The Use of Scan Statistics and Control Charts in Assessing Ventilator-Associated Pneumonia Quality Control Programs

Brian H. Nathanson[1]*, Ph.D. and Thomas L. Higgins[2], MD, MBA
[1]OptiStatim, LLC, Longmeadow, MA
[2]Baystate Medical Center, Springfield, MA, Tufts University School of Medicine, Boston, MA

ABSTRACT

Scan statistics are concerned with clusters of events over time. In the realm of critical care medicine, such clusters might include the occurrence of ventilator-associated pneumonia (VAP). Given N patients over time, the number of observations in a "moving window" of fixed length can be counted and the maximum cluster value becomes a scan statistic for which both parametric and exact methods exist to calculate its rarity. A statistically unusual cluster may indicate a breakdown in quality. Another approach to monitoring rare events is a g-type statistical process control chart where prospectively observing unusually long periods of time between events can indicate a significant improvement in quality. Both methods are presented in detail and applied to a 24-bed medical/surgical ICU's experience with VAP during a 27-month period.

Keywords: Ventilator Associated Pneumonia, scan statistic, ratchet scan statistic, g-type control chart

1. INTRODUCTION

Respiratory failure is a common, resource intensive comorbidity among patients in the intensive care unit (ICU). Over 25% of adult, non-cardiac surgical patients require mechanical ventilation within 1 hour of ICU admission and 36% of all ICU patients require mechanical ventilation at some time during their stay [1, 2]. When intubated patients are on a mechanical ventilator for more than 48 hours, up to 20% can develop Ventilator-Associated Pneumonia (VAP), a potentially deadly condition which occurs in 100,000 to 300,000 patients per year in the United States [3, 4, 5]. VAP is also associated with an increase in costs, length of stay, and over 50% of all antibiotic use in the ICU [5, 6, 7]. Consequently, there has been a concerted effort to implement procedures that would reduce VAP [8, 9], and since VAP may be preventable, a hospital's VAP rate is increasingly being used as a measure of hospital quality of care [5, 10, 11]. Yet, the correlation between recommended practices or "safety bundles" of care and preventing VAP is unclear [5] and the use of a hospital's VAP rate as a quality

*Corresponding author: OptiStatim, LLC, P.0. Box 60844, Longmeadow, MA
brian.h.nathanson@att.net

measure is controversial [10, 11]. Despite this, the Centers for Medicare & Medicaid Services recently proposed withholding reimbursement for hospital costs associated with treating VAP [5].

Regardless of how different government agencies or insurers eventually decide to structure payments in patients with VAP, clinicians and hospital managers will want to assess their effectiveness in its prevention. Additionally, if there is an increase in the number of VAP events during a particular period of time, they need to know how likely that occurrence can happen by chance alone since an unusual cluster may indicate a breakdown in a safety protocol or an outbreak of a more virulent strain of bacteria. The purpose of this study is to illustrate some new statistical methods that can be used for a VAP quality control program. Examples will be provided using data taken from a 24-bed medical/surgical closed ICU at a large tertiary care, teaching hospital in the urban northeast United States during January, 2006 to April, 2008. This information was collected as part of routine quality improvement efforts, and did not have identifiable patient data, and was thus not considered human subjects research requiring IRB approval.

2. THE DIAGNOSIS OF VAP

Currently, there is no single, unequivocal marker of VAP. Instead, physicians rely on several clinical signs such as the presence of a fever, increased pulmonary secretions, abnormal leukocyte counts, radiographic opacities, and cultures of pulmonary secretions to make the diagnosis. Prior studies have shown that diagnoses based on clinical signs alone leads to over-diagnosis while diagnoses based on bronchoalveolar lavage culture results alone may result in under-diagnosis [10, 12].   Further complicating matters is that other commonly encountered conditions in the ICU such as septicemia with pulmonary edema [10] can mimic VAP.

To accurately assess an ICU's VAP rate over time, the diagnosis of VAP must be consistent. To standardize the diagnosis, the Center for Disease Control's National Healthcare Safety Network (NHSN) (formerly known as the National Nosocomial Infection Surveillance System (NNIS)) has published a diagnostic algorithm (see Appendix A) [13]. There is also a validated Clinical Pulmonary Infection Score (CPIS) for which a score above 6 is consistent with a diagnosis of pneumonia (see Appendix B [3,14,15]). In our study's ICU, one specialized infection control nurse (with one back up nurse) has the job of determining VAP among ventilated patients based on the NHSN criteria. Since only one person is making the interpretation, the VAP rate is not subject to interrater variability which may be a problem in other ICUs.

Much of the literature cites VAP rates as cases per 1,000 ventilator days, rather than incidence, but for the purposes of cluster analysis, we will focus on discrete events. Also, since data to this study were not tied to specific patients, rates per 1,000 vent days must be estimated based on population, not individual data.

