Review paper

Moving towards process-based radiotherapy quality assurance using statistical process control

Vysakh Raveendran a,b,∗, Ganapathi Raman R c, Anjana P.T. a, Saju Bhasi d, Ranjith C.P. a, Rajesh Ashok Kinhikar e

a Department of Radiation Oncology, Advanced Centre for Treatment Research and Education in Cancer, Tata Memorial Centre, Homi Bhabha National Institute, Navi Mumbai, Maharashtra, India.

b Department of Physics, Noorul Islam Centre for Higher Education, Kumaracoil, Kanyakumari District, Tamil Nadu, India.

c Department of Physics, Saveetha Engineering College (Autonomous), Chennai, Tamil Nadu, India.

d Division of Radiation Physics, Regional Cancer Centre, Trivandrum, India.

e Department of Medical Physics, Tata Memorial Centre, Homi Bhabha National Institute Parel, Mumbai, India.

ARTICLE INFO

Keywords:
Statistical Process control
Radiotherapy Quality assurance
Linac QA
Patient specific IMRT QA

ABSTRACT

Monitoring Radiotherapy Quality Assurance (QA) using Statistical Process Control (SPC) methods has gained wide acceptance. The significance of understanding the SPC methodologies has increased among the medical physics community with the release of Task Group (TG) reports from the American Association of Physicists in Medicine (AAPM) on patient-specific QA (PSQA) (TG-218) and Proton therapy QA (TG–224). Even though these reports recommend using SPC for QA analysis, physicists have ambiguities and doubts in choosing proper SPC tools and methodologies. This review article summarises the utilisation of SPC methods for different Radiotherapy QAs published in the literature, such as PSQA, routine Linac QA and patient positional verification. QA analysis using SPC could assist the user in distinguishing between ‘special’ and ‘routine’ sources of variations in the QA, which can aid in reducing actions on false positive QA results. For improved PSQA monitoring, machine-specific, site-specific, and technique-specific Tolerance Limits and Action Limits derived from a two-stage SPC-based approach can be used. Adopting a combination of Shewhart’s control charts and time-weighted control charts for routine Linac QA monitoring could add more insights to the QA process. Incorporating SPC tools into existing image review modules or introducing new SPC software packages specifically designed for clinical use can significantly enhance the image review process. Proper selection and having adequate knowledge of SPC tools are essential for efficient QA monitoring, which is a function of the type of QA data available, and the magnitude of process drift to be monitored.

1. Introduction

Monitoring components involved in healthcare require standardisation and process monitoring through rigorous quality control protocols and guidelines. The statistical process control (SPC) technique has always been a powerful and effective tool for quality improvement in various manufacturing industries. Being the most common tool in SPC, the use of control charts in healthcare industries has been extensively reviewed by Suman and Prajapati [1], Tennant et al. [2], and Noyez et al. [3]. However, none of these reviews included the application of control charts in Radiotherapy. A straightforward interpretation of the application of SPC methods in Radiotherapy was demonstrated by Pawlicki et al. via multiple publications [4–6]. They demonstrated the difference between the conventional way of performing Quality Assurance (QA) based on standard deviation and SPC-based QA analysis. The traditional QA process consists of checking whether the QA results are within specification; in case of any failure, time and effort are put into rectifying and determining the causes of QA failure. However, this approach only gives a binary decision, which may lead to passive process negligence followed by occasions of intense troubleshooting as the QA fails. Pawlicki and Whitaker [5] demonstrated that this happens because of the lack of knowledge of how the process varies within the
A two-dimensional representation of the parameter under evaluation over time or sample number is called a control chart. The parameter representation can be the mean of the sample subgroup or individual sample value. Results from any QA process can be represented using an SPC-based control chart with statistically derived limits which can be used to acquire knowledge about the process stability and improve process outcomes. Control charts are used to separate the ‘special’ and ‘routine’ causes of variation, analogues to systematic and random error concepts [4]. A process is said to be ‘in control’ when it consists only of predictable routine causes of variation, while an ‘out of control process’ contains unpredictable special reasons for variation [4]. Ideally, all process variations should be eliminated if known, which is unrealistic in clinical practice. The utilization of standard control charts involves two distinct phases, namely Phase I and Phase II, each serving different purposes. In Phase I, a comprehensive analysis is conducted on collected sample data to establish trial control limits and assess process control, which helps achieve statistical control. The limits obtained from the stable Phase I data are utilised for Phase II monitoring, and subsequent data samples are monitored using the control charts [7]. Shewhart’s control charts are a good choice for Phase I application since these control charts are effective in detecting large process drifts, while time-weighted control charts can be utilised for Phase II application.

In a control chart, the mean value of the parameter under evaluation is plotted as the central line. Above and below the central line, the Control Limits (CL) and Action Limits (AL) are plotted. The CL for the SPC-monitored QA process are calculated using process attributes from a chosen subset. Conversely, the AL can be either universally defined (Universal AL) or locally defined and driven by local experience [10]. The AL are always chosen subset. Conversely, the AL can be either universally defined (Universal AL) or locally defined and driven by local experience [10]. The AL are always directly correlated with treatment outcome. However, locally defined AL can be utilised to ensure proper patient care by keeping it always below the Universal AL [10]. The CL and Tolerance Limits (TL) may not always be the same, as CL solemnly depends upon the process variability as it is calculated from a chosen set of stable subgroups. It is a user’s choice how to define TL for the monitored QA process. Some frequently utilised control charts for Radiotherapy QA purposes are briefly discussed below.

