queensland University of Technology, GPO Box 2434, Brisbane, QLD 4001, Australia. Tel: +61-7-31388313; Fax: +61-7-31382310; E-mail: mary.waterhouse@qut.edu.au, mary.waterhouse@gmail.com

Accepted for publication 18 July 2010

Abstract

Background. In most clinical monitoring cases there is a need to track more than one quality characteristic. If separate univariate charts are used, the overall probability of a false alarm may be inflated since correlation between variables is ignored. In such cases, multivariate control charts should be considered.

Purpose. This paper considers the implementation and performance of the T^2 , multivariate exponentially weighted moving average (MEWMA) and multivariate cumulative sum (MCUSUM) charts in light of the challenges faced in clinical settings. We discuss how to handle incomplete records and non-normality of data, and we provide recommendations on chart selection.

Data sources. Our discussion is supported by a case study involving the monitoring of radiation delivered to patients undergoing diagnostic coronary angiogram procedures at St Andrew's War Memorial Hospital, Australia. We also perform a simulation study to investigate chart performance for various correlation structures, patterns of mean shifts, amounts of missing data and methods of imputation.

Conclusions. The MEWMA and MCUSUM charts detect small to moderate shifts quickly, even when the quality characteristics are uncorrelated. The T^2 chart performs less well overall, although it is useful for rapid detection of large shifts. When records are incomplete, we recommend using multiple imputation.

Keywords: quality control, clinical indicators, multivariate control charts

Introduction

Control charts are becoming widely accepted in the health domain as a means of monitoring processes and outcomes [1–3]. To date, attention has mainly focused on the application of univariate control charts. In most clinical monitoring cases, however, there is a need to track more than one quality characteristic. If separate univariate charts are used to monitor each quality characteristic, the overall probability of a false alarm may be inflated (unless the control limits are adjusted accordingly) since any correlation between the variables is ignored. This suggests that it might be worthwhile adopting multivariate techniques.

In this paper we survey some of the charts available for monitoring the means of continuous variables. We consider their implementation and performance in light of the challenges faced in clinical settings. In particular, we address the fact that clinical records are frequently incomplete.

The paper proceeds as follows. In Section 2 we outline the general multivariate framework and explain how to construct

Hotelling's T^2 , the multivariate exponentially weighted moving average (MEWMA) and the multivariate cumulative sum (MCUSUM) charts. The discussion is supported by a case study involving the monitoring of radiation delivered to patients undergoing diagnostic coronary angiogram procedures at St Andrew's War Memorial Hospital, Australia. We also outline the methodology of a simulation study used to investigate how each chart performs for various correlation structures, patterns of mean shifts, amounts of missing data and methods of imputation. Results are given in Section 3, and we provide recommendations in Section 4 on chart selection and implementation.

Methods

Description of case study data

Our data set contains information for three variables linked to the radiation delivered to a patient, namely the dose area product, fluoroscopy time and the number of digital images

(frames) acquired during coronary angiogram procedures at St Andrew's War Memorial Hospital between April 2005 and December 2008. Dose area product, measured in mGy cm², provides a measure of the total radiation to which a patient is exposed. It is affected by the fluoroscopy time (low radiation dose rate component of the study associated with positioning catheters in the heart), the number of frames (high dose rate documentation phase), and other procedural and clinical factors such as the patient's weight. To minimize variations in dose area product associated with patient size, data included in the case study have been limited to female patients only. Under radiation safety and protection guidelines, every effort is taken to limit patient radiation exposure. To this end St Andrew's War Memorial Hospital routinely monitors and reviews dose area product, fluoroscopy time and frames separately in an effort to achieve an optimized risk versus benefit balance. Although the number of frames is technically discrete, we will regard it as continuous.

A general framework for multivariate monitoring

Let \mathbf{X}_i be the i th vector of observations for the p variables that we want to monitor. With respect to the case study, \mathbf{X}_i comprises the values of dose area product, fluoroscopy time and frames for the i th patient. For example, if the tenth patient had a dose area product of 30 638 mGy cm², a fluoroscopy time of 2.56 min and 622 frames were taken of their heart, then $\mathbf{X}_{10} = [30638 \ 2.56 \ 622]'$.

