

## Princípios básicos de procedimentos estatísticos aplicados na análise de dados de vigilância em saúde pública: uma revisão

Basic principles of statistical methods applied to public health surveillance data analysis: a review

Odécio Sanches <sup>1</sup>

---

<sup>1</sup> Av. Portugal 1690, apto. 81,  
Ribeirão Preto, SP  
14020-380, Brasil.

**Abstract** *Statistical methods in public health surveillance data analysis are important for detecting temporal and/or spatial disease clusters, which can indicate an outbreak or an epidemic. In this paper we present a review and literature survey introducing the public health surveillance process and focusing on some statistical methods which are indicated to the data analysis in routine procedure. Comparative studies of these methods are also analyzed.*

**Key words** *Public Health Surveillance; Health Statistics; Statistical Methods and Procedures; Statistics*

**Resumo** *Procedimentos estatísticos na análise de dados, obtidos rotineiramente pelos sistemas de vigilância em saúde pública, são importantes para detectar conglomerados temporais e/ou espaciais de casos de doenças, os quais podem indicar a ocorrência de um surto ou de uma epidemia. Neste trabalho são apresentados uma revisão e um survey da literatura, introduzindo o processo de vigilância, bem como a limitação dos procedimentos de análise estatística, determinada pela natureza desse processo, e centrando a atenção em alguns métodos estatísticos indicados para a análise rotineira de dados. Também são apresentados e discutidos estudos comparativos entre métodos.*

**Palavras-chave** *Vigilância em Saúde Pública; Estatística de Saúde; Métodos e Procedimentos Estatísticos; Estatística*

## Introdução

O tema vigilância em saúde pública, no Brasil, tem sido objeto de uma viva e constante discussão entre acadêmicos, pesquisadores e profissionais de serviços, como atestam as apresentações de trabalhos, conferências e mesas redonda nos congressos de epidemiologia e de saúde pública, bem como as publicações nos principais periódicos do país, nessas áreas. Essas discussões, no entanto, não têm abordado a questão dos procedimentos ou métodos estatísticos para análise rotineira de dados de vigilância em saúde pública.

No âmbito internacional, o interesse tem sido grande, crescente e atual, como é evidenciado pelo periódico *Statistics in Medicine*, que co-patrocinou duas importantes reuniões de trabalho sobre o tema: o *Symposium on Statistics in Surveillance*, juntamente com o CDC (*Centers for Disease Control and Prevention*) de Atlanta, Geórgia, Estados Unidos, no ano de 1988, quando os anos 80 foram considerados a década da estatística no CDC (Mason, 1989); e o *Workshop on Statistics and Computing in Disease Clustering*, em 1992. Parte substancial dos trabalhos apresentados nesses dois eventos foi publicada, respectivamente, nos volumes 8, agosto de 1989, e 12, números 19/20, outubro de 1993, desse periódico.

Tal interesse não parece estar, ainda, generalizado nos serviços, como destacam Thacker et al. (1989): “*Para muitos profissionais de saúde pública, entretanto, incluindo epidemiologistas, a estatística é o aspecto da vigilância mais difícil de ser entendido e utilizado*”.

Isso pode ser conseqüência de os textos acadêmicos para cursos de medicina preventiva, saúde pública, epidemiologia e áreas afins não abordarem métodos estatísticos relacionados à vigilância em saúde pública, sendo raros os textos que tratam dos aspectos mais gerais da vigilância, apesar do destaque dado por Thacker & Berkelman (1988): “*A vigilância é a pedra angular da saúde pública*”.

Com a intenção de levar uma contribuição aos profissionais de ensino e de serviços de vigilância em saúde pública, este trabalho de revisão tem por objetivo descrever, resumidamente, os princípios teóricos que fundamentam os procedimentos estatísticos que se adaptam à análise de dados de vigilância, apontando onde e quando surgiram aplicações práticas dos mesmos. São apresentados, ainda, estudos comparativos entre distintos procedimentos, discutindo-se vantagens ou desvantagens de cada um, assim como áreas que necessitam mais investigação, principalmente de caráter

aplicado. Pretende-se, ao mesmo tempo, que o trabalho seja um *survey* da literatura, que poderá servir de suporte às atividades dos profissionais em questão.

Por razão de competição por espaço, a apresentação de gráficos/diagramas ou de tabelas se reduz a um mínimo necessário para uma descrição sucinta.

## O processo de vigilância em saúde pública e a presença da estatística

O CDC, que tem sido um órgão de referência mundial para as atividades de vigilância, estabelece que qualquer sistema de vigilância deve ter capacidade funcional para coleta, análise e disseminação de dados ligados a programas de saúde pública, de forma contínua (CDC, 1986), isto é, um processo ininterrupto (Colton, 1989).

Essas três etapas, ou seja, coleta, análise e disseminação dos dados, compoem o aspecto operacional da vigilância em saúde pública. Mason (1989) aponta a necessidade de uma componente adicional, periódica, que consiste na avaliação do sistema.

No contexto em que se insere este trabalho, o acompanhamento contínuo da população é uma questão fundamental. Inquéritos pontuais ou esporádicos, por amostragem, por exemplo, não são considerados atividades de vigilância. Um sistema contínuo, também, não é suficiente para constituir um sistema de vigilância, a não ser que as informações coletadas e análises realizadas estejam integradas à administração e avaliação de programas específicos em saúde (Thacker & Berkelman, 1988).

Por outro lado, um sistema que só coleta, consolida e armazena as informações, ainda que coletadas de forma contínua, não constitui um sistema de vigilância em saúde pública; constitui apenas um arquivo de dados de vigilância, pois exclui a disseminação da informação e a conseqüente não-incorporação nas práticas de saúde pública (Thacker et al., 1989).

Isto posto, considerando-se as etapas operacionais do processo de vigilância, é fácil perceber que a estatística é parte integrante desse processo. De fato, excluindo-se a análise de dados, objeto principal deste trabalho, a estatística está presente: na etapa de coleta de dados, contribuindo com a discussão e elaboração dos instrumentos de coleta, do controle de qualidade desses instrumentos e dos dados coletados, cuidando de métodos para limpeza dos dados, assim como da metodologia apropriada para tratamento de dados perdidos, etc.; na etapa de disseminação dos dados que, em geral, é fei-

ta através de tabelas, gráficos e indicadores numéricos; e, finalmente, na etapa de avaliação, onde, igualmente, são utilizados indicadores numéricos, como sensibilidade, especificidade e valor preditivo do sistema, além de questões relacionadas à otimização de custos e otimização geral de tal sistema.

### Uma preliminar sobre a análise estatística de dados de vigilância

A situação geral que se apresenta no processo de vigilância, pelo prisma da análise estatística dos dados, pode ser resumida do seguinte modo: uma população, definida pelas características dos seus elementos e ocupando um especificado espaço geográfico, amplo ou restrito, constituída de  $N$  elementos, dos quais  $M$  são portadores de uma ou mais doenças, é monitorada continuamente no tempo.

O que interessa saber, considerando os objetivos da vigilância em saúde pública, é se essas  $M$  ocorrências são aleatórias ou não; mais especificamente, o objetivo é determinar a ocorrência ou não de conglomerados de doenças e/ou agravos à saúde.

O CDC define um conglomerado como uma agregação incomum, real ou percebida, de eventos de saúde que são agrupados no tempo e/ou no espaço e que são relatados às agências de saúde (CDC, 1990).

Assim, o objetivo da análise estatística, principalmente rotineira, é detectar a presença de conglomerados de casos e se tais conglomerados têm alta probabilidade de serem aleatórios ou não. Portanto, em função do processo de vigilância, as únicas hipóteses estatísticas que têm significado para dados de vigilância em saúde pública são aquelas de aleatoriedade ou não das ocorrências.

Pela natureza do processo de vigilância, não existe, associado ao mesmo, um procedimento de amostragem ou de alocação aleatória na coleta de dados. Conseqüentemente, não existe a presença do erro aleatório. Existem erros de registro, erros por falta de resposta, erros provocados por cobertura inadequada, erros na interpretação da definição de caso, etc., mas nenhum desses é aleatório. Ainda que se tome um subconjunto de dados, por exemplo, dentro de um especificado intervalo de tempo, isso constituiria uma amostra intencional, que não é, definitivamente, uma amostra aleatória (Kish, 1965).

Cabe, pois, um alerta: procedimentos clássicos de testes de hipóteses sobre parâmetros populacionais, para comparação de subgrupos

dentro da população monitorada e construção de intervalos de confiança, que se apóiam na existência de amostragem aleatória (Lehmann, 1983, 1986) devem ser vistos com cautela e, até mesmo, com certa dose de desconfiança (Sanches, 1993).

No que diz respeito às comparações internas, afirma Hall (1989) que: entre as  $M$  ocorrências detectadas, se for de interesse, é possível estabelecer estratificações, por grupos étnicos, classes sócio-econômicas, escolaridade etc. Mas não há porque comparar, via testes de hipóteses, parâmetros populacionais, simplesmente porque tais estratos são estratos populacionais ou, no máximo, amostras de tamanho um (1). Tais dados comportam, por exemplo, o teste exato de Fisher, para comparar dois subgrupos, no sentido de teste de aleatoriedade ou, no caso de  $k$  subgrupos ( $k > 2$ ), o teste de Pitman, que é um teste de permutação.

*“Um conglomerado pode ser útil para gerar hipóteses mas não é útil, provavelmente, para testar hipóteses. Frequentemente, as questões levantadas pela detecção de conglomerados não podem ser respondidas pela detecção, por si só; elas requerem uma investigação epidemiológica, alternativa, com desenho específico, embasado em algum procedimento de amostragem aleatória”* (CDC, 1990).

Nesse sentido de examinar conglomerados no espaço/tempo, pelo enfoque da investigação epidemiológica, recentemente, Werneck & Struchiner (1997) apresentaram uma excelente revisão. Aqui, neste trabalho, a preocupação é apenas com métodos que levam a testes de aleatoriedade, ainda que se faça analogia com os testes clássicos de Neyman-Pearson.

