Statistical Process Monitoring in Disease Surveillance

Karen Kafadar

Department of Mathematics

University of Colorado at Denver

kk@math.cudenver.edu

`http://www-math.cudenver.edu/ kk`

OUTLINE

1. Principles of SPC/SQC
2. Detecting outbreaks, aberrations, trends
3. Data types and time periods
4. Control charts: Shewhart, EWMA, CUSUM
5. Sample data from CDC
6. Time series modeling
7. Final thoughts

1. Principles of SPC/SQC

Western Electric-AT&T *Statistical Quality Control Handbook*:

"SQC is a scientific method of analyzing data and using the analysis to solve practical problems"

"Quality": Desirable characteristics of a product or process (sorting good from bad is inadequate: Need to focus on process)

"Control": Keep process within boundaries so it behaves in a predictable fashion

- minimize variation in output
- meet customer expectations (define "customer")
- institute procedures for continual improvement

SPC steps and tools

Steps	Tools
1. Understand process	Flowcharts
2. Define process measures	Surveys
(customer-oriented)	Gage R&R Studies
3. Collect/summarize data	Exp't design/analysis
4. Process monitoring	Control charts
5. Characterize current	Data analysis,
process/product performance	tolerancing
6. Continual improvement	Exp't design/analysis

Goals of Public Health Surveillance

1. Disease surveillance: changes in "known" patterns of incidence or mortality
2. Syndromic surveillance: more or less gradual trends
3. Epidemic surveillance: unusual disease or outbreak where none or very few cases are expected

Different SPC monitoring tools for each situation:

1. Shewhart control charts
2. Cumulative sum charts (CUSUM);
   Exponential weighted moving average (EWMA)
3. Defects charts

Based on straightforward statistical principles

3. Data types and time periods

- Rates: Need reliable denominators, and grouped time periods or areas for sufficiently reliable rates
- Counts:
  Large counts: Poisson $\rightarrow$ Gaussian
  Small counts: Poisson (not Gaussian)
- – Monthly (e.g., not highly contagious)
  – Weekly (e.g., influenza)
  – Daily (e.g, West Nile Virus in summer)
  – hourly (biological warfare)
- Length of time interval depends on data to be monitored

4. Control charts

"Extremely powerful (and deceptively simple) tool" to assess process stability, detect process aberrations
(Vardeman and Jobe 1999)

Walter Shewhart, Bell Labs 1920s–1930s:

observed variation = baseline variation + removable variation

"baseline": from measurement technology, random factors (temperature), short-term effects: "stable"

"removable": "special cause", "assignable", "non-random", elimination returns process to "stable"

Generic control chart

Statistic Q_t plotted as a point at time t

Ex: $Q_t = \bar{x}_t$ = mean of measurements on n samples (choice of n)

Plot "Center Line"

- "Standards given": nominal targets (e.g., 5 cases)
- "Retrospective": use historical data

Plot upper/lower control limits UCL_Q, $LCL_Q \ni$

$$\mathrm{P}\{\ Q_t \notin (LCL_Q, UCL_Q)\ \} \text{ is "small"}$$

Avoid multiple false alarms: UCL_Q, $LCL_Q = \bar{\bar{x}} \pm 3\hat{\sigma}/\sqrt{n}$

Types of Shewhart charts

Data type	Q_t	Center	(LCL, UCL)	Distribution
Measurement	$\bar{x}_t$	$\bar{\bar{x}}$	$\pm 3\hat{\sigma}/\sqrt{n}$	Normal
Proportions	p_t	$\bar{p}$	$\pm\sqrt{\bar{p}(1-\bar{p})/n}$	Binomial
Counts	u_t	$\bar{u}$	$\pm 3\sqrt{\bar{u}}$	Poisson
Spread	s_t	$\bar{s}$	$B_3\bar{s}$, $B_4\bar{s}$	Chi-Square

$p_t = x_t/n_t$; $\bar{p} = \sum x_t / \sum n_t$

B_3, B_4: product of 0.5%, 99.5% quantiles from χ^2_{n-1} and factor to make $\bar{s}$ unbiased for σ

Shewhart chart patterns

In-control process:

- No obvious pattern or trend
- Rarely fall outside control limits
- Cluster about center line, above/below equally often
- Approach control limits only occasionally

Out-of-control process:

- Systematic variations/cycles:
  seasonal patterns, changes in shift/operator, etc.
- Instability: many points outside control limits
- Changes in level:
  Abrupt: New equipment, reporting definitions
  Gradual: Change in practice, tool wearout
- Mixtures: aberrations, grouping, clumping
  (may require stratification)

Shewhart alarm rules

Western Electric signals:

- Single point outside 3σ limits
- 2 out of 3 consecutive points outside 2σ limits
- 4 out of 5 consecutive points outside 1σ limits
- 8 consecutive points above/below center line