When the process is in-control, it is assumed that \mathbf{X}_i follows a multivariate normal distribution, with mean vector $\boldsymbol{\mu}_0$ and covariance matrix $\boldsymbol{\Sigma}$, independent of other observations.

That is, $\mathbf{X}_i \stackrel{\text{iid}}{\sim} N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma})$. There will be many occasions, however, where clinical data do not satisfy this assumption. The MEWMA chart can be designed to be robust against deviations from normality [4, 5], and there exists a non-parametric version of the MCUSUM chart [6], but the T^2 chart is highly sensitive to the normality assumption [4].

When normality is questionable, it will often suffice to transform one or more variables. For example, although dose area product and fluoroscopy time are both strongly right-skewed, normality appears to hold for the natural log of dose area product and the inverse of fluoroscopy time. Consequently, instead of creating control charts using the raw data, we would construct them for $\mathbf{X}_i^* = [D \ T \ F]'$, where $D = \ln(\text{dose area product})$, $T = (\text{fluoroscopy time})^{-1}$ and F denotes the number of frames. Hence, $\mathbf{X}_{10}^* = [10.3 \ 0.39 \ 622]'$.

The parameters $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}$ can either be specified by management or estimated using a sample from a stable process. We assume that $\mathbf{X}_i^* \stackrel{\text{iid}}{\sim} N_3(\boldsymbol{\mu}_0, \boldsymbol{\Sigma})$, where

$$\boldsymbol{\mu}_0 = \begin{pmatrix} 9.5 \\ 0.55 \\ 586 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{bmatrix} 0.2 & -0.03 & 23.8 \\ -0.03 & 0.04 & -6.0 \\ 23.8 & -6.0 & 14882 \end{bmatrix}.$$

The objective is to detect a shift from $\boldsymbol{\mu}_0$ to $\boldsymbol{\mu}_1$. The T^2 , MEWMA and MCUSUM charts consider only the

magnitude of any shift and not its direction. Hence, they use only an upper control limit (UCL). If a statistic exceeds the UCL, the chart is said to 'signal', and the process should be investigated to determine if the signal is due to an error in the data, is indicative of a genuine shift, or simply the result of natural variability. Univariate charts and the raw data should be inspected to determine the variable(s) responsible for the signal and whether it is associated with a change in patterns of use of radiation or variation in imaging equipment performance. In terms of our case study, a signal would correspond to increased levels of radiation exposure.

From a clinical governance perspective, deviations from a stable process must be identified as quickly as possible, while limiting the occurrence of false alarms. Performance of a control chart is described in terms of the average number of observations that are monitored, average run length (ARL), before the chart 'signals'. For each chart, choice of the signal threshold is a trade-off between the ARL when the process is in-control (ARL_0) and when the process is out-of-control (ARL_1). Under ideal circumstances the chart should have a very low false alarm rate (long ARL_0) while rapidly detecting true changes (short ARL_1).

Before we can construct a multivariate chart, we need to deal with any missing data. One solution is to use imputation, methods of which include multiple imputation [7], insertion of the sample mean and regression-based imputation.

Multiple imputation is preferred because it preserves variability in the missing values and performs well for small sample sizes and/or large proportions of missing data. Generally speaking, it involves creating $3 \leq r \leq 10$ complete data sets, performing analyses on each of these data sets and then combining the results. For each variable with missing values, it is necessary to construct r imputation models. In practice, most researchers will not need to develop these models directly, since software, such as NORM [8, 9], is available that performs multiple imputation.

The rates of missing data for D , T and F are 3.0%, 1.4% and 1.5%, respectively. Instead of constructing control chart statistics for multiple data sets and then combining the results, we use multiple imputation to create five observations for each missing value, and we impute the average of these five values to create a single complete data set.

Control chart construction

We have chosen to concentrate on the T^2 , MEWMA and MCUSUM charts because they are extensions of univariate charts commonly used to monitor clinical data, namely the Shewhart, the exponentially weighted moving average (EWMA) and the cumulative sum (CUSUM) charts. In what follows, we briefly describe how these charts are constructed. A more detailed explanation can be found in [10]. See also [11] for a more comprehensive survey of research into multivariate charts.