### Controle estatístico de qualidade de um processo de produção industrial e vigilância em saúde pública: uma analogia

No controle estatístico de qualidade, em um processo de produção industrial, o problema consiste em monitorar a saída do processo, de modo contínuo, inspecionando lotes ou amostras do produto, no denominado esquema de Shewhart (1931) ou inspecionando itens produzidos continuamente, em um análogo de um processo seqüencial, no denominado esquema *Cusum* (Page, 1954), a fim de verificar se a qualidade de lotes ou itens produzidos se mantém dentro de um padrão pré-fixado. Caso isso deixe de ocorrer, o procedimento estatístico utilizado na análise dos resultados deve disparar um alarme, e, se o valor da qualidade ultrapas-

sar um determinado limite, denominado limite de controle, o processo de produção deve ser interrompido até que as causas da alteração sejam removidas. Após esse procedimento, o processo de produção é retomado.

Em todo processo de produção industrial, existe sempre uma certa variabilidade natural que é inerente ao processo e que é resultante de muitas pequenas causas essencialmente inevitáveis. Quando a variabilidade é relativamente pequena, considera-se o processo em nível aceitável de performance e diz-se, então, que o sistema está sob controle estatístico.

Outras causas de variabilidade, porém, podem estar presentes, como máquinas não ajustadas adequadamente, erros de operadores, impurezas do material utilizado, etc. Essa variabilidade, quando presente, é geralmente muito maior que a variabilidade natural. Quando isso ocorre, o processo de produção se apresenta em um nível de performance não aceitável, então o sistema de produção é dito fora de controle estatístico.

O processo de vigilância em saúde pública é, também, um processo contínuo no tempo, cujo objetivo é monitorar alguma característica relacionada à ocorrência de uma ou mais doenças, expressa, por exemplo, através de taxas médias de incidência, ou número médio de novos casos. Se estas taxas ou esses números médios se mantêm dentro de um intervalo de variação pré-fixado, a condição de saúde da população monitorada, em relação a essas doenças, é considerada estável, isto é, por analogia, sob controle.

Reforçando o que já foi expresso, toda a questão estatística na análise de dados de vigilância, através do monitoramento da população, é detectar possíveis alterações bruscas nos valores dessas taxas, ou dessas médias, e, quando detectadas, decidir se as mesmas ocorrem por mera chance ou não, nesse último caso, dando elementos para uma possível intervenção da saúde pública.

Dada esta analogia e dada a simplicidade de alguns métodos estatísticos de controle de qualidade, tem havido, na literatura internacional, uma maior divulgação e incentivo para a utilização dos mesmos na análise rotineira de dados de vigilância em saúde pública, apesar de uma subutilização (Reinke, 1991).

## Base estatística dos diagramas de controle de qualidade

### O procedimento de Shewhart

O diagrama de controle de Shewhart (1931) pode ter dois enfoques: se a característica de interesse pode ser mensurada e expressa como um número, em alguma escala contínua, o diagrama é dito de controle para variável; se a característica de interesse não é mensurável em uma escala contínua, ou até mesmo em qualquer escala quantitativa, a qualidade é julgada em termos de proporção de atributos como, por exemplo, defeituoso ou não defeituoso (por analogia, em vigilância, a doença está presente ou não presente), quando o diagrama é dito de controle para atributos. O diagrama mais utilizado para variável é o de controle de uma média aritmética, para o controle de uma medida de tendência central.

Basicamente, um diagrama de controle de Shewhart é constituído de um sistema de dois eixos coordenados cartesianos ortogonais: no eixo das abscissas são representados os instantes de registro das informações e no eixo das ordenadas, a medida da característica correspondente a cada instante (no caso de um diagrama para variáveis) ou a proporção dos itens com o atributo de interesse (no caso do diagrama de atributos). Uma linha central, paralela ao eixo das abscissas, representa o valor médio da característica correspondente ao estado do processo sob controle. Duas linhas paralelas à linha central, respectivamente, acima e abaixo desta, são traçadas correspondendo aos limites de alarme e de controle. Este diagrama é para um procedimento bicaudal; no caso monocaudal, bastam apenas os limites de alarme e de controle num só sentido, dependendo do interesse de controlar aumentos ou diminuições no valor de referência.

Essencialmente, um diagrama de controle é um teste de hipótese para decidir se o processo se encontra sob controle ou, equivalentemente, um teste de hipótese que somente causas atribuíveis à chance estão presentes no processo.

Assim, por analogia com a teoria de Neyman-Pearson, podemos falar nas probabilidades de erros tipo I e II no diagrama de controle: evidenciar que o processo está fora de controle quando realmente está sob controle ou evidenciar que o processo está sob controle quando realmente está fora de controle; analogamente em relação aos limites de alarme.

Para o controle de uma média  $\bar{x}$ , suponhamos que, nas condições em que só atua a chance,  $\mu_{\bar{x}}$  é a média do processo e  $\sigma_{\bar{x}}^2$  é a variância.

Então, no diagrama de controle, os limites de alarme e de controle são dados por:  $\mu_{\bar{x}} \pm k_1 \sigma_{\bar{x}}$  para os limites de alarme, e  $\mu_{\bar{x}} \pm k_2 \sigma_{\bar{x}}$  para os limites de controle. Em controle de qualidade, em geral, se escolhe  $k_1 = 2$  e  $k_2 = 3$ .

Os limites de alarme e de controle podem ser fixados, alternativamente, em termos de valores de probabilidades.

Especificar os limites de alarme e de controle é uma das decisões críticas do desenho do diagrama de controle, em razão da presença dos erros de decisão.

Outra característica importante associada aos diagramas de controle de qualidade é o *Average Run Length* (ARL) do procedimento.

Quando o processo se mantém sob controle, isto é, a qualidade permanece estável, o ARL associado é o número esperado de itens inspecionados antes que uma ação precise ser tomada. Para o sistema fora de controle, o ARL mede o atraso e a quantidade de defeituosos produzidos antes que uma ação seja tomada. É desejável, então, montar um esquema de monitoramento em que o ARL seja grande para o sistema sob controle e pequeno quando o sistema sai fora de controle. Pormenores, inclusive analíticos, dessas relações são facilmente encontrados em textos de controle estatístico de qualidade, entre eles, Duncan (1974) e Montgomery (1990).

### O procedimento *Cusum*

O procedimento *Cusum* (*Cumulative sums*) foi proposto por Page (1954) e leva, também, a um diagrama de controle, mas, modernamente, com a implementação de programas computacionais específicos de controle estatístico de qualidade, é possível obter o que se denomina forma tabular (Santos, 1997). O procedimento *Cusum* foi criado como uma alternativa ao de Shewhart por causa de algumas vantagens, como a de levar em consideração todas as informações obtidas no processo, e não apenas a última, entre outras que serão apontadas na discussão da comparação de procedimentos.

O procedimento *Cusum* também se aplica tanto ao controle de variáveis quanto ao de atributos.

Originariamente, a proposta é a seguinte: fazem-se observações, a intervalos regulares de tempo, assinalando-se um escore  $x_k$  à  $k$ ª observação e representando-se, em um diagrama de eixos coordenados, a soma acumulada:

$$S_n = \sum_{k=1}^n x_k$$

Um alarme será disparado e uma ação será tomada se  $(S_n - \min S_i) > h$ ;  $0 \leq i < n$  e  $h$  é o valor que define o intervalo de decisão, por exemplo,

em múltiplos de desvio padrão. Este é um teste monocaudal.

Alternativamente, é possível definir:

$S'_n = \max(S'_{n-1} + x_n, 0)$ ;  $n \geq 1$ ;  $S'_0 = 0$ , de modo que  $S'_n = 0$  quando  $S'_n < \min S_i$ ;  $i = 1, 2, \dots$

Uma decisão será tomada se  $S'_n \geq h$ .

Esta regra se decompõe em uma seqüência de testes seqüenciais de Wald (Page, 1954).

Uma dificuldade com essas regras de decisão, conforme aponta Duncan (1974), é que o procedimento pode apanhar pequenas alterações, que não são de interesse, dentro do processo. Introduce-se, então, uma modificação que consiste em se fixar um valor de referência um pouco maior, digamos  $k$  unidades a mais. Ou seja, se o valor de referência é uma média aritmética, digamos  $\mu_0$ , tomar um valor de referência  $(\mu_0 + k)$ . Embora a escolha de  $k$  seja feita em função do particular problema que se tem no momento, em geral,  $k$  é tomado como  $k \equiv \frac{d}{2}$ ,

onde  $d = (\mu_1 - \mu_0)$  e  $\mu_1$  é o valor da característica para o qual se considera o sistema fora de controle.

Então, as condições do teste são:

$S'_0 = 0$  e  $S'_n = \max[S'_{n-1} + (\bar{x}_n - k); 0]$ , no caso do controle de uma média aritmética, e  $S'_n > h$ , no caso de tomar a decisão de que o processo está fora de controle.

Em geral,  $h$  e  $k$  são escolhidos de modo a dar ao esquema algumas propriedades desejáveis em relação aos erros de decisão e aos valores do ARL.

Barnard (1959) analisa os esquemas de controle de qualidade, em particular o procedimento *Cusum*, sob a estrutura de processos estocásticos e, também, propõe um diagrama para o teste bicaudal, posteriormente denominado máscara em V (De Bruyn, 1968).

O cálculo do ARL, no esquema *Cusum* é mais complexo que no esquema de Shewhart. Ewan & Kemp (1960) e Goel & Wu (1971) construíram nomogramas para a determinação do ARL, no caso de uma distribuição Normal de variância unitária. Ewan & Kemp produziram, também, tabelas de valores do ARL para o caso de uma variável Poisson. Processo markoviano para o cálculo do ARL é discutido por Brooks & Evans (1972) e por Yashchin (1992).

Uma ilustração da construção e comportamento desses dois tipos de diagramas é feita na Tabela 1. Os dados utilizados são número de ocorrências semanais de casos de meningite meningocócica, nos anos de 1992 e 1994, na cidade do Rio de Janeiro, obtidos na Secretaria Municipal de Saúde, dessa localidade (Santos, 1997).

Tabela 1

Casos de meningite meningocócica na cidade do Rio de Janeiro, nos anos de 1992 e 1994\*, segundo a semana do ano.