Shewhart’s Individual and Moving Range Chart (I-MR Charts): The I-MR chart is a combination of the individual value control chart, and moving range chart, which is used for detecting rare events or outliers in a process (which are above or equal to 1.5 standard deviations). The I-chart indicates the data location using individual data points, while the range is the difference between two consecutive points, indicating the data dispersion. Subgrouping the data to find the limits is difficult if the data is added slowly; I-MR charts are the best choice for process monitoring in these scenarios, as suggested by Montgomery et al. [7]. I-chart is drawn with the parameters calculated with equations (1) to (3):

\[
\begin{align*}
\text{Upper control limit} & = \bar{x} + 3 \frac{MR}{d_2} \\
\text{Center line} & = \bar{x} \\
\text{Lower control limit} & = \bar{x} - 3 \frac{MR}{d_2}
\end{align*}
\]

Parameters for the MR chart are calculated with equations (4) to (6) as follows:

\[
\begin{align*}
\text{Upper control limit} & = \left(1 + 3 \frac{d_2}{d_3}\right) MR \\
\text{Center line} & = MR \\
\text{Lower control limit} & = \left(1 - 3 \frac{d_2}{d_3}\right) MR
\end{align*}
\]

where \(\bar{x}\) is process mean and \(MR\) is the mean moving range; \(d_2\) and \(d_3\) were chosen to be 1.128 and 0.8525 for I-MR charts.

**\(X\) and R control chart:** In an \(X\) chart changes in the process mean are analysed over time, for which the individual samples are grouped into subgroups and subgroup averages are plotted against subgroup number.

Typically, each subgroup contains about 4 to 5 individual samples [7]. The associated R chart plots the difference between each subgroup’s largest and smallest observations against the sample number. The \(X\) chart is more effective than the I chart at detecting small but consistent shifts in the process mean, which can be essential for maintaining process quality. For retrospective analysis of Radiotherapy QA results, these charts are frequently utilised.

**\(X\) and S control chart:** \(X\) and S control charts are preferred when the process standard deviation is unknown or difficult to estimate. S chart is used to estimate the process standard deviation, and the \(X\) chart is used to monitor the process mean. When sample or variable size is large (n greater than 10) these charts are preferred over the \(X\) and R control charts. This chart is more suitable when the process is stable and in control, and when small changes in the process mean are important to detect.

**Exponentially Weighted Moving Average (EWMA) Control Chart:**

Due to its peculiar construction, the EWMA chart allows it to find small drifts in the monitored processes. Each point in an EWMA plot is produced using the weighted mean of the prior points, with the most weightage assigned to the latest data point and then decreasing exponentially chronologically. Due to its peculiar construction, the CLs also changes with the sample number, and later it saturates in an EWMA chart. Mathematically, the EWMA value \(z_i\) of the \(x_i^{th}\) data point is:

\[
EWMA\ value, z_i = \lambda x_i + (1 - \lambda)z_{i-1}
\]
where for the $x_{i,j}$th data point, the EWMA is $z_{i,j}$, and the constant $\lambda$ decides the weight of prior data points ($0 < \lambda \leq 1$) via an exponential function $\lambda (1-\lambda), \lambda (1-\lambda)^2, \ldots$. The process target is the first EWMA value, $z_0 = T$. The following equations are used to calculate the parameters for an EWMA chart,

\[
\text{Upper control limit} = \mu_0 + L\sigma \sqrt{\frac{1}{1-\lambda}(1-(1-\lambda)^2)}
\]

(8)

\[
\text{Central line} = \mu_0
\]

(9)

\[
\text{Lower control limit} = \mu_0 - L\sigma \sqrt{\frac{1}{1-\lambda}(1-(1-\lambda)^2)}
\]

(10)

where, $\mu_0$ is the process mean, and $1$ gives the control limit width.

**The Cumulative Sum (CUSUM) Control Chart:** A CUSUM control chart is a time-weighted control chart that is effective in finding small drifts in the process observed. The CUSUM ($C_i$) of differences of the sample characteristic from a set target sample value is plotted against the sample number in a CUSUM chart. For a target value $T$ and sample value $x_j$ the $C_i$ up to and including the $i^{th}$ sample is given by

\[
C_i = \sum_{j=1}^{i} (x_j - T)
\]

(11)

If the sample parameter approaches $T$, the $C_i$ will swing around zero when the process is under control. If the process is out of control, it will signal as the $C_i$ will take different values. The one-sided tabular CUSUM values indicates the drifts in either, upper ($C_i^+ (x_j > T)$) or lower ($C_i^- (x_j < T)$) directions. These are calculated as follows:

Upper tabular CUSUM, $C_i^+ = \max(0, x_i - (T + K) + C_{i-1})$

(12)

Lower tabular CUSUM, $C_i^- = \max(0, (T - K) - x_i + C_{i-1})$

(13)

When $i = 0$, $C_i^+ = C_i^- = 0$, and $K$ is called the allowance value or the reference value, which is chosen based on the magnitude of the monitored process shifts. It is calculated as

\[
\text{Allowance value, } K = \frac{\Delta}{2}\sigma
\]

(14)

where the term $\Delta$ represents the standard deviation of the process shifts. Another important factor that decides the CUSUM control chart is the decision interval $H$. A process is said to be out of control if the CUSUM statistic of the sample/process surpasses the decision interval $H$, which is usually chosen as five times the standard deviation. Fig. 1 shows the control charts for Output constancy analysis of 10 MV photon beam as reported by Vysakh et al., [11]. The Fig. 1 demonstrates the difference in sensitivity of the four control charts in determining process drifts.