3. ANALYZING A CLUSTER OF VAP EVENTS

Assuming that there is a consistent and accurate diagnosis of VAP over a period of time in an ICU, a major question is how to decide whether a particular cluster of events is unusual. For example, in one 3-month (91-day) period (September through November

of 2006), there were 12 cases of VAP in this study's ICU. During this time, 278 patients received 1936 days of mechanical ventilation, implying a VAP rate of 6.2 VAPs/1000 vent days, well above the NHSN benchmark 50th and even 75th percentile at the time, and higher than the ICU's historic rate. In addition, 4 of these 12 cases occurred in the first week of October (See Figure 1). An analysis comparing 4 cases in 7 days (4/7) vs. 8 cases in the remaining 84 days (8/84) using a chi-square (χ^2) test or Fisher's exact test (a non-parametric version due to the sample size) reveals a p-value < 0.01.
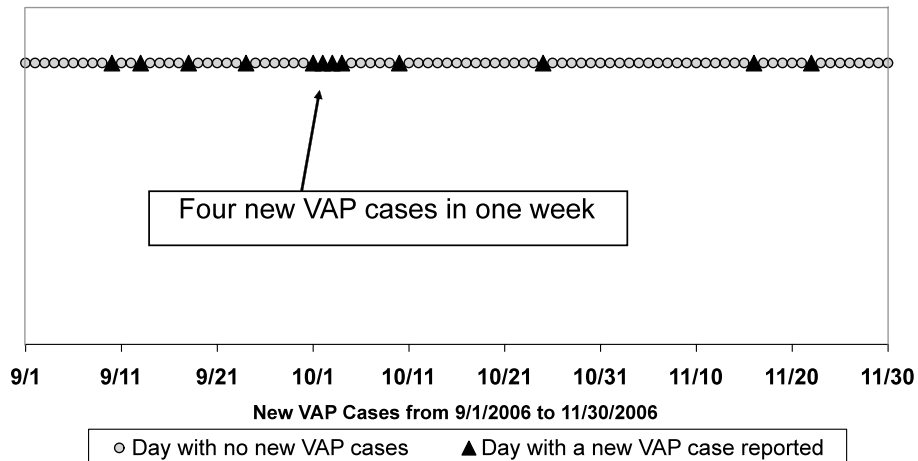

Figure 1.   A graphical representation of new VAP cases from 9/1/2006 to 11/30/2006

These 4 cases happened to be consecutive (which was not relevant to the prior analysis) and if we view the chance of a VAP case to be 13% on any one day as there were 12 cases in 91 days, then the probability of 4 consecutive cases is less than 0.03%, implying a seemingly more unusual occurrence.

However, the chi-square analysis only showed two arbitrarily defined proportions were different. That does not prove that the clustering was unusual. In fact, one can always "cherry pick" a particular sequence in a long run of observations and find clusters of events that appear unusual on the surface. One simply has to imagine the results of flipping a fair coin over time and realize that an occasional long run of consecutive heads or tails is more "natural" than a consistent alternating sequence of heads, tails, heads, tails. In our scenario, the appropriate question to ask is: What's the chance of observing 4 new VAP cases in *any* 7 day period (i.e., days 1 to 7, then days 2 to 8, then days 3 to 9, etc.) given that there were 12 cases in 91 days? To answer this question, we would need to rely on scan statistics.

3.1. Scan Statistics: The Discrete Case

Given N events distributed over a period of time, we can define S_w as the largest number of events in a window of fixed time of length w' observed retrospectively. In our analysis, w' will be normalized, leading to the definition of w as the ratio of the time length of interest divided by the total time observed. In our example above with 7 days out of 91, w = 7/91 or ≈ 0.077. S_w is called a scan statistic because one scans a series of time intervals (e.g., days 1 to 7, then days 2 to 8, etc.) to find the value of S_w. We can define W_k to be the smallest interval of time that contains k events. The distributions of S_w and W_k are related since $P(S_w \geq k) = P(W_k \leq w)$. As per convention, we write the common probability with the notation: P(k; N, w) [16].

Historically, exact formulas have been known for two special cases: when all events happened within time window w and when two events happened within time window w. Their formulas are given in Equations 1 and 2, respectively [16]:

$$P(N; N, w) = Nw^{N-1} - (N-1)w^N \text{ , for } 0 < w < 1 \tag{1}$$

$$
\begin{aligned}
P(2; N, w) &= 1 - [1-(N-1)w]^N \text{ , for } 0 < w < 1/(N-1) \\
&= 1, \text{ for } 1/(N-1) < w < 1
\end{aligned}
\tag{2}
$$

More recently, Wallenstein and Neff derived an approximate formula (eqn. 3) for P(k; N, w) [16]:

$$P(k; N, w) \approx (N-k+1) \, b(k-1; N, w) - (N - k - 1) \, b(k;N,w) + 2G_b(k; N, w) \tag{3}$$

where

$$b(k;N,w) = \left(\begin{array}{c} N \\ k \end{array} \right) w^k \left(1 - w\right)^{N-k} \tag{4}$$

and

$$G_b(k;N,w) = \sum_{i=k}^{N} b(i;N,w) \tag{5}$$

The approximation formula is accurate when the derived P(k;N,w) is low (typically < 0.1) and is exact when k > N/2 and w ≤ 0.5 [16]. There are also published tables that are based on similar algorithms. For example, Neff and Naus derived probabilities for a range of k's and N's at different w's [17].