Semana	1992	1994
01a	02	09
02a	11	09
03a	07	09
04a	07	16
05a	02	09
06a	06	03
07a	03	09
08a	09	06
09a	07	10
10a	04	06
11a	04	07
12a	06	09
13a	03	10
14a	06	06
15a	04	04
16a	05	10
17a	07	14
18a	05	05
19a	06	07
20a	03	06
21a	06	10
22a	08	10
23a	03	09
24a	06	07
25a	12	11
26a	09	12
27a	08	08
28a	11	13
29a	06	11
30a	04	14
31a	08	06
32a	10	16
33a	09	15
34a	08	11
35a	05	12
36a	08	07
37a	05	10
38a	09	08
39a	03	17
40a	02	06
41a	14	06
42a	09	08
43a	09	10
44a	02	13
45a	07	16
46a	03	13
47a	04	05
48a	03	10
49a	05	11
50a	04	08
51a	06	10
52a	07	07
53a	06	-

\* Dados da Secretaria Municipal de Saúde, RJ (Santos, 1997).

Como ilustração, o interesse é estudar o comportamento do número de casos em 1994, tomando como referência a média aritmética e o desvio padrão da distribuição de casos no ano de 1992. Denotando por  $\mu_0$  e  $\sigma_0$  esses parâmetros de referência, os dados da Tabela 1 fornecem  $\mu_0 = 6,15$  e  $\sigma_0 = 2,67$ .

A distribuição do número de caso para o ano de 1994 se ajusta a uma distribuição de Poisson (teste de Kolmogorov-Smirnov:  $d = 0,047$ , não significativo; teste do  $\chi^2$ :  $p = 0,67$ ); é feita, então, uma aproximação da Poisson para a distribuição Normal (teste de Kolmogorov:  $d = 0,13$ , não significativo; teste do  $\chi^2$ :  $p = 0,21$ ).

Os gráficos das Figuras 1 e 2 ilustram esses ajustes.

Os dados do ano de 1994, normalizados e padronizados usando a média e o desvio padrão da população de referência, isto é,  $\mu_0$  e  $\sigma_0$ , são denotados, nesta ilustração, por  $y_n$ ;  $n = 1, 2, \dots, 52$ .

Para a construção do diagrama *Cusum*, fixou-se um  $(ARL)_0$ , isto é, o ARL com base na hipótese  $H_0$  de que o sistema está sob controle, tal que  $(ARL)_0 = 1.000$ . Isso implica, em média, que o valor esperado de uma situação fora de controle, quando o sistema realmente está sob controle, deverá ocorrer uma vez a cada mil semanas, ou seja, aproximadamente, a cada 19,2 anos. É uma exigência bastante forte.

Fixou-se o valor de  $d$ , isto é, o deslocamento que deseja detectar em relação a  $\mu_0$ , igual a um desvio padrão da variável  $y_n$  padronizada; isto implica  $k = 0,5$ .

Com esses valores de  $(ALR)_0$  e  $k$ , determinou-se, a partir do nomograma de Ewan & Kemp (1960), o valor do limite superior de controle  $h$ , para um teste monocaudal;  $h = 4,4$ .

As somas acumuladas  $S'_n$  são calculadas utilizando os valores de  $y_n$ ;  $n = 1, 2, \dots, 52$ ; impondo uma interrupção do processo sempre que  $S'_n > h$ , após a qual o processo é reinicializado (supõe-se que a detecção de uma situação fora de controle interrompa o processo para que ações corretivas possam ser tomadas). Com o mesmo algoritmo de  $S'_n$  calculou-se  $S_n^*$ , as somas acumuladas sem interrupção e reinicialização do processo. Isso permite que o leitor visualize no diagrama como o processo se deteriora rapidamente se ações corretivas não forem tomadas quando  $S'_n > h$ .

O diagrama de Shewhart, monocaudal superiormente, foi construído utilizando também os valores transformados  $y_n$  para permitir uma melhor comparabilidade com o diagrama *Cusum*. Poderia, no entanto, ser construído a partir dos dados brutos. Para esse valor padronizado, o valor de referência se torna  $\mu_0 = 0$ . To-

mando o mesmo (ARL)<sub>0</sub> do procedimento *Cusum*, isso produz, para o diagrama de Shewhart,  $\alpha = 0,001$ , já que, nesse procedimento,  $\alpha = \frac{1}{(A.R.L.)_0}$  é a probabilidade de detectar

que o processo está fora de controle, quando realmente está sob controle. Com esse valor de  $\alpha$ , determina-se, a partir da Normal (0; 1), o valor do limite superior de controle (LSC) = 3,09. Sempre que  $y_n < 0$  se reinicializa o processo em  $y_n = 0$ .

As Figuras 3 e 4 apresentam a construção desses diagramas. Observa-se, a partir dos mesmos, que o diagrama *Cusum* capta mais rapidamente e em maior número as situações fora de controle.

É bom salientar que os parâmetros escolhidos o foram por razões meramente estatísticas. Se a escolha de tais parâmetros estivesse adequada às características e à história natural da doença, poder-se-ia dizer que, se ações corretivas de intervenção não fossem tomadas, a cidade teria enfrentado um surto a partir da 26ª semana do ano de 1994.

Barnard (1959), em suas considerações sobre esquemas de controle de qualidade como processos estocásticos, admite, por esse prisma, a dependência entre as variáveis, para o processo sob controle. No entanto, Johnson & Bagshaw (1974) e Bagshaw & Johnson (1975), indicam que, no procedimento *Cusum*, a correlação serial pode afetar não só o valor do ARL, mas também a natureza de sua distribuição de probabilidades e que, como consequência, o procedimento pode não ser robusto com respeito ao afastamento da pré-condição de independência quando o processo se encontra sob controle.

Considerando que esses dois procedimentos de controle estatístico de qualidade, por sua simplicidade e facilidade de implementação, são adequados para serem adaptados à rotina de análise de dados de vigilância em saúde pública, é conveniente citar que a literatura específica é rica em referências relativas a tais procedimentos, entre as quais: Montgomery (1980), relacionado a questões econômicas; Page (1957), Hinkley (1971), Pettitt (1979, 1980) e Assaf et al. (1992), relacionados ao problema *change point*; Nagendra & Rai (1971) e Frisén & De Maré (1991), relacionados a questões de otimalidade; Alwan (1986) e Pignatiello Jr. & Runger (1990), a questões de *Cusum* multivariado; Ncube & Woodall (1984), Westgard et al. (1977) e Lucas (1982), para questões relacionadas à combinação de esquema Shewhart – *Cusum*.

A Tabela 2 apresenta uma síntese de um *survey* da literatura com enfoque aplicado à prática da análise de dados de vigilância.

Figura 1

Distribuição de Poisson esperada para os dados observados de 1994 da Tabela 1.

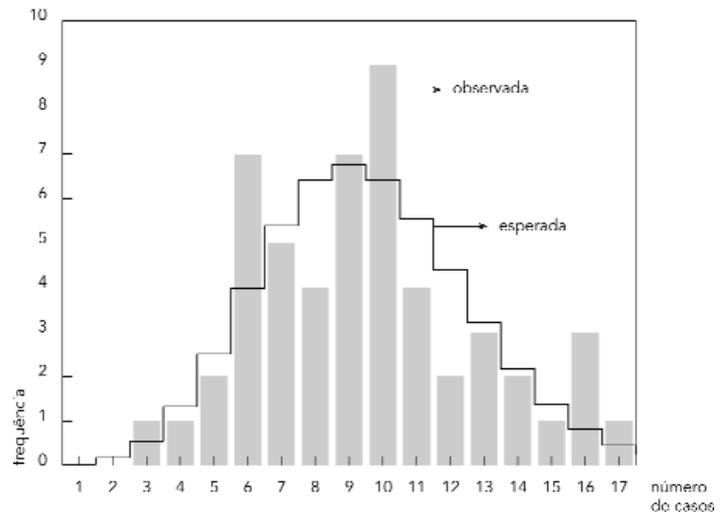


Figura 2

Distribuição Normal esperada para os dados observados de 1994 da Tabela 1.

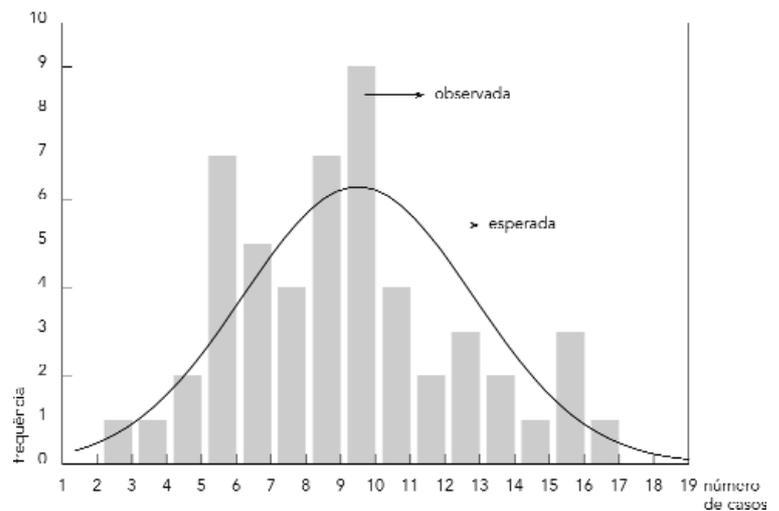


Figura 3

Diagrama de somas acumuladas (*Cusum*) para os dados de 1994 da Tabela 1.