**Bayesian Control Chart:** A Bayesian Control Chart is a SPC tool that uses Bayesian statistics, which applies probabilities based on prior knowledge. Unlike traditional control charts that use observed data alone, Bayesian control charts combine prior knowledge and new data to analyze process variability. Uses and advantages of Bayesian Control Charts include; 1) Improved Accuracy: They provide accurate results even when data is limited or uncertainty is high. 2) Prior Knowledge Integration: They allow the incorporation of expert or historical data into the analysis. 3) Dynamic Updating: They continuously update the probability distributions of process parameters as new data is gathered. 4) Flexibility and Robustness: They can handle different types of data and distributions. 5) Risk Analysis: They offer a probabilistic measure of risk for decision-making in quality control and risk management. In essence, Bayesian control charts are valuable for statistical process control in complex environments, offering dynamic interpretation and risk analysis.

**Heuristic control charts:** Heuristic control charts, also known as adaptive control charts, are statistical tools used in quality control and process monitoring. Unlike traditional control charts, heuristic control charts dynamically adjust their limits based on observed data, allowing for better detection of process changes. The primary purpose of heuristic control charts is to identify variations in a process that may be indicative of assignable causes or special events. These variations could include
A summary of studies which utilised SPC methods for analysing PSQA.

**Table 1**

<table>
<thead>
<tr>
<th>Author and Year</th>
<th>SPC tools used</th>
<th>Objective and QA analysed</th>
<th>Observations / Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Todd Pawlicki et al., 2008 [6]</td>
<td>I-MR charts</td>
<td>To establish limits for PSQA using SPC tools and compare these limits with those calculated using the standard deviation (σ) approach.</td>
<td>SPC methods were found to be better than the standard deviation-based approach in categorizing process performance, which could minimise protocol variation and guide process improvements.</td>
</tr>
<tr>
<td>Stephen L. Breen et al., 2008 [12]</td>
<td>I-chart</td>
<td>To review and analyse the PSQA results based on SPC methods and evaluate a new beam model with SPC for the TPS validation.</td>
<td>Control charts demonstrated that the HN IMRT dosimetry was stable over time.</td>
</tr>
<tr>
<td>Palaniswaamy et al., 2012 [14]</td>
<td>I-chart</td>
<td>To develop an SPC-based tool for monitoring PSQA results using MATLAB; To use this tool to analyse QA results for different sites and evaluate the possibility of site-specific tolerance limits for PSQA.</td>
<td>S-chart can be used to evaluate the variability in PSQA across sites.</td>
</tr>
<tr>
<td>J. Bellec et al., 2017. [17]</td>
<td>I-chart</td>
<td>To evaluate Cyberknife PSQA using SPC.</td>
<td>The TL for IMRT and VMAT for ( \chi ) - analysis were 62.5% and 70%, respectively.</td>
</tr>
<tr>
<td>Diana Binny et al., 2018. [18]</td>
<td>I-chart</td>
<td>To investigate the use of site-specific planning parameters by applying SPC tools on planning parameters and PSQA results for Tomotherapy.</td>
<td>The ANOVA test can be used to evaluate the need for site-specific tolerance limits, which can be calculated using the automatic MATLAB tool developed.</td>
</tr>
<tr>
<td>Teodor Tîplica et al., 2020 [19]</td>
<td>Beta distribution</td>
<td>To demonstrate that the Beta distribution defines the 2D-GIPR.</td>
<td>SPC-based analysis provided useful information for assessing changes in the beam model.</td>
</tr>
<tr>
<td>Marco Esposito et al., 2020 [21]</td>
<td>I-charts</td>
<td>To analyse and establish TL and AL for pelvic and abdominal stereotactic treatment PSQA.</td>
<td>Prostate PSQA results were more statistically stable than the HN PSQA results.</td>
</tr>
<tr>
<td>Yuqing Xia et al., 2020 [22]</td>
<td>I-charts</td>
<td>To apply and report the TG-218 recommendations for SRS and SBRT cases, performed with different detector systems.</td>
<td>Developed dedicated software package for analysing PSQA results using control charts and process capability analysis.</td>
</tr>
</tbody>
</table>

**continued on next page**
shifts in the mean, changes in variability, or the occurrence of unusual patterns. By detecting such variations early on, organizations can take timely corrective actions to prevent defects or quality issues from occurring. Overall, heuristic control charts provide a flexible and responsive approach to monitoring and improving processes.

2.2. Pareto charts

Pareto charts are one of the SPC tools extensively utilized for process improvement. The categorical arrangement of attributed data represented as frequency distribution is known as the Pareto charts. Pareto charts help the user visualise and point out the most common defects occurring quickly. It must be kept in mind that the priority is not given to the fatal error, which might decrease the process quality; instead, the error frequency is given priority. However, if the errors of higher consequence need to be prioritized, either a Weighted Pareto plot or an Exponential Pareto chart can be used.