For our study's ICU, k = 4 events out of N = 12, in a window of w = 0.077 (7/91). Equation 3 gives us a P(4; 12, 0.077) value of 0.38 while the tables of Neff and Naus give a value of 0.36 [17], and so we can conclude with either inference result that this type of clustering is definitely not unusual.

Table 1 provides *p*-values adapted from the tables of Neff and Naus [17] for some common values of N and k at a time period of w = 0.08 and highlight in boldface cases where the *p*-value ≤ 0.05. For example, if we use 0.08 as our window length w, Table

1 shows that the *p*-value for 4 events out of 12 is 0.38887. Clinicians should find Table 1 with a length w of 0.08 useful for their own ICUs as it is approximately 7 days out of 3 months, 1 month out of one year, or 2 weeks out of 6 months, etc.

Table 1. P-values for the retrospective case of P(k,N,0.08) [17]‡

w = 0.08				
N	k = 3	k = 4	k = 5	k = 6	k = 7
3	**0.01818**				
4	0.06509	**0.00193**			
5	0.14556	**0.00867**	**0.00019**		
6	0.25772	**0.02344**	**0.00104**	0.00002	
7	0.39291	**0.04924**	**0.00329**	0.00012	0.00000
8	0.53720	0.08852	**0.00794**	0.00042	0.00001
9	0.67470	0.14277	**0.01616**	0.00115	0.00005
10	0.79168	0.21214	**0.02920**	0.00261	0.00015
11	0.88011	0.29520	**0.04833**	0.00520	0.00038
12	0.93897	0.38887	0.07472	**0.00942**	0.00084
15	0.99688	0.68453	0.20501	**0.03791**	0.00510
18	1.00000	0.89767	0.40830	0.10373	**0.01887**

3.2. Assessing a Seasonal Trend in the VAP Rate

Implicit in the prior analyses is that the incidence of VAP is random over time. Though this is often a reasonable simplifying assumption, it may not be true in a specific ICU. Rello et al. [18] found in a large U.S. database that trauma patients were at a higher risk for developing VAP. This is true in part because trauma patients are more likely to aspirate during their injury and a lung infection from this aspiration could appear 3 to 4 days later (presumably when the patient is still on a ventilator). There is seasonality to the presentation of trauma patients not only due to weather conditions, but also due to available hours of daylight [19]. Thus, a rise in an ICU's VAP rate may not indicate a breakdown in quality but merely a seasonal change in patient case-mix.

A common mistake to determine if a seasonal trend exists in the data is to use a chi-square test with each month (or quarter, etc.) serving as the "bins" for the counts of VAP. The chi-square test is the wrong test to use as the 12! possible different ways to order the monthly counts would not affect the *p*-value of the test. Instead, a helpful first step would be to order the counts on a circular plot rather than a rectangular histogram to visually observe temporal trends. The circular plot is needed to help discern if a spike in VAP rate occurs during the winter as it does not split December from January. (Figure 2)

There are several inference tests in the literature designed to detect seasonal trends. The first major improvement on the Chi-Square test was made by Edwards in 1961 [20]. The Edwards test orders data sequentially over the unit circle's rim and assuming there are monthly data, creates 12 sectors. The counts in each month become weights

---

‡Adapted with permission from the American Mathematical Society from material originally published in Neff N, Naus J. *Selected Tables in Mathematical Statistics*, Vol. 6: *The Distribution of the Size of the Maximum Cluster of Points on a Line*. American Mathematical Society, Providence, RI, copyright 1980.

and a weighted center of gravity is calculated using trigonometry. If the center of gravity is skewed away from the origin in the x-y plane, the null hypothesis is rejected and a seasonal trend is assumed [20]. Unfortunately, the test is sensitive to extreme values; it lacks power, and assumes a sinusoidal trend in the data. Walter and Elwood improved upon the Edwards test by allowing for the size of the population at risk to vary during the time in question, though otherwise it is similarly flawed [21]. Contemporaneously, Hewitt developed a simple, non-parametric test that had more power than the Edwards test but was designed for a 6 month cyclical trend and later Rogerson devised a more general version of Hewitt's method for 3, 4, or 5 month periods [22]. Gao et al. recently proposed a more complex methodology based on angular (i.e., trigonometric) regression and made use of the von Mises distribution as they examine event data on the unit circle [23]. Though this method and its extensions may have promise, they have acknowledged technical difficulties in its implementation [23].
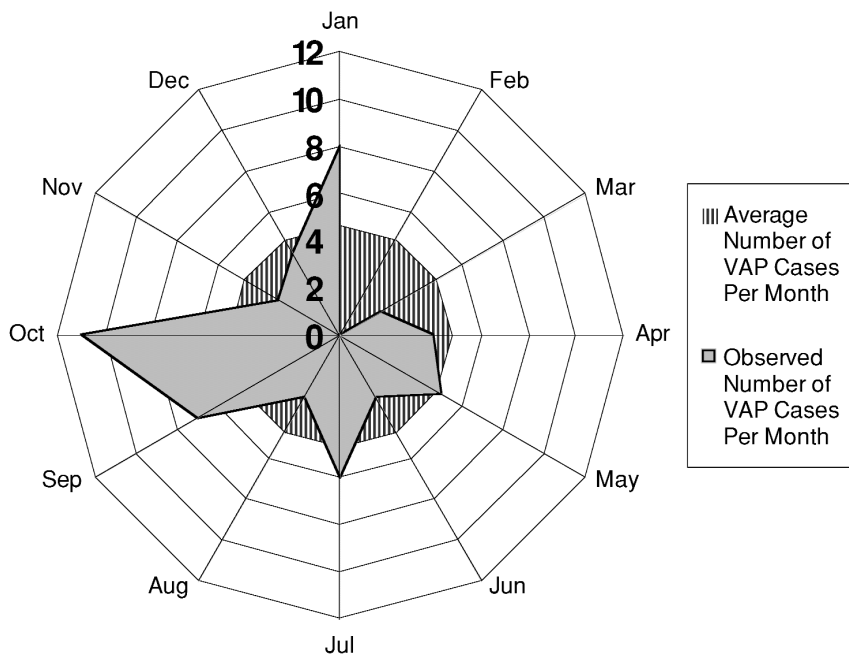