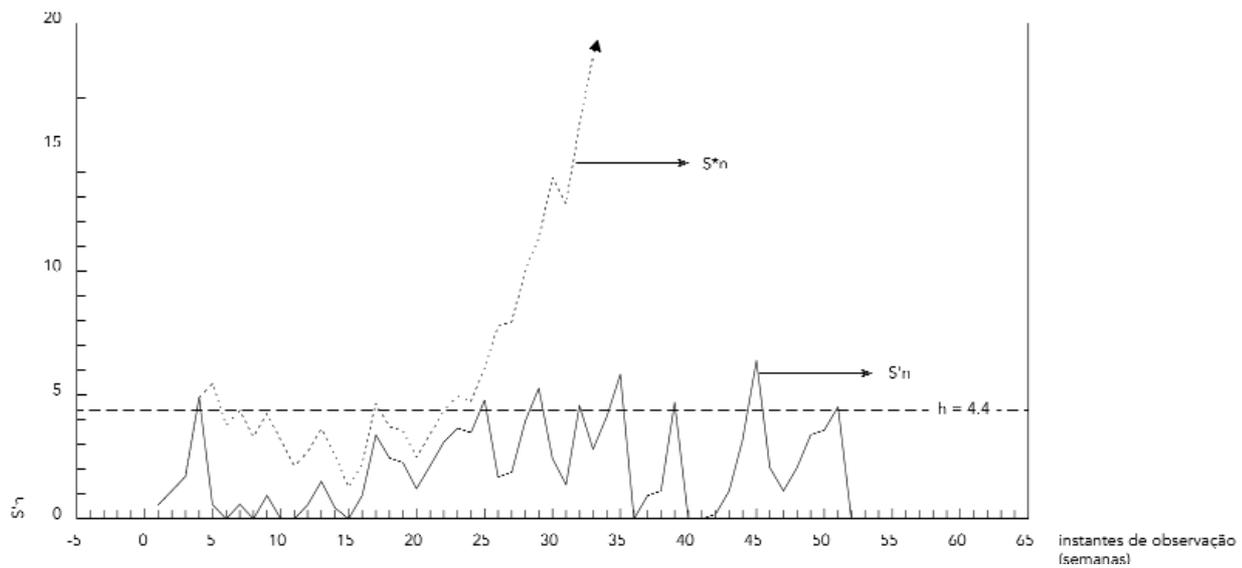


Figura 4

Diagrama de Shewart para os dados de 1994 da Tabela 1.

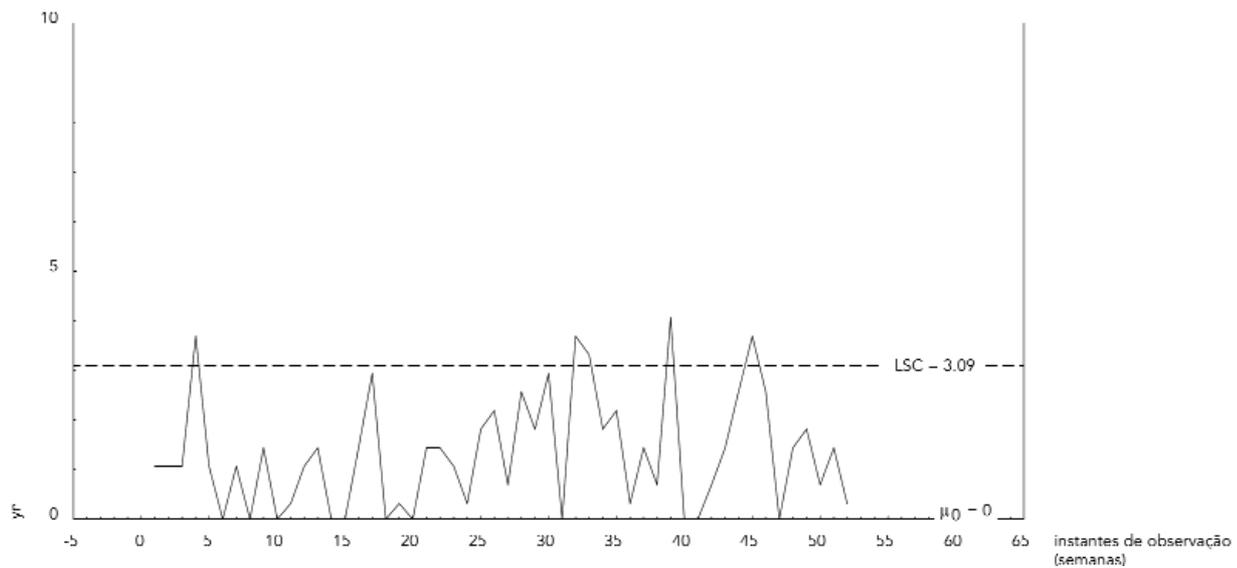


Tabela 2

Os procedimentos de Shewhart e *Cusum*, na literatura, com aplicações à vigilância em saúde pública e áreas correlatas.

Autor(es)	Procedimento	Área de utilização do procedimento
Serfling (1944)	Shewhart	Vigilância de poliomielite
Rich (1946)	Shewhart	Controle de qualidade de pessoal de enfermagem
Levey & Jennings (1950)	Shewhart	Controle de qualidade de exames laboratoriais em clínica médica
Hill et al. (1968)	<i>Cusum</i>	Vigilância de má-formação congênita
Bjerkedal & Bakkeiteig (1975)	<i>Cusum</i>	Vigilância de má-formação congênita
Weatherall & Haskey (1976)	<i>Cusum</i>	Vigilância de má-formação congênita
Westgard et al. (1977)	Shewhart- <i>Cusum</i> combinados	Controle de qualidade em química clínica
Kemp et al. (1978)	<i>Cusum</i>	Controle de qualidade em radioimunoensaio
Wilson et al. (1979)	<i>Cusum</i>	Controle de qualidade em radioimunoensaio
Edwards (1980)	<i>Cusum</i> e Máscara em V truncada	Controle de qualidade em química clínica
Rowlands et al. (1980)	<i>Cusum</i>	Controle de qualidade em química clínica
Royston & Abrams (1980)	<i>Cusum</i>	Monitoramento de temperatura corporal basal em mulheres em idade de procriar
Westgard et al. (1981)	Shewhart	Controle de qualidade em química clínica
Thompson (1983)	<i>Cusum</i> Modificado	Monitoramento de doenças isquêmicas do coração
Barbujani & Calzolari (1984)	<i>Cusum</i>	Vigilância de má-formação congênita
Levin & Kline (1985)	<i>Cusum</i> Modificado	Monitoramento de abortos espontâneos
Raubertas (1989)	<i>Cusum</i>	Vigilância envolvendo várias unidades de saúde
Reinke (1991)	Shewhart	Controle de qualidade de pessoal de saúde pública
Royston (1991)	<i>Cusum</i>	Monitoramento de temperatura corporal basal em mulheres em idade de procriar
Santos (1997)	Shewhart e <i>Cusum</i>	Análise de dados de vigilância para três doenças distintas

## Outros procedimentos

### A estatística *scan*

A extensão ou duração do monitoramento de uma doença e/ou agravo à saúde pode ser fixado e finito ou pode se estender indefinidamente para o futuro. A análise de dados de um monitoramento finito é, geralmente, retrospectiva. Dados históricos são examinados para detectar algum deslocamento em relação a um valor padrão, a fim de evidenciar tendências ou, ainda, detectar conglomerados de casos. Na literatura, um procedimento que tem sido proposto para monitoramento finito é a denominada estatística *scan* (Raubertas, 1989).

A estatística *scan* se originou em uma série de trabalhos de J. I. Naus. Suponha que  $N$  eventos ocorram em um intervalo  $(0, T]$ , onde as unidades de medida do intervalo podem ser de

tempo ou de distância, assumindo-se, também, que  $N$  é conhecido e fixado e que os eventos observados são identicamente distribuídos, segundo uma distribuição uniforme, e independentes. A hipótese a ser testada é a de aleatoriedade das ocorrências no intervalo  $(0, T]$  contra hipóteses alternativas de ocorrência de conglomerados no mesmo intervalo (Naus, 1965, 1966a, 1966b).

Sem perda de generalidade assume-se  $T = 1$ .

Considerando  $x_1, x_2, \dots, x_N$ , indicando os tempos ou posições de ocorrência dos eventos e considerando, ainda,  $Y_{t,t+w}$ , representando o número de eventos sujeito à varredura através de uma janela de amplitude  $(t, t + w]$ ;  $0 \leq t \leq 1 - w$ , define-se:  $S_w = \max Y_{t,t+w}$ .  $S_w$  é conhecida como estatística *scan*.

A hipótese nula de aleatoriedade é rejeitada, em nível de significância  $\alpha$ , sempre que o valor da estatística *scan* atinge ou supera o valor crítico  $m$ , isto é,  $P_r [S_w \geq m] \leq \alpha$ .

Portanto, para implementar o teste é necessário calcular as probabilidades, na cauda, da distribuição de  $S_{i|}$ . A questão da determinação destas probabilidades, exatas ou aproximadas, é a grande questão teórico-prática para a implementação do teste (Wallestein & Naus, 1973; Huntington & Naus, 1975; Wallestein, 1980; Wallestein et al., 1989; Glaz, 1993; Sahu et al., 1993; Wallestein et al., 1993). Outras questões teórico-práticas, ainda em aberto, também são descritas por Glaz (1993).

Aplicações práticas são referenciadas em Wallestein (1980), monitorando a ocorrência de trisomias; Wallestein et al. (1989), monitorando a ocorrência de suicídios entre adolescentes; e Levin & Kline (1985). Glaz (1993) referencia outros estudos que utilizam a estatística *scan* na área médica.

#### **Os procedimentos de Chen para dados de vigilância de más-formações congênicas e para doenças crônicas**

Chen (1978) propõe um procedimento para análise de dados de vigilância de más-formações congênicas, que se aplica tanto a áreas restritas, como um só hospital, quanto a uma reunião de áreas restritas.

O procedimento se baseia no número de nascimentos consecutivos que ocorrem entre o nascimento de uma criança portando uma específica má-formação e o nascimento seguinte com essa má-formação. Este grupo de nascimentos constitui um conjunto ou agrupamento e o seu tamanho, ou amplitude, é uma variável aleatória com distribuição geométrica, que constitui a estatística do teste. Se surge uma seqüência de agrupamentos, de forma que cada agrupamento tenha uma amplitude abaixo de um certo tamanho, o sistema dispara um alarme.

Uma tal seqüência deverá ocorrer com baixa probabilidade sob a hipótese nula de que a taxa de má-formação ocorre somente por chance; caso contrário, esta seqüência deverá ocorrer com alta probabilidade.