The T^2 chart plots $T_i^2 = (\mathbf{X}_i - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}^{-1} (\mathbf{X}_i - \boldsymbol{\mu}_0)$ [12]. If $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}$ have been specified by management or estimated using a sufficiently large sample (in excess of 100

observations), then the UCL is $\chi^2_{\alpha,p}$ [13]. If they have been estimated using a ‘small’ sample, the UCL depends upon whether the researcher is performing a retrospective (Phase I) analysis or wants to monitor future values (Phase II). Phase I and II UCLs are $\beta_{\alpha,p/2,(n-p-1)/2}[(n-1)^2/2]$ and $F_{\alpha,p,n-p}[p(n+1)(n-1)/(n^2-np)]$, respectively, where n is the sample size [14].

Construction of the MEWMA chart [15] requires specification of a weight λ ($0 \leq \lambda \leq 1$) that is used to assign importance to observations, with recent observations being weighted more heavily than observations more distant in time. Letting $Z_i = \lambda(X_i - \mu_0) + (1 - \lambda)Z_{i-1}$, where $Z_0 = 0$, it plots $Z'_i \Sigma^{-1} Z_i$, where $\Sigma_{Z_i} = (\lambda/(2 - \lambda))[1 - (1 - \lambda)^{2i}] \Sigma$. Prabhu and Runger [16] determined the optimal weight and corresponding UCL for selected combinations of p , the size of the shift to be detected, and the desired ARL_0 . For cases not considered, the UCL can be obtained through simulation in order to achieve a desired ARL_0 .

Several versions of the MCUSUM chart have been proposed [17, 18]. We consider a version proposed by Crosier [17]. It plots $(L'_i \Sigma^{-1} L_i)^{1/2}$, where

$$L_i = \begin{cases} 0 & \text{if } C_i \leq k, \\ (L_{i-1} + X_i - \mu_0)(1 - k/C_i) & \text{if } C_i > k \end{cases}$$

and $C_i = ((L_{i-1} + X_i - \mu_0)' \Sigma^{-1} (L_{i-1} + X_i - \mu_0))^{1/2}$. Crosier [17] recommended setting $L_0 = \mathbf{0}$ and $k = ((\mu_1 - \mu_0)' \Sigma^{-1} (\mu_1 - \mu_0))^{1/2}/2$, and we follow the convention of resetting the MCUSUM chart following a signal. The UCL is calculated by simulation in order to achieve a desired ARL_0 .

Statistics were generated for all 884 records in the case study data set using code written in Matlab. In Figs 1–3 we show the T^2 , MEWMA and MCUSUM charts for

procedures performed in November 2005. The figures were produced using R [19]. We discuss chart interpretation in Section 3.1.

Outline of simulation study

The simulation study considers the monitoring of $X_i = [V_1 V_2 V_3]'$, where the correlation between variables V_i and V_j ,

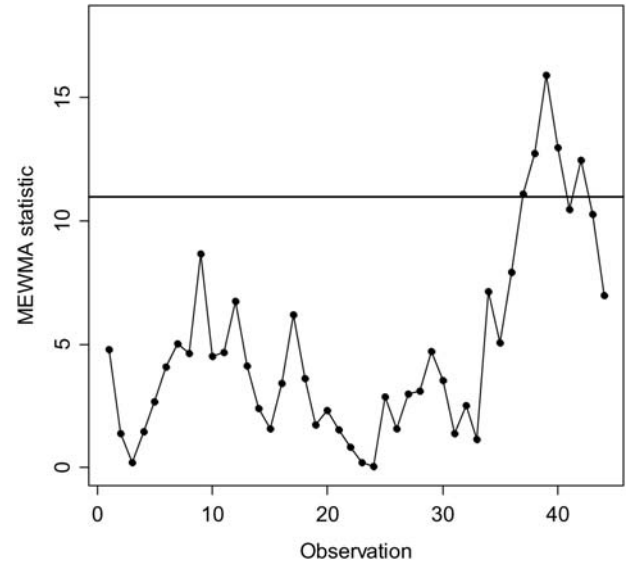


Figure 2 MEWMA chart for the simultaneous monitoring of $\ln(\text{dose area product})$, $(\text{fluoroscopy time})^{-1}$ and frames for females undergoing a coronary angiogram in November 2005.