Figure 2.   Two years of new VAP cases by month from 2006 to 2007

Wallenstein et al. also expanded upon Hewitt's method to create what is called a *ratchet* scan statistic [16, 24]. The algorithm for deriving the ratchet scan statistic R is as follows (see Equation 4). To detect a three month seasonal trend (M =3) in yearly data when counts for each month are known, first find the "moving" 3-month maximum total in all 3-month "groupings" (i.e., find the number of VAP cases during {Jan, Feb, Mar}, {Feb, Mar, Apr},…, {Dec, Jan, Feb}) and denote the largest 3-month sum as R_{Max}.

Let w = 0.25 as 3 months is 25% (3/12) of the year, and N = the grand total number of events, then the ratchet scan static R is:

$$R = (R_{Max} - 1 - Nw)/\sqrt{Nw(1-w)} \qquad (6)$$

Figure 2 shows a circular plot of our study's VAP rate for each month over two years (e.g., January = 8 as there were 8 total cases in the two months of January during years 2006 and 2007). We do see an above average spike starting in September and peaking in October but curiously, August, November, and December had below average VAP counts. Is a seasonal trend present? The Edwards test gives a *p*-value of 0.011 on this dataset though as previously noted, it is sensitive to extreme values. The 3-month ratchet scan statistic is 1.83. Using Table 2 that provides critical thresholds for this statistic [16], we see that α falls between 0.5 and 0.1, which indicates no statistical significance at the traditional α of 0.05. If we calculated a one month (M=1) ratchet scan statistic (i.e., the maximum multinomial), we would get an R = 2.58 and see that it is somewhat unusual to have a month like October with 11 VAP cases, but again the statistic is between the thresholds of α = 0.1 and 0.05.

**Table 2. Critical values for the ratchet scan statistic for 1, 2,
and 3 month periods [16]#**

| α	M = 1	M =2	M = 3
0.5	1.66	1.58	1.52
0.1	2.38	2.34	2.32
0.05	2.63	2.62	2.58

3.3. Prospective Scanning and the Use of Control Charts
The statistics that have been examined so far were retrospective. To derive probabilities for prospective events (e.g., the chance that there will be a certain number of events in fixed time period in the future), different unconditional scan statistics based on the Poisson process would be needed [16]. Unfortunately, the computation of prospective scan statistic probabilities is more complex than the retrospective discrete case, though exact probability formulas are available and some limited results have been published in tabular form [16, 17, 25].

From a clinical perspective, there are alternatives to a prospective scan statistic per se, that are more useful when assessing an ICU's VAP rate in an on-going manner. One

---

#Reprinted with kind permission of Springer Science and Business Media from Glaz J, Naus J, Wallenstein S. *Scan Statistics*. Springer-Verlag, New York, NY. copyright 2001, page 330.

of the simplest types of benchmarking over time is illustrated in Figure 3 which is an approximation of the quarterly data from the study's ICU compared to published national VAP rates from the NHSN. Though lacking in statistical rigor, the figure shows that after a series of safety bundles (e.g., Head of Bed Elevation, blood glucose control, etc.) were implemented in the first quarter of 2005, the VAP rate declined. The figure also shows that by the start of 2008, this ICU was performing well as its VAP rate was lower than the NHSN's 25th percentile.
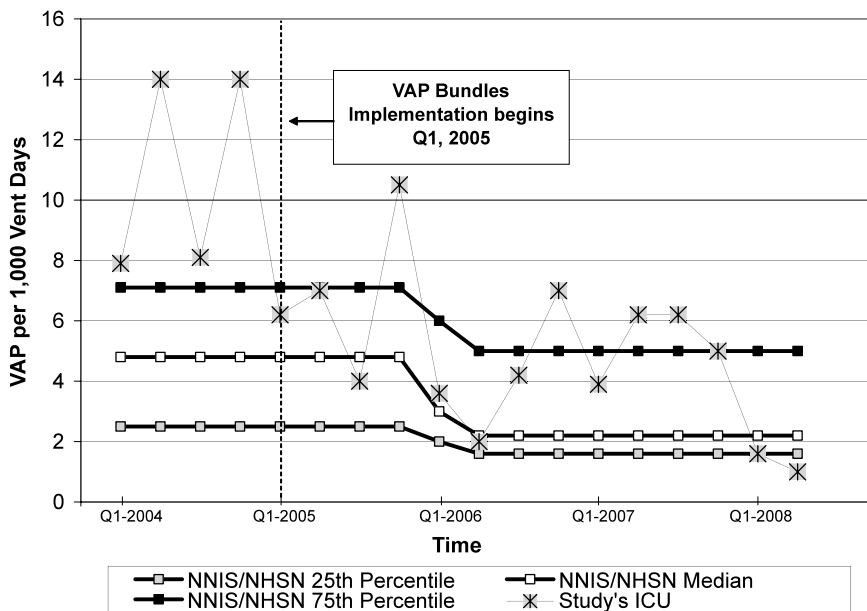
The vertical dashed line marks the implementation of VAP safety bundles in the study's ICU.