Se  $X_i$  é a amplitude do *imo* agrupamento, e  $\pi_0$ , a taxa padrão sob  $H_0$ , a amplitude esperada dos agrupamentos é  $\eta_0 = E[X_i] = \frac{1 - \pi_0}{\pi_0}$  e,

em geral, para qualquer valor de amplitude expresso como um múltiplo de  $\eta_0$  tem-se  $P_r[X_i < k\eta_0] \cong 1 - e^{-k}$ ;  $k > 0$ ;  $i = 1, 2, \dots$

Pode-se, então, calcular a função geradora de momentos e obter o valor esperado de  $X_i$ , condicionado a  $X_i < k\eta_0$ , obtendo-se  $E[X_i | X_i < k\eta_0] \cong \frac{\eta_0(1-k)}{e^k - 1}$ ;  $i = 1, 2, \dots$ ,

e sob  $H_1$ , sendo  $\eta_1 = E[X_i] \cong \frac{\eta_0}{\gamma}$ ,

$$E[X_i | X_i < k\eta_0] \cong \frac{\eta_1(1 - k\gamma)}{e^{k\gamma} - 1}; \gamma = \frac{\pi_1}{\pi_0}; i = 1, 2, \dots$$

O sistema para um só hospital é baseado em  $n$  agrupamentos consecutivos. Um alarme será disparado por uma seqüência de agrupamentos se cada um dos agrupamentos estiver abaixo de um certo tamanho.

As probabilidades sob as hipóteses nula e alternativa, respectivamente, que uma certa seqüência disparará um alarme dependem de  $n$ ,  $k$  e  $\gamma$ , como abaixo:

$$P_r[X_{\max} \leq k\eta_0] = (1 - e^{-k})^n \cong P_0$$

$$e P_r[X_{\max} \leq k\eta_0] = (1 - e^{-k\gamma})^n \cong P_1,$$

onde  $X_{\max}$  é o tamanho do maior agrupamento na seqüência.  $P_0$  é determinado por  $\lambda$ ; o número de alarmes falsos que se espera durante um dado intervalo de tempo em que se esperam, por exemplo,  $\beta$  nascimentos. São, então, estabelecidas as relações entre  $P_0$ ,  $\lambda$  e  $\beta$ .

Para um sistema amplo que cubra uma grande área, compreendendo hospitais distintos, o sistema proposto se baseia nos mesmos dados registrados em uma única unidade de pequena escala. Nesse sistema amplo, a análise de dados será feita ao final de intervalos de tempo constantes, por exemplo, a cada dois meses. O número e o tamanho dos agrupamentos completados dentro desse intervalo são registrados. Será disparado um alarme se cada agrupamento completado for menor que um certo valor, digamos  $k_m \cdot \eta_0$ , onde  $m$  é o número de agrupamentos completados. Se  $q_0$  é a probabilidade de um falso alarme, sob a hipótese nula, isto é, dado que pelo menos um agrupamento foi completado,  $k_m$  está relacionado a  $q_0$  pela relação  $q_0 = (1 - e^{-k_m})^m$  e  $k_m = -\ln(1 - q_0^{1/m})$ .

Se o número de alarmes falsos deve ser, por exemplo, um em  $a$  anos e a análise é conduzida  $b$  vezes por ano, então:  $P_0 \cdot a \cdot b = 1$ .

É possível mostrar, ainda, que se o número  $N$  de nascimentos é grande, nessa área ampla sob vigilância, então  $q_0 \cong P_0$ .

São apresentadas ilustrações com exemplos numéricos para cada um dos sistemas considerados.

Uma adaptação dos procedimentos anteriores para o monitoramento de doenças crônicas que pode, por analogia, ser utilizada em áreas restritas ou em áreas ampliadas foi proposta por Chen et al. (1982). Ambos os sistemas se baseiam, também, na observação dos intervalos de tempo consecutivos entre diagnósticos da doença objeto de vigilância.

### O procedimento de Tango e um procedimento alternativo

Tango (1984) propõe um índice  $C$  para mensurar o nível de conglomerados no tempo e conglomerados cíclicos no tempo, que é dado por uma forma quadrática  $C = r'Ar$ , considerando  $f = (N_1, N_2, \dots, N_m)$  o vetor aleatório de frequências observadas em  $m$  sucessivos intervalos de tempo de amplitudes iguais, com frequência total  $N = \sum N_i$ ; considerando a frequência relativa

$$r_i = \frac{N_i}{N},$$

então  $r = (r_1, r_2, \dots, r_m)$  é o vetor de frequências relativas e  $A = (a_{ij})$ ;  $i, j = 1, 2, \dots, m$  é uma matriz, pré-definida, de proximidade entre o  $i$ mo e o  $j$ mo intervalos de tempo.

A distribuição assintótica do índice de conglomerados gerou alguns questionamentos discutidos por Whittmore & Keller (1986), Whittmore et al. (1987), Tango (1990) e Rayens & Kriscio (1993), estes últimos afirmando terem resolvido a controvérsia. Tanto Tango quanto Rayens & Kriscio ilustram a aplicação do procedimento utilizando os mesmos dados sobre trisomias utilizados por Wallestein (1980).

Um procedimento alternativo ao de Tango e que se baseia em um teste de *goodness of fit* de uma distribuição uniforme discreta, usando a estatística  $X^2$  de Pearson, foi proposto por Best & Rayner (1991). Mais especificamente, este procedimento consiste no exame das componentes da estatística  $X^2$ , mais particularmente a primeira e a segunda, de cujos valores se pode obter indicações sobre conglomerados no tempo ou conglomerados cíclicos.

### Um método analítico-gráfico em uso no CDC para doenças de notificação compulsória

Segundo Stroup et al. (1993), o CDC incorporou, para a apresentação rotineira de dados sobre doenças de notificação compulsória, no *Morbidity and Mortality Weekly Report (MMWR)*, um procedimento gráfico de barras proposto por Stroup et al. (1989).

Como os dados registrados podem variar em função de uma série de fatores não relacionados à doença, os autores os agrupam em um período de quatro semanas, denominado genericamente de mês. Denotando por  $x_0$  o número de casos de uma dada doença registrado no período de um mês terminando na semana corrente, este valor é comparado com um valor básico de referência construído a partir dos quinze totais prévios,  $x_2$ , tomados do correspondente mês, em análise, e dos dois meses vizinhos (anterior e posterior), em cada um dos

cinco anos anteriores. É assumido, mesmo na presença de variação sazonal, que os  $x_i$ ;  $i = 1, 2, \dots, 15$  são variáveis aleatórias independentes e identicamente distribuídas.

Os resultados são apresentados em termos de razão, digamos  $r$ , do valor  $x_0$  dividido por alguma medida de tendência central com base nos quinze valores tomados como referência. Como medidas de tendência central, foram experimentadas a média aritmética e a mediana.

A construção do diagrama de barras é simples: utiliza-se um eixo horizontal, em escala logarítmica e um eixo vertical lançado do ponto 1 (um) do eixo horizontal. Sobre esse eixo vertical, para cada doença em foco, é construído um retângulo, ou barra, cuja base se assenta no eixo vertical e cuja altura é o valor de  $r$ . Os retângulos são dispostos à direita ou à esquerda do eixo vertical, respectivamente, se a razão é maior ou menor que um. Intervalos de confiança, baseados na teoria da distribuição Normal, são construídos para  $r$ , e a altura do retângulo que supera tal limite de confiança é representada em destaque, por exemplo, por hachurado ou negrito.

Como ilustração, considere uma doença hipotética  $D1$  para a qual registrou-se, no período de 4 semanas do mês de março de 1999, o número  $x_0 = 64$  casos.

A Tabela 3 apresenta o número de casos dessa doença ocorridos nos meses de fevereiro, março e abril, nos cinco anos anteriores.

Da tabela temos:  $\bar{x} = 34,80$  casos/mês;  $S_{\bar{x}} = 0,83$  casos/mês

$$\text{Então: } r = \frac{x_0}{\bar{x}} = 1,85 > 1,00.$$

Como  $r > 1$ , interessa determinar (LS), o limite superior do intervalo de confiança para  $r$ . Fixando em 95% o nível de confiança, resulta (LS)  $\cong 1,05$ . Portanto  $r$  excede (LS) em 0,80, indicando que  $x_0$  supera, significativamente, os valores históricos.

Tabela 3

Número de casos da doença  $D1$ , nos meses de fevereiro, março e abril, nos anos de 1994 a 1998\*.

	Fevereiro	Março	Abril
1998	35	40	37
1997	32	38	31
1996	37	35	39
1995	29	36	39
1994	35	34	30

\* Dados hipotéticos.

A Figura 5 ilustra o procedimento gráfico descrito, onde  $D_2$ ,  $D_3$ ,  $D_4$ , ... seriam outras possíveis doenças submetidas à análise, para as quais se teria seguido o mesmo procedimento.

Também com base na teoria da distribuição Normal, porém distinto do procedimento anterior, e também utilizando um sistema de barras que os autores denominam Cartas de Avaliação, muito utilizado em monitoramento em saúde pública no Japão, Arantes et al. (1997) fazem a avaliação de serviços através do monitoramento de 30 indicadores epidemiológicos operacionais de programas antituberculose, em 62 unidades de saúde.

#### Um procedimento de memória curta para detectar um aumento na média de uma distribuição de Poisson

Shore & Quade (1989) propõem um procedimento para detectar aumento súbito na taxa de ocorrência de eventos raros, baseando-se somente em dados registrados mais recentemente, daí o nome de memória curta. O número de casos em uma dada semana, digamos a semana corrente, é comparado com o número total de casos ocorridos nas  $s$  semanas prévias;  $s$  é o parâmetro que representa a memória do sistema.

Apesar do desenvolvimento do procedimento considerar  $s$  como uma amostra no tempo e utilizar as hipóteses de nulidade e alterna-

tivas no sentido clássico de testes de hipóteses, elas são, realmente, adaptações que se reduzem a hipóteses de aleatoriedade. São estabelecidas condições para o cálculo do ARL sob condições epidêmicas e não-epidêmicas, e a performance do procedimento é comparada àquela do procedimento *Cusum*, partindo de dados simulados.