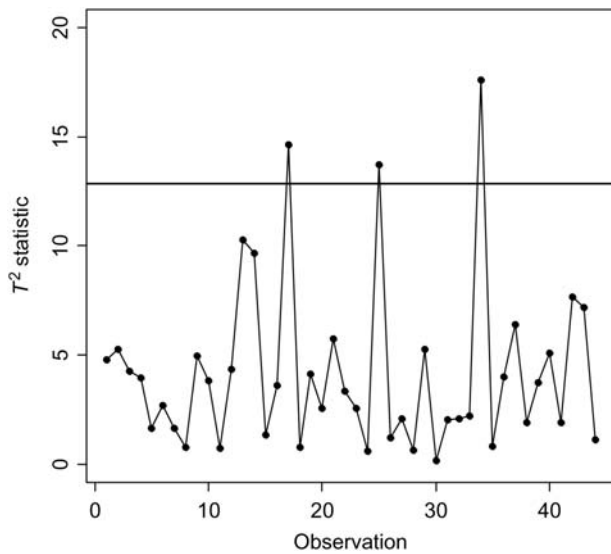


Figure 1 Hotelling’s T^2 chart for the simultaneous monitoring of $\ln(\text{dose area product})$, $(\text{fluoroscopy time})^{-1}$ and frames for females undergoing a coronary angiogram in November 2005.

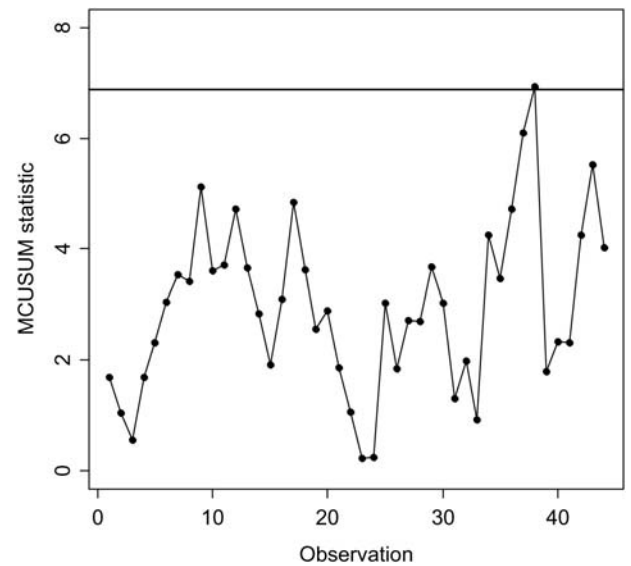


Figure 3 MCUSUM chart for the simultaneous monitoring of $\ln(\text{dose area product})$, $(\text{fluoroscopy time})^{-1}$ and frames for females undergoing a coronary angiogram in November 2005.

denoted by ρ_{j_1} , takes the values 0, 0.2 and 0.8. To study ARL_0 , we simulate data from $N_3(\mathbf{0}, \mathbf{\Sigma})$, where $\mathbf{\Sigma}$ is in correlation form. To study ARL_1 , we generate data such that the first 50 records are drawn from an in-control process, and the remaining records are drawn from $N_3(\boldsymbol{\delta}, \mathbf{\Sigma})$, where $\boldsymbol{\delta} = [\delta_1 \delta_2 \delta_3]' \neq \mathbf{0}$. We let δ_i be 0, 0.5 or 2, and we allow for shifts of different magnitudes among the variables. Data are generated such that V_i is missing with probability γ , where $\gamma = 0, 0.05$ or 0.2 , subject to the constraint that a record cannot have all three of its observations missing. When $\gamma > 0$, a complete data set is then obtained by multiple imputation, imputation of the mean or regression-based imputation. When the process is stable, the run length is the number of records before the chart signals. When the process is unstable, the run length is the number of out-of-control records before the chart signals. Run lengths for 10 000 data sets are averaged to produce $ARLs$.