Figure 3.   Quarterly ventilator-associated pneumonia (VAP) rates for the study ICU from January 2004 to June 2008

Another prospective quality control approach for an ICU is statistical process control (SPC) charts. Because VAP events are uncommon, a g-chart should be considered as it is a relatively new SPC chart ideally suited for rare events [26, 27, 28]. In SPC, instead of monitoring the number of events like more traditional SPC charts, a g-chart plots the number of days *between* VAP events. The Center Line, Upper Control Limit, and Lower Control Limit for the g-chart are derived as follows [26, 27, 28]:

Center Line (CL) $= \bar{X}$ (average number of days between VAP events)

Upper Control Limit (UCL) $= \bar{X} + 3\sqrt{\bar{X}(\bar{X}+1)}$

Lower Control Limit (LCL) $= \bar{X} - 3\sqrt{\bar{X}(\bar{X}+1)}$, or 0 if $\bar{X} - 3\sqrt{\bar{X}(\bar{X}+1)} < 0$

Note that in a g-chart, as the process improves, the number of days between events would increase, and the data will move above the Center Line. In our study's ICU, the

average number of days between events was 14.0 days. The UCL formula above then yields a value of 57.5 days. Because the LCL value is negative, we set it to 0 as per convention. The downside to doing this is to lose the ability to detect a rate increase. Benneyan provides several ways to address a negative LCL (e.g., use a value smaller than 3) which can be useful [26,27]. As a caveat, the UCL and LCL are based on an approximation to the normal distribution which may not always be appropriate to a particular dataset. The g-chart for the study's ICU is presented in Figure 4. In our example, the first VAP event occurred at 1/6/06 and the y-coordinate of this point is at 6 as it is 6 days from the start of the study period. The second event occurred at 1/30/06 and is plotted as the second point from the left (y = 24) and due to space constraints this date and some others are not shown. On 1/30/06 another event also occurred and so the third point has its y coordinate at 0.
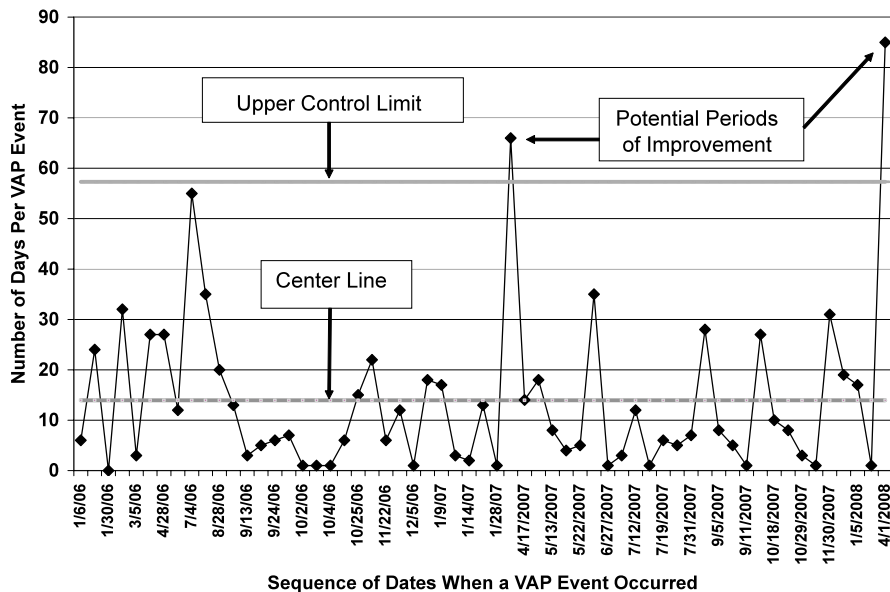

Figure 4.   G-Chart for number of days between ventilator associated pneumonia (VAP) events

With a g-chart in place, accepted SPC "rules" can then be applied to determine unusual patterns indicating deteriorating or improving quality over time. For example, having 8 or more consecutive values below or above the CL is significant, as is 12 of 14 consecutive values on the same side of the CL or observing a value (i.e., a time period with no VAP events) more than 3 times the CL.