2.3. Cause-and-effect diagram or fishbone diagram

A cause-and-effect diagram or the fishbone diagram can be used when the causes for a process state are unclear. Once a defect is defined, a team for cause-and-effect analysis is formed; this team is responsible for uncovering the potential reasons through brainstorming. Potential

<table>
<thead>
<tr>
<th>Author and Year</th>
<th>SPC tools used</th>
<th>Objective and QA analysed</th>
<th>Observations / Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vysakh et. al., 2021</td>
<td>• I-MR charts</td>
<td>• Retrospective evaluation of PSQA (275 plan QAs and 1214 field QA deliveries) performed using four different detector systems Delta4, Portal Dosimetry, ArcCHECK, and SRS MapCHECK.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Cause and effect diagram for evaluating failed PSQA and Linac Output constancy QA.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. Cui et.al., 2021</td>
<td>• I-MR charts</td>
<td>• To demonstrate a PSQA method by combining gamma and DVH based evaluation based SPC methods.</td>
<td>• Site-specific ALs were determined for PSQA. The combined DVH and Gamma evaluations added more insights to PSQA.</td>
</tr>
<tr>
<td></td>
<td>• EWMA charts.</td>
<td>• To implement site-specific tolerance limits and ALs for PSQA based on AAPM TG-218.</td>
<td>• Cause-and-effect analysis of failed PSQAs pointed out six primary potential sources of errors in the results.</td>
</tr>
<tr>
<td></td>
<td>• The TL and AL are calculated using TG-218 principles.</td>
<td>• Retrospective evaluation of 241 PSQA results in terms of both Gamma passing rates and DVH evaluation results</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• I-Charts</td>
<td>• To improve the TLs and ALs for PSQA using SPC methods.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• PCI</td>
<td>• Retrospective evaluation of 324 portal dosimetry PSQA results (GIPR)</td>
<td></td>
</tr>
<tr>
<td>Qing Xiao et al., 2021</td>
<td>• Box-Cox power transformation method</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Three heuristic methods and the quantile method for specific distributions.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robert A. Price Jr. et.</td>
<td>• I-charts</td>
<td>• To use the heuristic control charts for PSQA analysis.</td>
<td></td>
</tr>
<tr>
<td>et.al.,2022 [25]</td>
<td>• PCI</td>
<td>• To implement an iterative “Identify-Eliminate Recalculate” technique for determining TLs in PSQA processes with unknown states based on retrospective GIPR.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The Johnson transformation system</td>
<td>• Retrospective evaluation of 1671 VMAT PSQA results (GIPR measured with a 2D detector array).</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The TL and AL are calculated using TG-218 principles.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sarah Strand et. al.,</td>
<td>• I-chart</td>
<td>• To analyse the PSQA plans of Elekta Unity MR Linac using TG-218 protocol.</td>
<td></td>
</tr>
<tr>
<td>2021[26]</td>
<td>• PCI</td>
<td>• To compare the PSQA results between the reference and adaptive plans, SBRT plans and conventionally fractionated plans.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The TL and AL are calculated using TG-218 principles.</td>
<td>• Retrospective analysis of 512 PSQA measurements performed with MR-compatible 2D array. (Gamma evaluation)</td>
<td></td>
</tr>
<tr>
<td>Despoina Stasinou et. al.</td>
<td>• I-chart</td>
<td></td>
<td></td>
</tr>
<tr>
<td>et. al., 2022 [27]</td>
<td>• PCI</td>
<td>• To analyse the PSQA plans of Elekta Unity MR Linac using TG-218 protocol.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The TL and AL are calculated using TG-218 principles.</td>
<td>• To compare the PSQA results between the reference and adaptive plans, SBRT plans and conventionally fractionated plans.</td>
<td></td>
</tr>
<tr>
<td>G. Li et al., 2023</td>
<td>• I-chart</td>
<td>• Retrospective analysis of 100 HN and 73 Prostate PSQA results (2D and 3D GIPR).</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Three heuristic methods and the quantile method for specific distributions.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• To investigate the impact of sample variability on the performance of individual charts (i-charts) for PSQA and to develop a robust and reliable technique for unknown PSQA processes.</td>
<td></td>
</tr>
</tbody>
</table>

reasons are categorised and plotted in effect boxes connected with a centre line. The order for categories is ranked to identify those that seem most likely to impact the problem. Conventionally the cause-and-effect diagram accounts for the errors or effects of materials involved, methods used, personnel involved, and environmental factors such as temperature and pressure in the process.

2.4. Process capability indices (PCI)

In-process monitoring, it is essential to understand process variability over time within its tolerance limits and how centered the process is over the target value. The process capability indices address both requirements. Any process monitored using PCI requires the process to remain in a state of statistical control, which is ensured using control charts until the process achieves statistical stability. Two familiar PCI indices used in SPC are \( C_p \) and \( C_{pk} \) [7]. \( C_p \) accounts for the process variability relative to the 6 \( \sigma \) spread in the process (one-sided PCI), while \( C_{pk} \) accounts for the proximity of the process mean to the process target (two-sided). \( C_{pm} \) was introduced later, which can describe both the process variability with respect to \( AL \) and the location of the process mean with respect to the process target [7]. For a process with normally distributed data with standard deviation \( \sigma \), and process mean \( \mu \), the two-sided \( C_{pm} \) index is given as:

\[
C_{pm} = \frac{UAL - LAL}{A/\sqrt{\sigma^2 + (\mu - T)}}, \tag{15}
\]

The constant \( A \) represents the quality level required. A satisfactory process can be described by a \( C_{pm} \) index of 1.33.