Parameters for the multivariate charts have been chosen such that ARL_0 is ~ 200 . The UCL of the T^2 chart is 12.85. We use $\lambda = 0.1$ and a UCL of 10.97 for the MEWMA chart, and $k = 0.5$ and a UCL of 6.88 for the MCUSUM chart.

In addition to creating multivariate charts, we construct separate univariate charts for each variable. For a particular univariate chart, we define the ‘overall’ run length to be the shortest of the three run lengths. For each Shewhart chart, we use the standard lower and upper control limits of -3 and 3 , respectively. For each EWMA chart, we use $\lambda = 0.1$ and $L = 2.814$. Given these values, the control limits for the i th observation are $\pm 2.814\sqrt{0.1[1 - 0.9^{2i}]/1.9}$. For each CUSUM chart we use $K = 0.5$ and a UCL of 5.

Results

Case study

The T^2 chart (Fig. 1) registered out-of-control signals at times 17, 25 and 34. The MEWMA chart first signalled at time 37 and the MCUSUM chart signalled at time 38 (Figs 2 and 3). Since the T^2 statistic only uses information from the current observation, investigation into the causes of the T^2 signals requires inspection of only the 17th, 25th and 34th records. In contrast, when interpreting MEWMA and MCUSUM signals, we should also consider records preceding a signal.

To help identify the causes of signals, we constructed Shewhart, EWMA and CUSUM charts for each variable. The Shewhart chart for D suggested that high values of dose area product may be the cause of the signals observed in the T^2 chart. Moreover, $\mathbf{X}_{17} = [50011 \ 5.26 \ 506]$, $\mathbf{X}_{25} = [67508 \ 2.27 \ 803]$ and $\mathbf{X}_{34} = [84965 \ 7.14 \ 816.2]$. In each case, the dose area product is considerably higher than the stable mean of $\exp(9.5) = 13360 \text{ mGy cm}^2$. The MEWMA and MCUSUM signals are perhaps due to the high dose area product in \mathbf{X}_{34} , but it is also interesting to note that the EWMA chart for F exhibited an increasing trend starting at observation 20. The fluoroscopy time (T) was stable during November 2005. There were no false alarms associated with any of the EWMA or CUSUM charts, or with the Shewhart charts used to

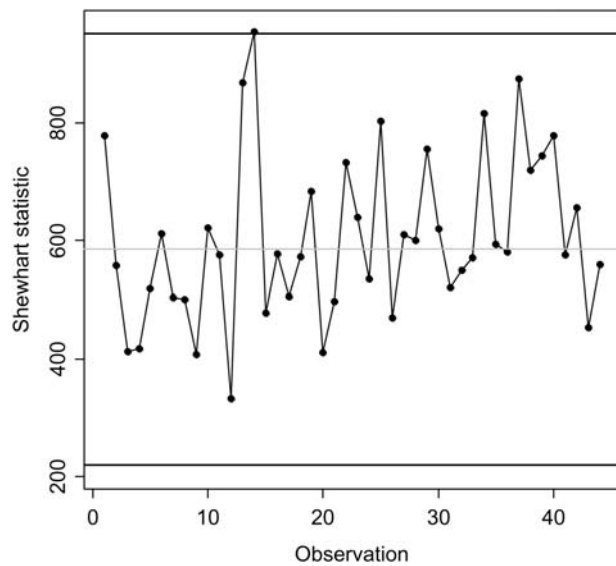


Figure 4 Shewhart chart for frames for females undergoing a coronary angiogram in November 2005.

monitor D and T . However, the Shewhart chart for F generated a false alarm at time 14 (Fig. 4). The number of frames (955) is reasonable when considered simultaneously with the dose area product of $28\,283 \text{ mGy cm}^2$ and the fluoroscopy time of 3.7 min. Moreover, an increasing trend in F does not begin until several observations later. As an aside, six of the 44 records from November 2005 required imputation. The inverse of fluoroscopy (F) was imputed in \mathbf{X}_9 , \mathbf{X}_{23} , \mathbf{X}_{24} and \mathbf{X}_{34} , T was imputed in \mathbf{X}_{26} and D was imputed in \mathbf{X}_{27} .