An interesting example of a g-chart application can be found in the work of Wall et al. studying catheter infections in the ICU [29]. They constructed a g-chart before and after a quality initiative was implemented. They observed no values above the baseline

UCL during the pre-initiative period but several values above the UCL during the post-initiative period, and so concluded that the initiative had a significant impact on quality. Following this line of reasoning, we observe two periods in Figure 4 where there was a potentially significant improvement in VAP rates as they were both above the UCL, though these may also have been periods where some intermittent special cause variation entered or left the ICU.

Both of these "well-performing" episodes happened in the spring (and a poor performing episode happened in the fall of 2006). This ICU had essentially the same personnel of intensivist physicians, nurses, and respiratory therapists, and compliance with established safety protocols for pneumonia patients was consistently close to 100%. Explanation for the consistent decline in VAP over time and during the spring in particular is unknown, though we did observe differences in case-mix of medical, surgical, and trauma patients throughout the year and note that these statistics do not adjust for case-mix (i.e., varying probabilities of getting VAP) which can distort the results. Thus, when a possible VAP cluster (or any outcome which has a variable diagnostic criteria) arises, the first task should be to confirm the validity of the data by checking if a different person is doing surveillance, if there's been a change in diagnostic method, or if there are changes in the ICU case mix. In this study's ICU, the person doing surveillance and the diagnostic criteria remained constant, though case mix (particularly the prevalence of trauma patients) did vary over time.

3.4. Alternative SPC Methods

A more common but more complex SPC method is the cumulative sum (CUSUM) chart. In this method, rather than plotting an individual statistic of interest (e.g., a mean value or days between events), a statistic based on cumulative sums is derived and plotted. In our context, a VAP rate value is determined a priori—perhaps based on an ICU's rate from the last year. Going forward, the sums of the current rate minus the target (historic) rate are added. If there is a constant shift above this rate, a visual shift is evidenced in the plot and there are methods to determine if the shift is statistically significant [30]. An advantage of the CUSUM chart over traditional SPC methods is that it can detect a shift sooner; i.e., it requires a shorter Average Run Length (ARL) [30,31]. A notable variant of the CUSUM is Bernoulli CUSUM (BCUSUM) developed by Reynolds and Stoumbos [32].

Sego et al. compared a BCUSUM method to a variety of other SPC procedures that have been proposed to monitor changes in rates when the event of interest is rare [33]. Their conclusion was that the BCUSUM method was the most efficient in terms of ARL needed to detect a true shift [33]. In a related paper, Joner et al. [34] compared the BCUSUM to a prospective scan statistic method described by Naus and Wallenstein [35]. They found that the BCUSUM was slightly more efficient in detecting a rate increase but the scan method was easier to use in practice. They also show how both methods can be used together for surveillance and their methodology has much to recommend [34].

3.5. Case Mix and its Role in VAP

If case mix plays a substantial role in VAP rates, then a more advanced analysis based on risk-adjusted sequential probability ratios may be warranted, particularly in ICUs that treat a heterogeneous mix of patients over time. A practical example of this methodology can be found in [36]. An alternative approach is to derive standardized VAP ratios similar to standardized mortality ratios (i.e., to make each VAP case the event of interest instead of each hospital death) and to adjust with a regression model predicting VAP, that is analogous to mortality prediction models such as APACHE IV or MPM III familiar to the intensive care community [37,38].

We reiterate that a hospital's VAP rate is a controversial measure of quality. There are safety protocols that physicians follow to prevent VAP and there are certain interventions that focus on reducing the risk of VAP. These include weaning protocols to limit the time spent on the ventilator, protocols that prevent the aspiration of contaminated secretions, and a judicious use of antibiotics to avoid digestive tract colonization of harmful bacteria. In addition, new technologies such as subglottic suctioning and silver tipped endotracheal tubes may lower VAP rates [39, 40]. The fact that the VAP rate has decreased both in this study's ICU and nationally as evidenced by Figure 3 shows that VAP can be reduced. Nevertheless, some VAP cases are likely unavoidable (i.e., the difference between special cause and common cause variation) and physicians should carefully assess the patient's risk factors and the therapeutic interventions the patient did and did not receive when each VAP case occurs.

4. CONCLUSION

Scan statistics and control charts have been described in detail and applied to an ICU's experience with VAP during a 27-month period. Statistical methods discussed in this paper can be applied to not only VAP but also other serious events such as central line infections, medication errors, or falls to name just a few. Traditional statistical inference methods like the Chi-Square test and traditional SPC charts are often inappropriate for these scenarios. We encourage clinicians to adopt the new statistical methods of scan statistics, ratchet scan statistics, and g-type SPC charts or more advanced CUSUM methods for rare events in designing, implementing, and evaluating quality improvement initiatives.