#### Uma proposta para a detecção de conglomerados espaciais de doenças raras

Besag & Newel (1991) propõem testes de aleatoriedade para detectar possíveis conglomerados espaciais de uma doença rara em uma região extensivamente grande, subdividida em pequenas regiões ou zonas administrativas. O procedimento é desenvolvido, inicialmente, para pequenas regiões, tais como distrito de enumeração censitária, ou setor censitário, para as quais existam dados censitários disponíveis e, então, é generalizado para uma ampla região que cubra tais regiões administrativas.

Se não há mapas digitados dos distritos de enumeração censitária, trabalha-se com as coordenadas de seus centróides. Cada caso é examinado por vez, e um teste de hipóteses simples é realizado para decidir se o centróide associado constitui o centro de um conglomerado de um certo tamanho pré-fixado.

#### **Comparações entre procedimentos estatísticos na análise de dados de vigilância em saúde pública**

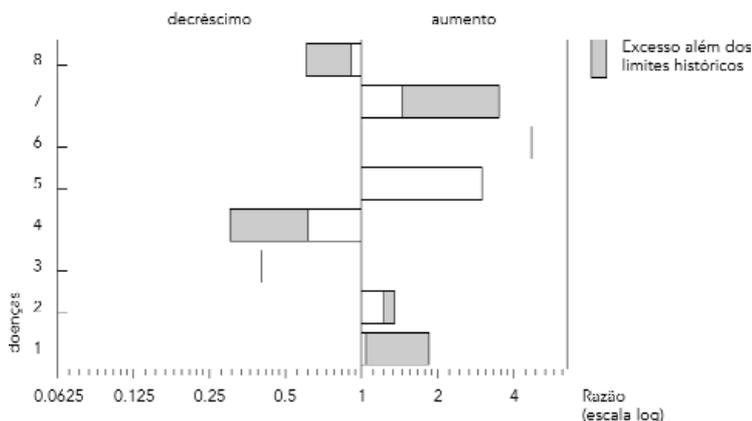
Na teoria clássica de testes de hipóteses, que exige amostragem aleatória, em geral no sentido de amostragem casual simples, o procedimento usual para comparação de distintos testes de hipóteses com o mesmo objetivo é utilizar a função poder do teste, que depende, entre outros fatores, do tamanho da amostra.

Na vigilância em saúde pública, dada a natureza do processo de coleta de dados, como já foi acentuado, não existe o processo de amostragem aleatória, nem alocação aleatória. Há, então, dificuldades de fundamentação teórica para estabelecer comparações entre métodos estatísticos, para análise de dados de vigilância, com apelo ao uso do poder do teste.

Nesse sentido, Chen (1987) considera que: “desde que o controle de qualidade e o monitoramento de doenças envolvem um número indefinido de análises, o resultado significativo, denominado alarme, neste contexto, é limitado à ocorrência aleatória, ou não, da taxa de itens defeituosos ou da taxa de novos casos diagnosti-

Figura 5

Diagrama de barras em uso no C.D.C.\*



\* Dados hipotéticos.

*cados da doença sob vigilância, então o poder do teste, em sentido estrito, é destituído de significado para a comparação entre procedimentos”.*

Assim, para procedimentos estatísticos na análise de dados de vigilância em saúde pública, considerando a natureza do processo de coleta de dados, outros meios têm sido utilizados, que não excluem possíveis generalizações do procedimento clássico.

Nessa linha, Frizén (1992) propõe quatro medidas para comparar o comportamento de procedimentos distintos que têm um sistema de alarme associado: o ARL; a probabilidade de falso alarme; a probabilidade de detecção, com sucesso, de um alarme verdadeiro; e o valor preditivo de um alarme.

A partir da definição dessas medidas e utilizando simulações em computador, a autora compara três procedimentos para vigilância em saúde pública: o de Shewhart, o *Cusum* monocaudal e o *Cusum* bicaudal de máscara em V.

Não existe, nesse trabalho, definitivamente, uma conclusão de supremacia de um procedimento em relação ao outro, porque essa supremacia é dependente das características da doença em particular ou condição sob monitoramento e dos valores propostos para os parâmetros associados às medidas, em função dos quais são estabelecidas as simulações. No entanto, a autora faz uma advertência, no final: “*Em alguns casos, como no monitoramento dos batimentos cardíacos de um feto, é muito importante ter um bom método e saber quão bom ele é. Deve-se saber muito de quanto se pode confiar em um alarme disparado e como interpretá-lo. Nestas situações, várias medidas como aquelas aqui descritas são necessárias*” (Frizén, 1992).

Duas técnicas para confirmar ou rejeitar um alarme, a partir de dados de monitoramento de incidência de câncer renal, são propostas em um estudo desenvolvido por Chen et al. (1993). É possível afirmar que o mesmo não é conclusivo, porque é muito difícil, no final, fazer um balanço entre as vantagens e desvantagens dessas técnicas.

Barbujani & Calzolari (1984) fazem uma comparação entre o procedimento *Cusum* e o procedimento de agrupamento de Chen, no estudo de máis-formações congênitas. São estabelecidas simulações em computador, usando o método Monte Carlo, a fim de verificar se alarmes disparados são falsos, verdadeiros, ou se nenhum alarme é disparado em um período de 49 meses de simulação. No final, os autores concluem pela maior eficiência do procedimento *Cusum*.

Chen (1987), também utilizando simulações em computador, faz o mesmo tipo de comparação. A medida de eficiência utilizada é o tempo médio de atraso para detectar um alarme verdadeiro. A conclusão é que o procedimento de agrupamentos deve ser preferido na situação de epidemia dita de baixo nível, isto é, quando a taxa básica anual esperada não é maior que cinco; caso contrário, a preferência deve ser pelo procedimento *Cusum*.

Nix et al. (1987) discutem comparações entre métodos estatísticos de controle interno em química clínica. O critério de avaliação se baseia na função de distribuição de probabilidades do ARL em AQL (*Acceptable Quality Level*) e em RQL (*Rejectable Quality Level*) e em um critério de otimização para o procedimento de controle especificado. As comparações são feitas entre os procedimentos de Shewhart, Shewhart modificado e *Cusum*. A conclusão, segundo os critérios utilizados, é que o procedimento *Cusum* é potencialmente o melhor dos três.

No trabalho de Stroup et al. (1989), as comparações são entre o procedimento analítico, proposto por eles, com a estatística *scan* e um terceiro procedimento, baseado na técnica *bootstrap*, utilizando somente os dados observados, coletados no CDC e relativos a onze doenças de notificação compulsória.

Critérios de avaliação da qualidade de métodos analíticos utilizando somente dados observados são passíveis de crítica e levam a resultados discutíveis. Aliás, isso deve ter sido reconhecido pelos autores, pois, no final do trabalho, anunciam que um estudo de simulação está sendo planejado para avaliar a sensibilidade do método. Além do mais, como já assinalamos anteriormente, também é passível de crítica a construção de intervalos de confiança baseados no pressuposto da validade da teoria Normal de distribuição de probabilidades, usando um conjunto de observações oriundas da coleta de dados de vigilância como amostra aleatória.

Shore & Quade (1989) fazem um estudo comparativo entre o esquema de memória curta e o procedimento *Cusum*, partindo de dados tabelados por Ewan & Kemp (1960) para o ARL de uma variável Poisson. Concluem que o procedimento *Cusum* é mais eficiente quando a amplitude da memória é pequena, mas, mesmo em condições assintóticas para a amplitude da memória, o procedimento *Cusum* parece ser mais eficiente.

Ainda, com base em simulações em computador, Rayens & Kriscio (1993) comparam o procedimento de Tango com outros três, que

rivalizam com aquele: o procedimento EMM (Ederer et al., 1964), que é considerado uma versão discreta da estatística *scan*; a própria estatística *scan*; e o teste  $\chi^2$  de *goodness of fit*, de uma distribuição de probabilidades. O critério utilizado é aquele de comparar poder de teste, em relação ao qual já foi estabelecida observação crítica. As simulações indicam que a maior ou menor eficiência de um procedimento é dependente dos particulares intervalos de tempo ou células onde ocorrem os conglomerados.

Royston (1991), comparando métodos ditos de limiar, métodos Bayesianos para *change point* e o procedimento *Cusum*, na área de fertilidade humana, opta pelo procedimento *Cusum*, e Santos (1997), comparando o procedimento Shewhart com o *Cusum*, via simulações em computador, também define o procedimento *Cusum* como mais eficiente.

Comparações entre procedimentos de controle de qualidade, com esse enfoque exclusivo, são referenciadas na literatura específica.

### Considerações finais

Intencionalmente, métodos de análise para detecção de conglomerados espaciais utilizando mapas geográficos, procedimentos analíticos que utilizam técnicas de regressão e técnicas de séries temporais não foram abordados nesta revisão.

A razão para a não citação dos primeiros é que, hoje, com a evolução de programas para produção de mapas computadorizados, existe uma ampla literatura específica, internacional e brasileira, o que, por si só, exigiria uma revisão específica, por especialista da área.

Quanto aos métodos de regressão e de séries temporais, estes têm sido utilizados mais com o objetivo de predição e estimação, no âmbito da vigilância da influenza e da pneumonia, no qual a detecção de conglomerados é feita de modo indireto: calculando o excesso de mortalidade atribuível a essas doenças (Collins & Lehmann, 1951; Eickhoff et al., 1962; Serfling, 1963; Clifford et al., 1977; Alling et al., 1981; Choi & Thacker, 1981; Tillett et al., 1983). Especificamente, no que diz respeito ao uso de séries temporais, é bem conhecido o poder e a flexibilidade de tais métodos, no entanto o seu uso na área de saúde pública, de um modo geral, depende ainda de estudos e refinamentos desses instrumentos por parte dos estatísticos vinculados a essa área (Thacker, 1989).

A análise comparativa, no item anterior, evidencia que, mesmo para aqueles procedimentos em que existem, com mais frequência,

indicações para a sua aplicação no âmbito da vigilância em saúde pública, há, ainda, necessidade de mais investigação, pelo menos, pela ótica de aplicação à prática.

Também, métodos gráficos na linha exposta por Tukey (1977), Tufte (1983) e Cleveland (1985), apesar do grande número de programas computacionais gráficos disponíveis no mercado, parecem não ter merecido a devida atenção por parte dos responsáveis pela análise dos dados de vigilância em saúde pública.