Starting around May 2006, there was a sharp increase in number of signals across all types of charts, both multivariate and univariate. This was due to a change in equipment and processes used at St Andrew’s War Memorial Hospital. If monitoring the situation in ‘real time’, we would have allowed the process to stabilize, before re-estimating $\boldsymbol{\mu}_0$ and $\mathbf{\Sigma}$ to determine parameters appropriate for the changed conditions.

Simulation study

In this section we summarize broad trends. Tables of results for all charts, combinations of parameters and types of imputation are available upon request from the first author.

In keeping with our choice of parameters, if $\gamma = 0$, we expect a false alarm every 200 records, on average. As the amount of missing data, and hence imputation, increases, so too does ARL_0 . For example, when $\gamma = 0.2$ and multiple imputation is used, ARL_0 is ~ 300 for the T^2 chart, and ~ 400 for the MEWMA and MCUSUM charts. For almost all scenarios considered, using separate univariate charts resulted in a quicker false alarm than using the multivariate counterpart.

Regardless of correlation structure, the MEWMA and MCUSUM charts detect small to moderate shifts significantly quicker than the T^2 chart. For small shifts, the MCUSUM chart marginally outperforms the MEWMA chart. For large shifts, the performance of all charts is essentially the same.

The charts are slower to detect changes as the amount of imputation performed increases. However, the effect is negligible for large shifts. The T^2 chart is most affected by imputation.

The MEWMA and MCUSUM charts are affected to the same extent. Figure 5 is representative of the trends described above. It plots ARL_1 versus the shift size, as summarized by $\|\delta\| = \sqrt{\delta_1^2 + \delta_2^2 + \delta_3^2}$.

When there is little or no correlation between the variables, using multiple Shewhart charts detects small shifts more quickly than a T^2 chart. In contrast, the MEWMA charts detects small shifts almost as quickly as multiple EWMA charts, and the MCUSUM chart is actually superior to multiple CUSUM charts, even when the variables are completely uncorrelated. When the variables are highly correlated, the multivariate charts tend to detect an unstable process more quickly than multiple univariate charts. The multivariate charts are slower when the means of all three variables have shifted by the same amount.

If at least one pair of variables is highly correlated, then small shifts are detected most quickly when the average has been imputed. However, under these conditions, imputing the average produces the worst ARL_0 . For example, when $\gamma = 0.2$, $\rho_{12} = 0.2$, $\rho_{13} = 0.2$ and $\rho_{23} = 0.8$, ARL_0 is 160 if the average is imputed. When multiple imputation and regression-based imputation are used, ARL_0 is 422 and 510, respectively. If a large amount of imputation is required, multiple imputation tends to produce better results than regression-based imputation, but the choice of imputation technique is largely irrelevant if interest lies in detecting moderate to large shifts.

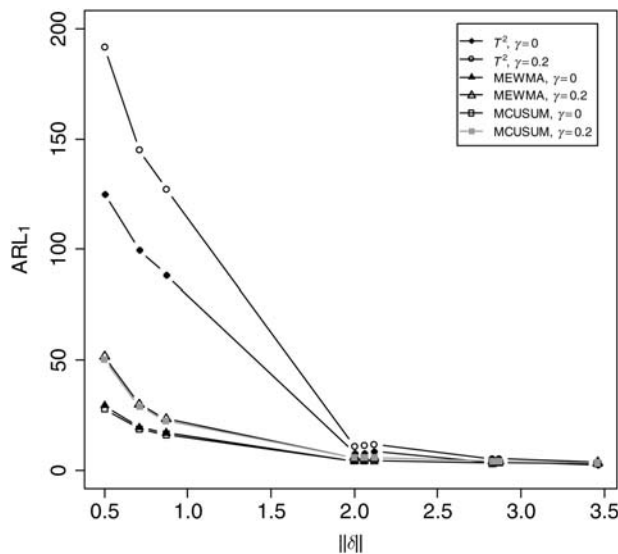


Figure 5 Plot of ARL_1 versus the magnitude of the shift in the mean vector, $\|\delta\|$, for the T^2 , MEWMA and MCUSUM charts, given $\rho_{12} = \rho_{13} = \rho_{23} = 0.2$. Results are shown for the cases where no data are missing ($\gamma = 0$) and when $\gamma = 0.2$. In the latter case, multiple imputation has been used to impute for missing values.