NOMENCLATURE

k	A fixed number of events of interest that are $\leq N$ where N is the total number of events.
M	Number of consecutive months representing the seasonal trend in the ratchet scan statistic.
N	The total number of events in the study period
$P(k; N, w)$	The probability of k events out of N events occurring during time window w.
R	Ratchet scan static
S_w	The largest number of events in a window of fixed time length
w	Normalized time length; defined as the ratio of time length of interest (e.g., 1 week) divided by the total time observed (e.g., 1 year).
W_k	The smallest interval of time that contains k events

Greek

α   Significance level of the hypothesis test

REFERENCES

[1] Higgins TL, Teres D, Copes WS, Nathanson BH, Kramer AA, Stark M. Assessing contemporary ICU outcome: An updated mortality probability admission model (MPM$_0$-III). *Crit Care Med* 2007, 35, 827-835.

[2] Dasta JF, Mclaughlin TP, Moody SH, et al. Daily cost of an intensive care unit day: The contribution of mechanical ventilation. *Crit Care Med* 2005, 33, 1266–1271.

[3] Porzecanski I, Bowton DL. Diagnosis and treatment of ventilator-associated pneumonia. *Chest* 2006, 130, 597-604.

[4] Koenig SM, Truwitt JD. Ventilator-associated pneumonia: diagnosis, treatment, and prevention. *Clinical Microbiology Reviews* 2006, Oct, 637-657.

[5] Zilberberg MD, Shorr AF, Kollef MH. Implementing quality improvements in the intensive care unit: Ventilator Bundle as an example. *Crit Care Med* 2009, 37, 305-309.

[6] Niederman MS, Craven DE, Bonten MJ, et al. Guidelines for the management of adults with hospital-acquired, ventilator-associated, and healthcare-associated pneumonia. *Am J Respir Crit Care Med* 2005, 171, 388-416.

[7] Safdar N, Dezfukian C, Collard HR, et al. Clinical and economic consequences of ventilator-associated pneumonia: A systematic review. *Crit Care Med* 2005, 33, 2184-2193.

[8] Bonten MJM, Kollef MH, Hall JB. Risk factors for ventilator-associated pneumonia: from epidemiology to patient management. *Clinical Infectious Diseases* 2004, 38, 1141-9.

[9] Caruso P, Denari S, Ruiz SAL, Demarzo SE, Deheinzelin D. Saline instillation before tracheal suctioning decreases the incidence of ventilator-association pneumonia. *Crit Care Med* 2009, 37, 32-38.

[10] Klompas M, Platt R. Ventilator-associated pneumonia—the wrong quality measure for benchmarking. *Ann Intern Med* 2007, 147, 803-805.

[11] Lisboa T, Craven DE, Rello J. Safety in critical care and pulmonary medicine: Should ventilator-associated pneumonia be a quality indicator for patient safety? *Clinical Pulmonary Medicine* 2009, 16, 28-32.

[12] Canadian Critical Care Trials Group. A randomized trial of diagnostic techniques for ventilator-associated pneumonia. *N Engl J Med*. 2006, 355, 2619-30.

[13] Miller PR, Johnson JC, III, Karchmer T, et al. National nosocomial infection surveillance system: from benchmark to bedside in trauma patients. *J Trauma* 2006, 60, 98-103.

[14] Luyt CE, Chastre J, Fagon JY. Value of the clinical pulmonary infection score for the identification and management of ventilator-associated pneumonia. *Intensive Care Med* 2004, 30, 844-852.

[15] Fartoukh M, Maitre B, Honore S, et al. Diagnosing pneumonia during mechanical ventilation: the clinical pulmonary infection score revisited. *Am J Respir Crit Care Med* 2003, 168, 173-179.

[16] Glaz J, Naus J, Wallenstein S. *Scan Statistics*. Springer-Verlag, New York, NY 2001.

[17] Neff N, Naus J. *Selected Tables in Mathematical Statistics, Vol 6: The Distribution of the Size of the Maximum Cluster of Points on a Line*. American Mathematical Society, Providence, RI 1980.

[18] Rello J, Ollendorf DA, Oster G, Vera-Llonch M, Bellm L, Redman R, Kollef M. Epidemiology and outcomes of ventilator-associated pneumonia in a large U.S. database. *Chest* 2002, 122, 2115-2121.

[19] Fantus RJ, Fildes J: Trauma season. *Bull Am Coll Surg* 2006, 91, 58-59.

[20] Edwards JH. The recognition and estimation of cyclic trends. *Annals of Human Genetics*. 1961, 25, 83-87.

[21] Walter SD, Elwood JM. A test for seasonality of events with a variable population at risk. *British Journal of Preventive and Social Medicine*. 1975, 29, 18-21.

[22] Rogerson PA. A generalization of Hewitt's test for seasonality. *International Journal of Epidemiology*. 1996, 25, 644-648.
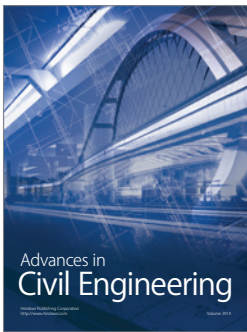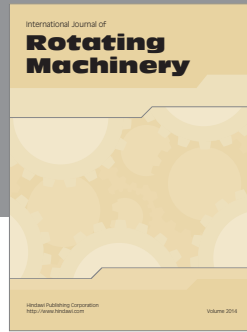
[23]　Gao F, Chia KS, Krantz I et al. On the application of the von Mises distribution and angular regression methods to investigate seasonality of disease onset. *Statistics in Medicine*. 2006, 25, 1593-1618.