Quando falamos em vigilância em saúde pública, a atenção geral parece estar voltada para as denominadas doenças de notificação compulsória. Mas, atualmente, existe um conceito de vigilância ampliada, em que outras situações ou eventos são objeto da vigilância, como, por exemplo, vigilância de pessoas, vigilância de pacientes internados, vigilância de doenças crônicas, vigilância de fatores de risco ambiental, vigilância de efeitos adversos de drogas após sua liberação para comercialização, etc., para os quais os procedimentos aqui revistos são também aplicáveis.

A despeito de todas as limitações dos dados coletados no processo de vigilância em saúde pública, quer nos sistemas em nível nacional ou em níveis regionais, os dados podem ser, e em geral são, úteis para evidenciar tendências ou alterações bruscas que apontam para a necessidade de ações de intervenção da saúde pública. O importante é ter em mente o alerta seguinte.

Os profissionais responsáveis pela condução das análises estatísticas rotineiras de dados de vigilância devem selecionar o procedimento a ser utilizado, de acordo com as circunstâncias analisadas, como, por exemplo, a natureza do evento, o tipo de dados disponíveis sobre casos diagnosticados, a viabilidade do uso de tal procedimento em função de seus pressupostos teóricos formais e, como consequência, conhecer as potencialidades e limitações do procedimento. Isso permitirá decidir sobre a maior ou menor eficiência do mesmo em relação aos resultados esperados com a sua utilização.

O CDC ressaltou a década de 80 como sendo a da estatística, naquele órgão, em virtude da contribuição dada pelo seu corpo de profissionais estatísticos. No Brasil, por outro lado, pelo menos até o ano de 1995, nos sistemas governamentais de vigilância em saúde pública, seja em nível federal, estadual ou municipal, não havia estatísticos oficialmente vinculados a tais sistemas, nem mesmo através de consultorias ou assessorias, periódicas ou esporádicas (Sanches et al., 1995).

Não há receio ou risco de erro ao se afirmar que, em nosso país, existe um campo de trabalho imenso a ser explorado, no contexto dos procedimentos estatísticos aplicados à análise de dados de vigilância em saúde pública, quer no sentido restrito ou no ampliado. O sucesso dessa tarefa depende de uma maior integração, nos serviços de vigilância, entre epidemiologistas, sanitaristas e estatísticos, estes com expe-

riência nas áreas de epidemiologia e saúde pública; integração esta que parece, ainda, nem mesmo existir. Para tanto, é necessário que as coordenações desses serviços entendam que o papel dos profissionais em estatística não deve, e não pode, ser substituído por outros profissionais que conhecem e sabem manipular um ou mais programas computadorizados de estatística.

## Agradecimentos

Ao CNPq, processos nºs 511716/93-4 e 300035/89-6 (NV); ao Diretor do Instituto de Saúde da Secretaria Estadual da Saúde de São Paulo, Dr. José da Rocha Carvalheiro / FAPESP (Processo nº 1995/3762-9), que, com seus apoios financeiros, cada um a seu tempo, conduziram a este trabalho.

Ao editor desta revista pela sugestão do tema e ao *referee* pelas sugestões apropriadas que contribuíram para uma melhor apresentação do trabalho.

## Referências

- ALLING, D. W.; BLACKWELDER, W. C. & STUART-HARRIS, C. H., 1981. A study of excess mortality during influenza epidemics in the United States, 1968-1976. *American Journal of Epidemiology*, 113:30-43.
- ALWAN, L. C., 1986. Cusum quality control-multivariate approach. *Communication in Statistics - Theory and Methods*, 15:3531-3543.
- ARANTES, G. R.; BELLUOMINI, M.; ALMEIDA, M. M. B. & LIMA, M. M., 1997. Monitoração das ações antituberculose: Cartas gráficas baseadas em unidades desvio padrão e tabelas de classificação baseadas em média-síntese. *Medicina*, 2:243-258.
- ASSAF, D.; POLLACK, M. & RISTOV, Y., 1992. A new look and action lines of surveillance schemes. *Journal of the American Statistical Association*, 87:889-895.
- BAGSHAW, M. & JOHNSON, R. A., 1975. The effect of serial correlation on the performance of cusum tests II. *Technometrics*, 1:73-80.
- BARBUJANI, G. & CALZOLARI, E., 1984. Comparison of two statistical techniques for the surveillance of births defects through a Monte Carlo simulation. *Statistics in Medicine*, 3:239-247.
- BARNARD, G. A., 1959. Control charts and stochastic process (with discussion). *Journal of the Royal Statistical Society*, 2:239-257.
- BESAG, J. & NEWELL, J., 1991. The detection of clusters in rare diseases. *Journal of the Royal Statistical Society*, 154A:143-155.
- BEST, D. J. & RAYNER, J. C. W., 1991. Disease clustering in time. *Biometrics*, 47:589-593.
- BJERKEDAL, T. & BAKKETEIG, L., 1975. Surveillance of malformations and other conditions of newborn. *International Journal of Epidemiology*, 4: 31-36.
- BROOKS, D. & EVANS, D. A., 1972. An approach to the probability of cusum run length. *Biometrika*, 59: 539-549.
- CDC (Centers for Disease Control and Prevention), 1986. *Comprehensive Plan for Epidemiological Surveillance*. Atlanta: CDC.
- CDC (Centers for Disease Control and Prevention), 1989. Proposed changes in format for presentation of notifiable disease report data. *MMWR*, 38: 805-809.
- CDC (Centers for Disease Control and Prevention), 1990. Guideline for investigating clusters of health events. *MMWR*, 39:1-23.
- CHEN, R., 1978. A surveillance system for congenital malformations. *Journal of the American Statistical Association*, 73:323-327.
- CHEN, R., 1987. The relative efficiency of the sets and the cusum techniques in monitoring the occurrence of a rare event. *Statistics in Medicine*, 6:517-525.
- CHEN, R., 1988. The relative efficiency of the sets and the cusum techniques in monitoring the occurrence of a rare event. *Statistics in Medicine*, 7:998-999.
- CHEN, R.; CONNELLY, R. R. & MANTELL, N., 1993. Analyzing post-alarm data in a monitoring system in order to accept or reject the alarm. *Statistics in Medicine*, 12:1807-1812.
- CHEN, R.; MANTEL, N.; CONNELLY, R. R. & ISACSON, P., 1982. A monitoring system for chronic diseases. *Methods of Information in Medicine*, 21:86-90.
- CHOI, K. & THACKER, S. B., 1981. An evaluation of influenza mortality surveillance: 1962-1979. Time series forecasts of expected pneumonia and influenza deaths. *American Journal of Epidemiology*, 113:215-226.

- CLEVELAND, W. S., 1985. *The Elements of Graphing Data*. New Jersey: Murrey Hill/ Bell Telephone Laboratories.
- CLIFFORD, R. E.; SMITH, J. W. G.; TILLET, H. E. & WHERRY, P. J., 1977. Excess mortality associated with influenza in England and Wales. *International Journal of Epidemiology*, 6:115-128.
- COLLINS, S. D. & LEHMANN, J., 1951. Trends and epidemics of influenza and pneumonia. *Public Health Report*, 66:1487-1516.
- COLTON, T., 1989. Remarks. *Statistics in Medicine*, 8:255-256.
- CROISER, R. B., 1988. Multivariate generalization of cumulative sum quality control schemes. *Technometrics*, 30:291-303.
- DE BRUYN, C. S. V., 1968. *Cumulative Sums Tests: Theory and Practice*. London: Griffin's Statistical Monographs & Courses.
- DUNCAN, A. J., 1974. *Quality Control and Industrial Statistics*. Homewood: Irwin Inc.
- EDERER, F.; MYERS, M. & MANTEL, N., 1964. A statistical problem in space and time: Do leukemia cases come in clusters? *Biometrics*, 20:626-637.
- EDWARDS, R. W., 1980. Internal analytical quality control using the cusum chart and truncated. V-Mask procedures. *Annals of Clinical Biochemistry*, 17:205-211.
- EICKHOFF, T. C.; SHERMAN, I. L. & SERFLING, R. E., 1962. Observations of excess mortality associated with epidemic influenza. *JAMA*, 76:776-782.
- EWAN, W. D. & KEMP, K. W., 1960. Sampling inspection of continuous process with no autocorrelation between successive results. *Biometrika*, 4:363-380.
- FRISÉN, M., 1992. Evaluation of methods for statistical surveillance. *Statistics in Medicine*, 11:1489-1502.
- FRISÉN, M. & DE MARÉ, J., 1991. Optimal surveillance. *Biometrika*, 78:271-280.
- GALLUS, G., 1988. The relative efficiency of the sets and the cusum techniques in monitoring the occurrence of a rare event. *Statistics in Medicine*, 7:997-998.
- GAN, F. I., 1991. An optimal design of cusum quality control charts. *Journal of Quality Technology*, 4:279-286.
- GLAZ, J., 1993. Approximations for the tail probabilities of the scan statistic. *Statistics in Medicine*, 12:1846-1852.
- GOEL, A. L. & WU, S. M., 1971. Determination of ARL and a contour nomograms for cusum charts to control normal mean. *Technometrics*, 13:221-230.
- HALL, D. B., 1989. Analysis of surveillance data: A rationale for statistical tests with comments on confidence intervals and statistical models. *Statistics in Medicine*, 8:273-278.
- HILL, G. B.; SPICER, C. R. & WEATHERALL, J. C., 1968. The computer surveillance of congenital malformations. *British Medical Bulletin*, 24:215-218.
- HINKLEY, D. V., 1971. Inference about the change-point from cumulative sums tests. *Biometrika*, 3:509-523.
- HUNTINGTON, R. & NAUS, J. I., 1975. A simple expression for the  $k^{\text{th}}$  nearest neighbor coincidence probabilities. *Annals of Probability*, 3:894-896.
- JOHNSON, R. A. & BAGSHAW, M., 1974. The effect of serial correlation on the performance of cusum tests I. *Technometrics*, 1:103-112.
- KAFADAR, K. & STROUP, D. F., 1992. Analysis of aberrations in public health surveillance data: Estimating variances on correlated samples. *Statistics in Medicine*, 11:1551-1568.
- KEMP, K. W.; NIX, A. B. J.; WILSON, D. W. & GRIFFINS, K., 1978. Internal quality control of radioimmunoassay. *Journal of Endocrinology*, 76:203-210.
- KISH, L., 1965. *Survey Sampling*. New York: John Wiley & Sons.
- LEHMANN, E. L., 1983. *Theory of Point Estimation*. New York: John Wiley & Sons.
- LEHMANN, E. L., 1986. *Testing Statistical Hypotheses*. New York: John Wiley & Sons.
- LEVEY, S. & JENNINGS, E. R., 1950. The use of control charts in the clinical laboratories. *American Journal of Clinical Pathology*, 20:1059-1066.
- LEVIN, B. & KLINE, J., 1985. The cusum test of homogeneity with an application in spontaneous abortion epidemiology. *Statistics in Medicine*, 4:469-488.
- LUCAS, J. M., 1982. Combined Shewhart-cusum quality control schemes. *Journal of Quality Technology*, 14:51-59.
- MASON, J. O., 1989. Opening remarks – Symposium on statistics in surveillance. *Statistics in Medicine*, 8:253-254.
- MONTGOMERY, D. C., 1980. The economic design of control charts: A review and literature survey. *Journal of Quality Technology*, 2:75-86.
- MONTGOMERY, D. C., 1990. *Introduction to Statistical Quality Control*. New York: John Wiley & Sons.
- NAGENDRA, Y. & RAI, G., 1971. Optimum sample size and sampling interval for controlling the mean of a non-normal variables. *Journal of the American Statistical Association*, 66:637-640.
- NAUS, J. I., 1965. The distribution of the size of the maximum cluster of points on a line. *Journal of the American Statistical Association*, 60:532-538.
- NAUS, J. I., 1966a. Some probabilities expectations and variances of the largest clusters and smallest intervals. *Journal of the American Statistical Association*, 61:1191-1199.
- NAUS, J. I., 1966b. A power comparison of two tests on nonrandom clustering. *Technometrics*, 8:493-517.
- NCUBE, M. M. & WOODALL, W. H., 1984. A combined Shewhart cumulative score quality control chart. *Applied Statistics*, 33:259-265.
- NIX, A. B. J.; ROWLANDS, R. J.; KEMP, K. W.; WILSON, D. W. & GRIFFITHS, K., 1987. Internal quality control in clinical chemistry: A teaching review. *Statistics in Medicine*, 6:425-440.
- NORTH, W. R. S., 1982. The quangle: A modification of the cusum chart. *Applied Statistics*, 28:126-135.
- PAGE, E. S., 1954. Continuous inspection schemes. *Biometrika*, 41:100-115.
- PAGE, E. S., 1957. On problems in which a change in a parameter occurs to an unknown point. *Biometrika*, 44:248-252.
- PETTITT, A. N., 1979. A non-parametric approach to the change-point problem. *Applied Statistics*, 2:126-135.
- PETTITT, A. N., 1980. A sample cumulative sum type statistic for the change-point problem with zero-one observations. *Biometrika*, 67:79-84.
- PIGNATIELLO Jr., J. J. & RUNGER, G. C., 1990. Comparisons of multivariate cusum charts. *Journal of Quality Technology*, 22:173-186.