Discussion

When there is more than one quality characteristic to be monitored, we advise using multivariate charts to avoid excessive false signals associated with using separate univariate charts. Of the charts considered in this paper, the MCUSUM chart showed the best overall performance. However, it is only marginally superior to the MEWMA chart, with differences becoming negligible for moderate to large shifts. Indeed, many clinicians may feel that the ability of the MCUSUM chart to detect small shifts quicker than the MEWMA chart is outweighed by the increased complexity of its construction. This is especially pertinent given that many statistical software packages do not include an in-built function for creating MCUSUM charts. We recommend strongly against relying on the T^2 chart. However, if the data follow a multivariate normal distribution, then the T^2 chart can be used in a supplementary manner for the purposes of quickly detecting large shifts, as demonstrated in our case study.

Our case study highlighted some standard transformations that can be used when data do not follow a normal distribution. If multivariate normality is questionable and transformations prove unsatisfactory, then the MEWMA chart should be used. In this case, the clinician should follow the design recommendations in [4] or [5].

In addition to having a skewed distribution, dose area product is strongly related to patient weight, with heavier patients exposed to higher levels of radiation, on average. As such, what is considered a ‘normal’ dose area product depends upon the patient’s size. Since weight was only recorded for $\sim 4\%$ of the patients in the case study data set, it was not feasible for it to be used as another covariate in our charts. As such, we used gender as a surrogate measure of weight, constructing charts for only the females. Partitioning records in this way is useful if the distribution of one or more of the quality characteristics is multimodal. If setting up an ongoing multivariate chart for the case study, in the absence of weight data, we would continue to monitor records for males and females separately.

Multivariate charts are known to perform well for a moderate number of variables. However, as the number of variables increases they become less efficient in detecting shifts. If more than 10 variables are to be monitored, we recommend using principal components analysis to reduce the dimensionality of the problem. Details on this procedure with respect to control charts can be found in [11].

Supplementary use of univariate charts can be used to investigate the cause(s) of multivariate signals. They can be used to examine the behaviour of individual variables and to identify the direction of any shift(s). If the signal is associated with an improvement, management should make efforts to maintain whatever procedures precipitated the change. If a signal suggests an undesirable change, investigations should be conducted to determine whether the signal is a result of a genuine deterioration in the process, a mistake in the data collection process or the result of natural variability.

If data are missing, we recommend using multiple imputation to create complete data sets, particularly if a large amount of imputation is required. In our study, we imputed the average of five observations for each missing value to create a single complete data set. An alternative approach would be to use multiple imputation to create five distinct data sets, calculate statistics for each data set and plot the average of the statistics. This method is more difficult to implement and interpretation of signals becomes more complicated.

We advise caution when using imputation based on the sample mean or regression because these methods artificially reduce variability and may also distort relationships between variables. We strongly advise against deleting incomplete records. Considering our case study, if X_{34} had been discarded on account of the number of frames being missing, then the unusually large dose area product associated with that record would have been overlooked.

Because imputation reduces the amount of variability in the data, it has the effect of increasing both ARL_0 and ARL_1 . While the former is not considered a problem, a higher ARL_1 is less desirable. This effect can be countered somewhat by setting a lower UCL than would be used for a data set with no missing values.

When estimating Σ in the case study, we used the sample covariance matrix. An alternative approach uses the differences between successive vectors of observations [20]. This is analogous to using the moving range to estimate the standard deviation. In this case the estimator of Σ is $S = (1/2(n-1))VV'$, where V is a $p \times (n-1)$ matrix, the i th column of which is given by $X_{i+1} - X_i$, for $i = 1, 2, \dots, n-1$. If successive observations are independent, this estimator should be used because it results in a chart that is better able to detect sustained shifts in the mean vector [21]. It should not be used if observations are autocorrelated because it results in a large false alarm rate [22].

We considered continuous clinical data. If the institution wants to monitor multiple discrete quality characteristics, a multi-attribute chart should be used instead. There has been less research into multi-attribute charts, but the interested reader is referred to [23–25].