3. Analysis of patient-specific quality assurance using SPC

Intensity Modulated Radiotherapy (IMRT) plan verification via PSQA is an important QA test performed in any Radiotherapy department. Over the years, SPC tools have been extensively used for analysing the PSQA results. A summary of published works which used SPC methods in analysing the PSQA is given in Table 1. Sanghanthum et al. [10] used Shewhart’s I-chart to ensure the process is ‘in a control’ state, and then the process capability index \( C_{pm} \) was used to estimate the AL. The recommendations by TG-218 [8] for process-based TL and AL for PSQA analysis were based on this study.

A two-stage approach is inferred in implementing SPC for PSQA analysis. In the first stage, the statistical stability of the PSQA process is monitored over a certain period using Shewhart’s I-chart. The TLs are usually calculated from these datasets using a minimum of 20–25 in-control data points (QA results) if the process is statistically stable. As mentioned in session 2.1, it is a user’s choice how to define TL for the monitored QA process. The CL obtained from the control charts solely depends upon the process variability as it is calculated from a chosen set of stable subgroups. Sanghanthum et al. [10] suggest using the CL of control charts used in stage one PSQA monitoring as TL for PSQA, in which the calculated TL acts as a sign that a process is drifting and requires attention. An ‘Identify-Eliminate-Recalculate’ approach is adopted in case of any process drifts during the initial stage. However, Sanghanthum et al. [29] suggest that those points whose reason for drifts are known and can be prevented on future occurrence only can be excluded during limit estimation. Though Shewhart’s I-charts are the commonly used control charts to monitor the process variability in PSQA, authors have used Exponentially Weighted Moving Average (EWMA) Control Charts to identify the minor process drifts [13,14,23]. Cui et al. [23] showed that the EWMA could mask the shortcomings of the I-chart in defining the minute process drifts, and they suggest adopting a combination of the I-chart and EWMA chart in analysing the PSQA process. The process capability indices are the other essential tool used along with the control charts in finding the AL. Authors have used various PCI such as \( C_{pp} \), \( C_{pko} \), \( C_{pm} \) and \( C_{pmk} \) in analysing the PSQA [11,13,16,18,20]. For determining the ALs, equation (15) can be rearranged to find the AL for PSQA as the equation below:

\[
\Delta AL = \beta \sqrt{\sigma^2 + (\mu - T)^2}, \tag{16}
\]

where \( T \) is the target value and \( \sigma \) and \( \mu \) are the standard deviation and mean of the chosen PSQA results. Usually, since the maximum gamma passing rate is 100%, the chosen value of \( T \) in equation (16) is 100% for \( \gamma \) analysis. The constant \( \beta \) is obtained by combining \( A \) and \( C_{pm} \) in equation (15), and Sanghanthum et al. [10] suggest using a value of 6 for obtaining AL for PSQA. When utilising PSQA data to decide on process performance, the value chosen for \( \beta \) balances type I errors (rejecting the true null hypothesis) and type II errors (accepting the false null hypothesis). The value for \( \beta \) in equation (16) was modified if the obtained TL exceeded the AL. Although various authors have used various control charts and approaches in monitoring PSQAs, the AAPM TG-218 [8] recommends using the methodologies of Sanghanthum et al. [10] in finding the limits for PSQA, and readers can adopt these methodologies in implementing SPC-based PSQA analysis.

One of the aspects that the user should keep in mind when monitoring PSQA results using control charts is the Average Run Length (ARL). ARL is the average number of points on a control chart that must be added before a signal is generated. The signal can either represent an in-control state (in-control ARL (ARL0)) or an out-of-control state (out-of-control ARL (ARL1)). ARL0 needs to be as large as possible, while ARL1 needs to be as small as possible. The inverse of ARL indicates the probability of a False Alarm Rate (FAR) for an observation when the process is in control. As defined by Montgomery [7], ARL determines the sampling frequency and sample size for the data from the process used in constructing the control charts. For a normal process with known parameters, the conventional 3-sigma Shewhart control chart exhibits an in-control performance of a FAR of 0.27% and ARL0 of 370.4 [7], which indicates the need for more than 300 data samples for ensuring the in-control performance of the Shewhart’s Individual value chart [28]. However, in practice, most practitioners use a dataset with a sample size of less than 30 to design an I-chart for the PSQA process. Li et al. [28] proved that the sampling variability significantly affects the performance of the I-chart, especially for low sample numbers, and the width of the 95% confidence interval of control limits tended to decrease with the increase in sample size. They suggest adopting a weighted standard deviation (WSD) approach for monitoring unknown PSQAs based on the “Identify-Eliminate-Recalculate” procedure [28].