[24]　Wallenstein S, Weinberg CR, Gould M. Testing for a pulse in seasonal event data. *Biometrics*. 1989, 45, 817-830.

[25]　Hwang FK. A generalization of the Karlin-McGregor theorem on coincident probabilities and an application to clustering. *Annals of Probability*. 1977, 5, 814-817.

[26]　Benneyan JC. Number-between g-type statistical quality control charts for monitoring adverse events. *Health Care Management Science*. 2001, 4, 305-318.

[27]　Benneyan JC. Performance of number-between g-type statistical control charts for monitoring adverse events. *Health Care Management Science*. 2001, 4, 319-336.

[28]　Benneyan JC, Lloyd RC, Plesk PE. Statistical process control as a tool for research and healthcare improvement. *Quality and Safety in Healthcare*. 2003, 12, 458-464.

[29]　Wall RJ, Ely EW, Elasy TA, et al. Using real time process measurements to reduce catheter related bloodstream infections in the intensive care unit. *Quality and Safety in Healthcare*. 2005, 14, 295-302.

[30]　Kennet RS, Zacks S. *Modern Industrial Statistics*. Brooks/Cole Publishing. Pacific Grove, CA 1998.

[31]　Montgomery DC. *Introduction to Statistical Quality Control* (5th ed) J. Wiley & Sons, New York, *NY* 2004.

[32]　Reynolds Jr MR, Stoumbos ZG. A CUSUM chart for monitoring a proportion when inspecting continuously. *Journal of Quality Technology*. 1999, 31(1), 87-108.

[33]　Sego LH, Woodall WH, Reynolds jr MR. A comparison of surveillance methods for small incidence rates. *Statistics in Medicine*. 2008, 27, 1225-1247.

[34]　Joner MD, William WH, Reynolds MR. Detecting a rate increase using a Bernoulli scan statistic. *Statistics in Medicine*. 2008, 27, 2555-2575.

[35]　Naus J, Wallenstein S. Temporal surveillance using scan statistics. *Statistics in Medicine*. 2006, 311-324.

[36]　Speigelhalter D, Grigg O, Kinsman R, Treasure T. Risk-adjusted probability ratio tests: applications to Bristol, Shipman and adult cardiac surgery. *Int J Qual Health Care.* 2003, 15, 7-13.

[37]　Zimmerman JE, Kramer AA, McNair DS, Malila FM: Acute Physiology and Chronic Health Evaluation (APACHE) IV: Hospital mortality assessment for today's critically ill patients. *Crit Care Med* 2006, 34(5), 1297-1310.

[38]　Higgins TL, Teres D, Copes WS, Nathanson BH, Stark M, Kramer AA. Assessing contemporary ICU outcome: An updated mortality probability admission model (MPM_0-III). *Crit Care Med* 2007, 35, 827-835.

[39]　Kollef MH. Prevention of hospital-associated pneumonia and ventilator-associated pneumonia. *Crit Care Med* 2004, 32, 396-1405.

[40]　Gastmeier P, Geffers C. Prevention of ventilator-associated pneumonia: analysis of studies published since 2004. *Journal of Hospital Infection*. 2007, 67, 1-8.

Appendix A. National Nosocomial Infection Surveillance System (NNIS) Clinical Criteria for Diagnosing Pneumonia [13]

Radiographic	Two or more serial chest radiographs with new or progressive and persistent infiltrate or cavitation or consolidation (one radiograph is sufficient in patients without underlying cardiopulmonary disease).	
Clinical	**One of the following**: • Fever > 38 degrees Celsius. • WBC count $\leq$ 4,000/μL or $\geq$ 12,000 μL. • Altered medical status in adults > 70 with no other recognizable cause.	**And at Least two of the following**: • New-onset purulent sputum or change in character of sputum, or increase in respiratory secretions or suctioning requirements. • New-onset or worsening cough, dyspnea, or tachypnea. • Rales or bronchial breath sounds.
Microbiology (optional)	Positive culture result (in one of the following): • blood (unrelated to other source) • pleural fluid • quantitative culture by BAL or PSB • $\geq$ 5% BAL-obtained cells contain intracellular bacteria.	

Appendix B. Clinical Pulmonary Infection Score Clinical Criteria for the Diagnosis of Pneumonia [3,14,15]

Variable	0	1	2
Temperature in Celsius	$\geq$ 36.1 to $\leq$ 38.4	$\geq$ 38.5 to $\leq$ 38.9	$\geq$ 39 to $\leq$ 36
WBC count, µL	$\geq$ 4,000 to $\leq$ 11,000	< 4,000 or >11,000	
Secretions	Absent	Present, non-purulent	Present, purulent
PaO_2/fraction of inspired oxygen	> 240 or ARDS		
Chest Radiography	No infiltrate	Diffuse or patchy infiltrate	Localized infiltrate
Microbiology	No or light growth	Moderate or heavy growth; add 1 point for same organism on Gram stain	