Acknowledgements

The authors thank Dr Tony Morton for helpful discussions.

Funding

This research was funded by an Australian Research Council Linkage Project.

References

1. Spiegelhalter D, Grigg O, Kinsman R *et al*. Risk-adjusted sequential probability ratio tests: applications to Bristol, Shipman and adult cardiac surgery. *Int J Qual Health Care* 2003;**15**:7–13.
2. Grigg OA, Farewell VT, Spiegelhalter DJ. Use of risk-adjusted CUSUM and RSPRT charts for monitoring in medical contexts. *Stat Methods Med Res* 2003;**12**:147–70.
3. Sherlaw-Johnson C, Gallivan S, Treasure F *et al*. Computer tools to assist the monitoring of outcomes in surgery. *Eur J Cardiothorac Surg* 2004;**26**:1032–6.
4. Stoumbos ZG, Sullivan JH. Robustness to non-normality of the multivariate EWMA control chart. *J Qual Technol* 2002;**34**:260–76.
5. Testik MC, Runger GC, Borrer CM. Robustness properties of multivariate EWMA control charts. *Qual Reliab Eng Int* 2003;**19**:31–8.
6. Qiu PH, Hawkins D. A nonparametric multivariate cumulative sum procedure for detecting shifts in all directions. *J Roy Stat Soc D-Sta* 2003;**52**:151–64.
7. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley, 1987.
8. Schafer JL. NORM: Multiple imputation of incomplete multivariate data under a normal model, version 2. Software for Windows 95/98/NT, available from <http://www.stat.psu.edu/~jls/misoftwa.html>.
9. Schafer JL, Olsen MK. Multiple imputation for multivariate missing-data problems: a data analyst's perspective. *Multivar Behav Res* 1998;**33**:545–71.
10. Montgomery DC. *Introduction to Statistical Quality Control*, 6th edn. New Jersey: Wiley, 2008.
11. Bersimis S, Psarakis S, Panaretos J. Multivariate statistical process control charts: an overview. *Qual Reliab Eng Int* 2007;**23**:517–43.
12. Hotelling H. Multivariate quality control—illustrated by the air testing of sample bombsights. In: Eisenhart C, Hastay MW, Wallis WA (eds). *Techniques of Statistical Analysis*. New York: McGraw-Hill, 1947, 111–84.
13. Seber GAF. *Multivariate Observations*. New York: Wiley, 1984.
14. Tracy ND, Young JC, Mason RL. Multivariate control charts for individual observations. *J Qual Technol* 1992;**24**:88–95.
15. Lowry CA, Woodall WH, Champ CW *et al*. A multivariate exponentially weighted moving average control chart. *Technometrics* 1992;**34**:46–53.
16. Prabhu SS, Runger GC. Designing a multivariate EWMA control chart. *J Qual Technol* 1997;**29**:8–15.
17. Crosier RB. Multivariate generalizations of cumulative sum quality control schemes. *Technometrics* 1988;**30**:291–303.
18. Pignatiello JJ, Runger GC. Comparisons of multivariate CUSUM charts. *J Qual Technol* 1990;**22**:173–86.
19. R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, 2009.
20. Holmes DS, Mergen AE. Improving the performance of the T^2 control chart. *Qual Eng* 1993;**5**:619–25.
21. Sullivan JH, Woodall WH. A comparison of multivariate control charts for individual observations. *J Qual Technol* 1996;**28**:398–408.

22. Capilla C. Application and simulation study of the Hotelling's T^2 control chart to monitor a wastewater treatment process. *Environ Eng Sci* 2009;**26**:333–41.
23. Patel HI. Quality control methods for multivariate binomial and Poisson distributions. *Technometrics* 1973;**15**:103–12.
24. Lu XS, Xie M, Goh TN *et al.* Control charts for multivariate attribute processes. *Int J Prod Res* 1998;**36**:3477–89.
25. Skinner KR, Montgomery DC, Runger GC. Process monitoring for multiple count data using generalized linear model-based control charts. *Int J Prod Res* 2003;**41**:1167–80.