The normality of the PSQA results is another essential aspect to be kept in mind while applying SPC methodologies to PSQA. In the case of skewed distributions, the control charts increase the risk of false positives, and process capability analysis could lead to wrong process performance interpretation [30,31]. According to Montgomery et al. [7] the X-R control charts are mostly robust to normality assumption, especially in Phase-I application; In the case of Shewhart’s I-chart, the non-normal data can reduce the ARL0 and the Phase-II performance. Also, the process capability indices utilised to calculate the PSQA limits strongly depend upon the normality of the data distribution. In cases where the data are not normally distributed, it may be necessary to transform the data or use a different control chart better suited for non-normal data [32,33]. Xiao et al. addressed the non-normality of data analysed using transformation, heuristic, and quantile methods based on the beta and optimally fitted distributions [20,24]. They recommended utilising heuristic and quantile approaches in clinical applications when sufficient PSQA results are available and when the process is in control, and for unspecified PSQA processes, they introduced an alternative approach to the ‘Identify-Eliminate-Recalculate’ method for determining TL and AL for PSQA analysis and recommended using this method in conjunction with the Heuristic approaches [24]. Although data transformation may be permissible, if a data point becomes out of control after transformation, it can be difficult for the user to pinpoint
the cause of the problem because it is the transformed data that is out of control rather than the original data.

Various authors [11,22,23,25–27] have adopted TG-218 methodologies for finding TL and AL for PSQA. The locally found AL and TL are compared with TG-218 recommended Universal TL (TL\text{uni}) and Universal AL (AL\text{uni}) of 95% and 90%, respectively. According to Xia et al., more stringent gamma criteria and higher AL and TL can be used for the stereotactic treatments [22]. The locally calculated AL and TL for individual structures can differ from the TG-218 recommended limits [11]. More studies are required to confirm whether the TG-218 recommended AL\text{uni} and TL\text{uni} for PSQA can be used for individual structure-based gamma analysis, as the percentage gamma passing points may vary abruptly among each structure.

4. SPC for routine accelerator QA analysis

Modern Radiotherapy QA seeks to ensure not just that parameter or process under evaluation is on target but also that it functions with low minimal variability [5]. SPC tools, such as control charts, evaluate the overall QA quality and help to take preventive, corrective measures. The AAPM TG-142 has classified the Linac QA frequency as daily, monthly, and annual, depending on the parameters observed, the type of treatment, and the machine manufacturers [34]. SPC methods can be effectively utilized in conjunction with the TG-142 recommendations for defining Institution and machine-specific and TL and AL for routine accelerator QA. Different authors have used control charts to plot time-ordered QA data and to infer conclusions regarding Linac behaviour over time. A summary of the works which used SPC tools in analysing the routine accelerator QA is given in Table 2.

The I-MR charts (or the X̄) can be used to monitor process drift in the order of 3σ, while the time-weighted control charts are sensitive to process drifts in the order of 1σ [7]. Time Weighted control charts are used to address the challenge of identifying subtle shifts in QA results over longer time periods. Thus, for a two-stage SPC-based analysis, the I-MR charts can be used for stage-1 monitoring, and after gaining enough confidence in the QA process, the time-weighted control charts can be used for further refinement. Most authors [4,36–38,40–43] used I-MR charts to analyse and define the limits of the Linac QA results. However, time-weighted control charts such as EWMA and CUSUM were also used to detect slow process drift or analyse QA processes with subtle variations between subgroups [29,35,39]. The parameters for the time-weighted control charts, such as the λ and L for EWMA charts, and H and K for CUSUM chart, can be modified depending on the magnitude of the process to be identified [7]. Since the X̄ charts are more effective in quickly detecting small process drifts than I-charts; authors have used X̄ charts with specific subgroup sizes for routine QA analysis [42]. The variation within the subgroup decides the sensitivity of the control charts; Thus, ideally, the largest practical subgroup size should be used. However, large subgroups with non-homogeneous data will mask the ability of control charts to detect a signal (i.e., widen the control limits), which eventually makes the control charts less sensitive in detecting process drifts. According to Pawlicki et al. [4], the subgroup size should be decided based on the process and the behaviour of the data originating from the process. The number of measurement points to be used for TL estimation is critical in process monitoring as too narrow TL will result in significant False positive cases, while larger TL would result in missing process drift identification. The selection of cases for TL estimation differed with studies, Sanghangthum et al. [29] suggest using 8–12 QA results to set the TL and to recalculate the limits in any case of deliberate changes in the process.

The physicists must ensure that the monitored system parameter or behaviour operates in a statistical control state and bring back the process within the TL if any process goes beyond the limits. The Fishbone diagram or Ishikawa diagram is an SPC tool for finding the causes of abnormal process behaviour. Pal et al. [39] demonstrated the use of the Ishikawa diagram for root cause analysis of failures in routine Linac QA. Vysakh et al. [11] demonstrated the use of these diagrams for analysing the possible reasons for PSQA failures and deviations in Linac output constancy.