- RAUBERTAS, R. F., 1989. An analysis of disease surveillance data that uses the geographic locations of the reporting units. *Statistics in Medicine*, 8:267-271.
- RAYENS, M. K., & KRISCIO, R. J., 1993. Properties of Tango's index for detecting clustering in time. *Statistics in Medicine*, 12:1813-1827.
- REINKE, W. A., 1991. Applicability of industrial sampling techniques to epidemiologic investigations: Examination of an underutilized resource. *American Journal of Epidemiology*, 134:1222-1232.
- RICH, W. H., 1946. The industrial "control-chart" applied to the study of epidemics. *Public Health Report*, 61:1501-1511.
- ROWLANDS, R. J.; GRIFFITHS, K.; KEMP, K. W.; NIX, A. B. J.; RICHARDS, G. & WILSON, D. W., 1983. Application of cusum techniques to the routine monitoring of analytical performance in clinical laboratories. *Statistics in Medicine*, 2:141-145.
- ROWLANDS, R. J.; WILSON, D. W.; NIX, A. B. J.; KEMP, K. W. & GRIFFITHS, K., 1980. Advantages of the cusum techniques for quality control in clinical chemistry. *Clinica Chimica Acta*, 108:393-397.
- ROYSTON, J. P., 1991. Identifying the fertile phase of the human menstrual cycle. *Statistics in Medicine*, 10:221-240.
- ROYSTON, J. P. & ABRAMS, R. M., 1980. An objective method for detecting the shift in basal body temperature in women. *Biometrics*, 36:217-224.
- SAHU, S. K.; BENDEL, R. B. & SISON, C. P., 1993. Effect on relative risk and cluster configuration on the power of the one-dimensional scan statistic. *Statistics in Medicine*, 12:1853-1865.
- SANCHES, O., 1993. Análise rotineira de dados de vigilância em saúde pública: Que procedimentos estatísticos utilizar? *Revista de Saúde Pública*, 27:300-304.
- SANCHES, O.; IGUCHI, T.; AMARANTE, C. M.; SANTOS, M. O. & OLIVEIRA, J. F. P., 1995. "Survey" nacional sobre o funcionamento do sistema de vigilância em saúde pública para coleta e análise de dados, em nível federal, estadual e municipal. In: 3º Congresso Brasileiro de Epidemiologia, *Resumos*, p. 79. Salvador: ABRASCO/SIAE/ALAMES.
- SANTOS, M. O., 1997. *A Utilização da Técnica de Somas Cumulativas (cusum) na Análise de Dados de Vigilância em Saúde Pública*. Dissertação de Mestrado, Rio de Janeiro: Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz.
- SERFLING, W., 1944. Analysis of the data of public health organization by the control chart method. *Journal of the American Statistical Association*, 39:311-324.
- SERFLING, R. E., 1963. Methods for current statistical analysis of excess pneumonia - Influenza deaths. *Public Health Report*, 78:494-506.
- SHEWHART, W. A., 1931. *Economic Control of Quality of Manufactured Product*. New York: MacMillan.
- SHORE, D. L. & QUADE, D., 1989. A surveillance system based on short memory scheme. *Statistics in Medicine*, 8:311-322.
- STROUP, D. F.; WHARTON, M.; KAFADAR, K. & DEAN, A. G., 1993. Evaluation of a method for detecting aberrations in public health surveillance data. *American Journal of Epidemiology*, 137:373-380.
- STROUP, D. F.; WILLIAMSON, G. D. & HERSON, J. L., 1989. Detection of aberrations in the occurrence of notifiable diseases in surveillance data. *Statistics in Medicine*, 8:323-329.
- TANGO, T., 1984. The detection of disease clustering in time. *Biometrics*, 40:15-20.
- TANGO, T., 1990. Asymptotic distribution of an index for disease clustering. *Biometrics*, 45:351-357.
- THACKER, S. B., 1989. Commentary. *Statistics in Medicine*, 8:363.
- THACKER, S. B. & BERKELMAN, R. L., 1988. Public health surveillance in the United States. *Epidemiological Reviews*, 10:164-190.
- THACKER, S. B.; BERKELMAN, R. T. & STROUP, D. F., 1989. The science of public health surveillance. *Journal of Public Health Policy*, 10:187-203.
- THOMPSON, S. G., 1983. A method of analysis of laboratory data in epidemiological study where time trends are present. *Statistics in Medicine*, 2:147-153.
- TILLET, H. E.; SMITH, J. W. G. & GOOCH, C. D., 1983. Excess deaths attributable to influenza in England and Wales: Age at death and certified cause. *International Journal of Epidemiology*, 12:344-352.
- TUFTE, E. R., 1983. *The Visual Display of Quantitative Information*. Cheshire: Graphics Press.
- TUKEY, J. W., 1977. *Exploratory Data Analysis*. Reading: Addison Wesley Press.
- WALLESTEIN, S., 1980. A test for detection of clustering over time. *American Journal of Epidemiology*, 111:367-372.
- WALLESTEIN, S.; GOULD, M. S. & KLEINMAN, M., 1989. Use of the scan statistic to detect time-space clustering. *American Journal of Epidemiology*, 130:1057-1065.
- WALLESTEIN, S. & NAUS, J. I., 1973. Probabilities for the  $k$ th nearest neighbor problem on the line. *Annals of Probability*, 1:188-190.
- WALLESTEIN, S.; NAUS, J. I. & GLAZ, J., 1993. Power of the scan statistic for detection of clustering. *Statistics in Medicine*, 12:1829-1843.
- WEATHERALL, J. A. C. & HASKEY, J. C., 1976. Surveillance of malformations. *British Medical Bulletin*, 32:39-40.
- WERNECK, G. L. & STRUCHINER, C. J., 1997. Estudos de agregados de doenças no espaço-tempo: Conceitos, técnicas e desafios. *Cadernos de Saúde Pública*, 13:611-624.
- WESTGARD, J. O.; BARRY, P. L. & HUNT, M. R., 1981. A multiple rule Shewhart chart for quality control in clinical chemistry. *Clinical Chemistry*, 27:493-501.
- WESTGARD, J. O.; GROTH, T.; ARONSON, T. & VERDIER, C. H., 1977. Combined Shewhart-CUSUM control chart for improved quality control in clinical laboratory. *Clinical Chemistry*, 23:1881-1887.
- WHITTEMORE, A. S.; FRIEND, N.; BROWN, B. W. & HOLLY, E. A., 1987. A test to detect clusters of disease. *Biometrika*, 74:631-635.
- WHITTEMORE, A. S. & KELLER, J. B., 1986. On Tango's index for disease clustering in time. *Biometrics*, 42:118.
- WILSON, D. W.; GRIFFITHS, K.; KEMP, K. W.; NIX, A. B. J. & ROWLANDS, R. J., 1979. Internal quality control of radioimmunoassay: Monitoring of error. *Journal of Endocrinology*, 80:365-372.
- YASHCHIN, E., 1992. Analysis of cusum and other Markov-type control schemes by using empirical distributions. *Technometrics*, 34:54-63.