5. SPC for analysing radiotherapy patient positioning accuracy

The introduction of new imaging modalities and treatment protocols has noted a significant improvement in patient positioning. SPC methods can be adopted to implement precise Image Guided Radiotherapy (IGRT) procedures and to analyse the efficacy of IGRT procedures for fractionated Radiotherapy [44–49]. Ung and Wee [44] used X̄ control charts to retrospectively analyse the fiducial markers matching to understand the potential problems with rigid-body matching algorithms for patients undergoing prostate Radiotherapy. They suggest using more than five days of registration information to estimate the limits for further positional verification. In a similar study [45], they adopted a non-zero fixed action level setup correction technique using CUSUM charts for analysing prostate IGRT, which proved effective for estimating and correcting systematic errors in patient setup. Moore et al. [47] used a combination of I-chart and EWMA charts to analyse the improvement in patient positioning for Head and Neck (HN) Radiotherapy over time. The intentional changes they made to improve the setup reproducibility were reflected in I-charts and EWMA charts. They used EWMA correlation plots between the EWMA value of systematic (EWMA\text{sys}) and random components (EWMA\text{rand}) for differentiating systematic and random changes, which proved to effectively analyse individual patient positioning accuracy by segregating systematic and random error components in patient positioning, from which objective decisions can be taken on whether treatment adaptation is required or not. Lowther et al. [48] utilized deformable image registration and EWMA charts to monitor the efficiency of adaptive HN Radiotherapy. They used deformable image registration between the sequential cone-beam CT images with planning CT images to verify the target volume change during fractions. Detection of systematic anatomical changes before the end of four weeks of treatment was possible due to structure-specific action tolerances derived using SPC charts. Satomi Shiraishi et al. [44] used patient specific control charts (X̄ charts and S-charts) for analysing the efficacy of fiducial matching in bilateral HN cases. They used two image similarity matrices (Normalized Cross-Correlation (NCC) and Mutual Information (MI)) as quality tracking metrics of the image registration data. The patient specific control charts revealed the variable setup consistency between the patients, with variable TLs. They proved that the current qualitative quality control review process for image registration is inadequate and suggest using similarity metrics based on patient specific TLs.

Because of its noninvasiveness, real-time imaging capacity, and radiation-free properties, Surface-Guided Radiation Treatment (SGRT) has received great interest in modern Radiotherapy. Li et al. [49] used SPC tools to assess the performance of the SGRT system. They monitored the process using a combination of the I-charts and EWMA charts and found a systematic error (~1.6 mm) between the SGRT and the cone-beam CT systems. They explained that this could be due to the fundamental difference between the two system’s image acquisition and matching. The systematic error was eradicated by acquiring SGRT and CBCT images during the first five fractions for every patient, then applying the determined shifts to all subsequent fractions in which only the SGRT system was used for patient positioning. Using process capability evaluation, they found that the performance of the SGRT system was allowable (C\text{pk} = 1.0) for HN patients but not acceptable for breast cancer patients (C\text{pk} < 0.75).

Most studies utilized the initial five days of image registration information for finding patient-specific control charts with customised control limits. The simultaneous usage of Shewhart’s control charts and time-weighted control charts for analysing patient positioning errors
was more effective than using only one type of control chart [47,49]. Most of the studies mentioned above utilised SPC methods for retrospectively reviewing the IGRT procedures. The logistic issues associated with the image review process might have prevented a prospective SPC-based image review process. This can be made possible by either embedding the SPC tools into the commercially available image review modules such as ARIA™ and Mosaïq™ or by introducing specially written SPC software packages into clinical use.

6. Other significant works which utilised SPC for Radiotherapy QA

The chain of processes involved in Radiotherapy is always subject to variability, which might adversely affect the treatment outcome. The contouring and planning are always subject to intra-practitioner and inter-practitioner variability [50,51]. SPC methods were utilised to assess plan quality and variability across clinicians in contouring and physicists in planning [52–54]. Pitkänen et al. [52] and Holli et al. [53] used $X - R$ charts to show the reproducibility and variability across clinicians in breast radiotherapy contouring and planning. Roy et al. [54] developed a risk-adjusted control chart to assess the plan quality. The risk-adjusted chart was built using eighteen risk factors and was demonstrated in 69 HN cases. The risk-adjusted charts were found to be more robust to patient and treatment-specific factors such as inter-patient variations in anatomy, tumour dose prescription, use of chemotherapy, and surgery than the conventional I-MR charts [54]. The process capability analysis of the 2D gamma index was utilised to compare and rank treatment plans for Lung radiotherapy by Chaikh et al. [55]. They showed that $C_p$ and $C_{pk}$ indices could be used as complementary information for judging the need for adaptive Lung Radiotherapy. Breen and Zhang [56] used Pareto charts to analyse the reasons for changes in HN IMRT plans. They added an automated checklist to the treatment planning process for HN IMRT and studied how it affects the outcome.

SPC methods can be used for forecasting system performance and avoiding downtime. Able et al. [57] used SPC to analyse retrospective Linac beam steering parameters to understand Linac behaviour and avoid downtime. They claimed that the use of SPC could have initiated

| Table 2 |
| A summary of works which utilised SPC methods for analysing Routine Linac QA |

<table>
<thead>
<tr>
<th>Author and Year</th>
<th>SPC tools used</th>
<th>QA Parameter analysed</th>
<th>Additional Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Todd Pawlicki et. al, 2005 [4]</td>
<td>$X$-chart</td>
<td>Linac output constancy</td>
<td>Subgroup averages are plotted in I-chart</td>
</tr>
<tr>
<td>Tawseep Sanghanthum et. al, 2012 [29]</td>
<td>$X$-chart, EWMA chart, PCI</td>
<td>Linac output constancy</td>
<td>The ARL for output constancy was found as</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>o $X$-chart: 8-12 points.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>o EWMA chart: 4 points.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The appropriate EWMA parameters for output constancy monitoring:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\lambda = 0.10, L = 2.703$.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Need 25–30 in-control data points for $C_p$ and $C_{pk}$ estimation.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Kolmogorove - Smirnov test for checking the normality.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Reference value, $k$ for CUSUM chart = 0.5.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Decision interval, $H$ for CUSUM chart = 5.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>For MR chart, $n = 4$ and ARL = 8.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Kolmogorove-Smirnov test for checking the normality.</td>
</tr>
<tr>
<td>Josué Manuel de la Vega et. al, 2012 [35]</td>
<td>CUSUM Chart, MR-Chart</td>
<td>Electron beam quality index ($F_q$)</td>
<td>For all control charts, the subgroup size ($n$) was chosen 6, and initial 7 subgroups were used to calculate the limits.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Kolmogorove-Smirnov test for checking the normality.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>All available data points were considered to find the TL in the I-MR chart.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Anderson-Darling test for checking Data Normality.</td>
</tr>
<tr>
<td>Kwang-Ho Cheong et. al, 2015 [37]</td>
<td>I-MR chart, PCI</td>
<td>Linac Log file analysis</td>
<td>For all control charts, the subgroup size ($n$) was chosen 6, and initial 7 subgroups were used to calculate the limits.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Kolmogorove-Smirnov test for checking the normality.</td>
</tr>
<tr>
<td>Diana Binny et. al, 2017 [38]</td>
<td>$X$-chart, PCI</td>
<td>Routine Linac QA</td>
<td>For all control charts, the subgroup size ($n$) was chosen 6, and initial 7 subgroups were used to calculate the limits.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Kolmogorove-Smirnov test for checking the normality.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Kolmogorove-Smirnov test for checking the normality.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Kolmogorove-Smirnov test for checking the normality.</td>
</tr>
<tr>
<td>Sandra M. Meyers et. al, 2019 [40]</td>
<td>I-MR chart</td>
<td>MLC QA results</td>
<td>The stipulated MLC performance test limit of ±1mm is sufficient to assess the MLC positional uncertainties.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Subgroup size ($n$) = 1.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Initial 85% data were used to build the control chart and create the model; the remaining 15% of data was used to validate the model.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Initial 30–40 points were used to calculate the Limits.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The AIs are calculated for different A values ($A = 4, 5, 6$).</td>
</tr>
</tbody>
</table>

an evaluation of the beam steering mechanism before a significant loss in beam flatness and symmetry, increasing the likelihood of detecting the water leak responsible for machine performance changes. Like this, Schlesinger et al. [58] employed log file analysis with SPC for the Gamma Knife radiosurgery equipment to forecast behaviour and reduce downtime. They made in-house software to extract the difference between the position of the planned and delivered patient positioning system, and retrospectively 8-years of data were extracted and analysed via EWMA charts.

SPC tools were also used for monitoring Brachytherapy QA results [59] and Proton therapy QA [60]. Using I-MR control charts, Rah et al. [60] analysed the daily QA results of proton machines, such as the beam range and the Dose per MU. PCI such as Cpm, Cpm and Ccpck were used to analyse the PSQA results on a site-specific basis. In clinical proton plans, their analysis utilising process capacity indices revealed that patient-specific range measurements were feasible at a specification limit of 2%. They concluded that the SPC techniques could aid in the customisation of AI to improve the accuracy of treatment delivery in proton beams. Able et al. [59] utilised X and MR charts to monitor the simulated high-dose-rate (HDR) Brachytherapy prostate treatment in a surrogate phantom. They simulated different treatment scenarios with added errors, which were found to have maximum dosimetric effects at the periphery and apex parts of the prostate. They conclude that SPC is a promising tool for evaluating HDR Brachytherapy QA.

7. Conclusions

SPC methods can be employed to monitor various Radiotherapy QAs and might potentially replace traditional QA approaches. Proper selection and having adequate knowledge of SPC tools is essential for proper QA monitoring, which is a function of the type of QA data available, and the magnitude of process drift to be monitored.

Machine-specific, site-specific and technique-specific TLs and ALs could be adopted for better PSQA monitoring. A two-stage approach is implemented in implementing SPC for PSQA analysis. In the first stage, the statistical stability of the PSQA process is monitored over a certain period using Shewhart’s I-chart. Then once the process is said to be in control, the TLs and ALs are calculated using control chart information and process capability indices, and the process is monitored further. The user should be aware of the sampling variability of I-charts before using them on PSQA results, especially for small sample sizes. The normality of the data distribution can affect the PSQA analysis, which can be tackled by transforming the data using a different control chart better suited for non-normal data.

Analysis of Routine Linac QA using SPC could help the user to segregate the ‘special’ and ‘routine’ causes of fluctuations in the QA process. An approach combining Shewhart’s control charts and time-weighted control charts for QA monitoring could add more insights to routine Linac QA. Incorporating SPC tools into existing image review modules or introducing new SPC software packages specifically designed for clinical use has the potential to significantly enhance the image review process.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors are thankful to Mr Sreejith P K for his helping hand in writing this article.

References

V. Raveendran et al.