Other sets of rules in Nelson (*JQT* 1984) and Duncan *Quality Control & Engineering Statistics*

Demonstrable usefulness

Contaminated process (Thompson & Koronacki 2001):

$X_0 \sim N(10, 0.01)$ underlying process

$X_1 \sim N(0.4, 0.02)$ contamination (Prob $p_1 = 0.1$)

$X_2 \sim N(-.2, 0.08)$ contamination (Prob $p_2 = 0.05$)

Observe $Y_t =$

$X_0 \sim N(10, 0.01)$ with probability 0.855

$X_0 + X_1 \sim N(10.4, 0.03)$ with probability 0.095

$X_0 + X_2 \sim N(9.8, 0.09)$ with probability 0.045

$X_0 + X_1 + X_2 \sim N(10.2, 0.11)$ with probability 0.005

$\mathrm{E}(Y_t) = 10.03$, $\mathrm{Var}(Y_t) = 0.0323$

$\mathrm{E}(s^2) = 0.01 + 0.1(0.02) + 0.05(0.08) = 0.016$

$\mathrm{E}(LCL, UCL) \approx 10.03 \pm 3\sqrt{0.016/5} = (9.86,\ 10.2)$

```
     [-- in-control --]
    9.86              10.2
--X---------|---------X---------X
 9.8        10.0                 10.4
```

control chart has $\geq$50% chance of detecting contamination, even when based on erroneous means and variances

How quickly can the chart detect an aberration?

Average Run Length

ARL = average time T to first out-of-control signal

Not easy to calculate due to combinations of rules

Consider only $Q_t = \bar{x}_t$, Alarm = One point outside (LCL, UCL):

$$f_T(t) = q(1-q)^{t-1} \text{ (Geometric)} \Rightarrow E(T) = 1/q$$

$$q = \Pr\{Q_t \text{ lies outside control limits}\}$$

In-control ARL (shift = 0):

$q = \Pr\{|Z| > 3\} = 0.0027 \Rightarrow ARL_0 = 370$

Out-of-control ARL ($n = 5$, shift = σ):

$q = \Pr\{|\bar{x} - \mu|/(\sigma/\sqrt{5}) > 3|\mu + \sigma\} = 0.2236 \Rightarrow ARL_1 = 4.5$

Poisson u-chart ARL

Assume $\lambda = 1.5$ defects (cases) per unit (area):

$(LCL, UCL) = (0,\ 1.5 + 3\sqrt{1.5}) = (0,\ 5.2)$ ($\approx$ exact)

$q = \Pr\{X > 5.2\} = 0.005 \Rightarrow ARL_0 = 200$

Change in $\lambda = 4.5$:

$q = \Pr\{X > 5.2\} = 0.297 \Rightarrow ARL_{4.5} = 3.4$

Two samples at once: $\lambda = 3.0$: $(LCL, UCL) = (0,\ 8.2)$

$q = \Pr\{X > 8.2\} = 0.0038 \Rightarrow ARL_0 = 263$

Change in $\lambda = 9.0$:

$q = \Pr\{X > 8.2\} = 0.544 \Rightarrow ARL_{9.0} = 1.8$

Fewer false alarms, faster signal of change, twice the inspection

Exponentially Weighted Moving Average

Plot W_t, where $W_t = \lambda Q_t + (1-\lambda)W_{t-1}$

"Smoothed" control chart

One-step-ahead prediction forecast for IMA(1,1) = integrated moving average model (ϵ_t = white noise):

$$x_t - x_{t-1} = \epsilon_t - \lambda\epsilon_{t-1}$$

"Optimal" choice of λ: Minimize ARL_δ

Control limits depend on λ:

$$(LCL, UCL) = \mu_Q \pm H \cdot \sigma_Q \sqrt{\lambda/(2-\lambda)}$$

Signals faster than Shewhart for small changes, but more slowly for very large process changes (outliers are smoothed out)

EWMA Tables

"Optimal" values of λ

All-OK	Shift (δ/σ_Q)				
ARL	0.5	1.0	2.0	3.0	4.0
50	.08	.22	.56	.83	.96
100	.07	.18	.49	.78	.94
370	.05	.14	.37	.70	.90

EWMA Tables

Allowances H for EWMA(λ)

All-OK	"Optimal" λ^*					
ARL	0.05	0.1	0.2	0.3	0.4	1.0
50	1.52	1.81	2.06	2.17	2.23	2.33
100	1.88	2.15	2.36	2.45	2.50	2.58
370	2.49	2.70	2.86	2.93	2.96	3.00

Cumulative Sum Chart (CUSUM)

Sequential Likelihood Ratio Test Statistic:

Given target level T and statistics Q_t, plot W_t:

$$\begin{aligned} W_t &= (Q_t - T) + W_{t-1} \\ &= \textstyle\sum_{j=1}^{t} Q_j - tT \\ &= t(\bar{Q}_t - T),\ \bar{Q}_t = \text{mean}\{Q_1, \ldots, Q_t\} \end{aligned}$$

Changes in slope of plot reflect changes in T

"Acceptance" interval (L_t, U_t), where

$$U_t = \max\{0, (Q_t - K_1) + U_{t-1}\}$$
$$L_t = \min\{0, (Q_t - K_2) + L_{t-1}\}$$

"Out of control" if $U_t > h$ or $L_t < -h$

CUSUM, cont'd

Decision constants K_1, K_2, h depend on desired ARL_0

For small ARL if μ_Q changes from μ to $\mu + \delta$, choose optimal $K_1 = \mu_Q + \delta/2$, $K_2 = \mu_Q - \delta/2$.

Disease surveillance: interested in high-side only CUSUM and $h = \delta/2$ (δ = desired detectable difference)

ARL Comparisons: $ARL_0 = 370$, EWMA(λ^*)

	$\lvert\mu_Q - target\rvert/\sigma_Q$						
Chart	$\frac{1}{2}$	1	1h	2	2h	3	3h
Shewhart	155	43	15	7	3.2	2.0	1.4
CUSUM	35	10	5h	4	3.0	2.5	2.2
EWMA*	37	11	5h	3	2.4	1.8	1.4
SC-ARMA	164	49	17	7	3.5	2.1	na
Optimal λ^*	.05	.14	.25	.37	.54	.70	.82

SC-ARMA: Shewhart chart on residuals of fitted ARMA(1,1)
($\phi_1 = 0.475$, $\theta_1 = 0.45$, $\rho_1 = 0.025$)

Vardeman and Jobe 1999; Crowder 1987; Wardell et al. 1994

Correlated data

Alwan and Roberts (1988): Fit ARMA(p,q)

1. "Common-cause control chart":
   Plot *forecasted values*
   Accounts for systematic variation
2. "Special Cause" charts:
   Principle: Systematic variation can be removed via dynamic process control ("feedback loops") or design of experiments focused on minimizing variation
   Fit ARMA; Plot *residuals* from fitted ARMA
   Shewhart control chart on residuals

Effect of fitting wrong model?

Cyclical behavior

Control charts to detect cycles

Beneke et al. (1988): Periodogram $I_j \equiv I(2\pi j/T)$

$$Q_t = \max(I_j) \; / \; \sum I_j$$

(relative contribution to periodogram at j^{th} Fourier frequency)

Spurrier and Thombs (1990): Harmonic analysis ($t = 1,...,T$)

$$x_t = \mu + \sum_{j=1}^{m} \alpha_j cos(2\pi jt/T) + \beta_j sin(2\pi jt/T) + \epsilon_t$$

(cf. Bloomfield 1976)

Plot maximum reduction in sum of squares by fitting j^{th} Fourier frequency, compared to $\sum(x_t - \bar{x})^2$

Computationally involved; not easily interpretable;
designed to detect very specific behavior (cycles)

Multivariate Charts

x_t = vector of p measurements (counts, rates)
per unit (county, tract) at time t

Hotelling's T^2:
Plot $Q_t = T^2 = n(\bar{x} - \mu)'V^{-1}(\bar{x} - \mu) = \sum_i \sum_j d_i r^{ij} d_j$
μ = target mean, $d_i = (\bar{x}_i - \mu_i)/s_j$
V = covariance matrix among measurements
$r^{ij} = ij^{th}$ element of inverse correlation matrix
T^2 = nominally distributed as χ^2_p
Centerline $= p$, $UCL = p + 3\sqrt{2p}$

Independence $\Rightarrow T^2 = \sum(tstat_j)^2 \Rightarrow$ signal cumulative small changes but perhaps not one large change

Robustness?

Multivariate CUSUM (Healy 1987):

Plot $Q_t = \max(0,\, Q_{t-1} + a'(x_t - \mu_Q) - D/2)$

$a' = \delta'\Sigma^{-1}/D$, $D = (\delta'\Sigma^{-1}\delta)^{1/2}$

δ = change from in-control mean

To detect shift in mean of multivariate normal random vector, Multivariate CUSUM reduces to univariate normal CUSUM

Power depends on μ_Q, μ'_Q, Σ only through D (not on p).

Same one-sided ARLs as for CUSUM.

MEWMA: $Q_t = T^2$; Plot W_t, where $W_t = \lambda Q_t + (1-\lambda)W_{t-1}$

Judgment composites?

5. Sample data from CDC

Goal: Disease surveillance (more or less known patterns of incidence) CDC's NNDSS
(National Notifiable Disease Surveillance System)
Reports in MMWR (Mortality and Morbidity Weekly Report)

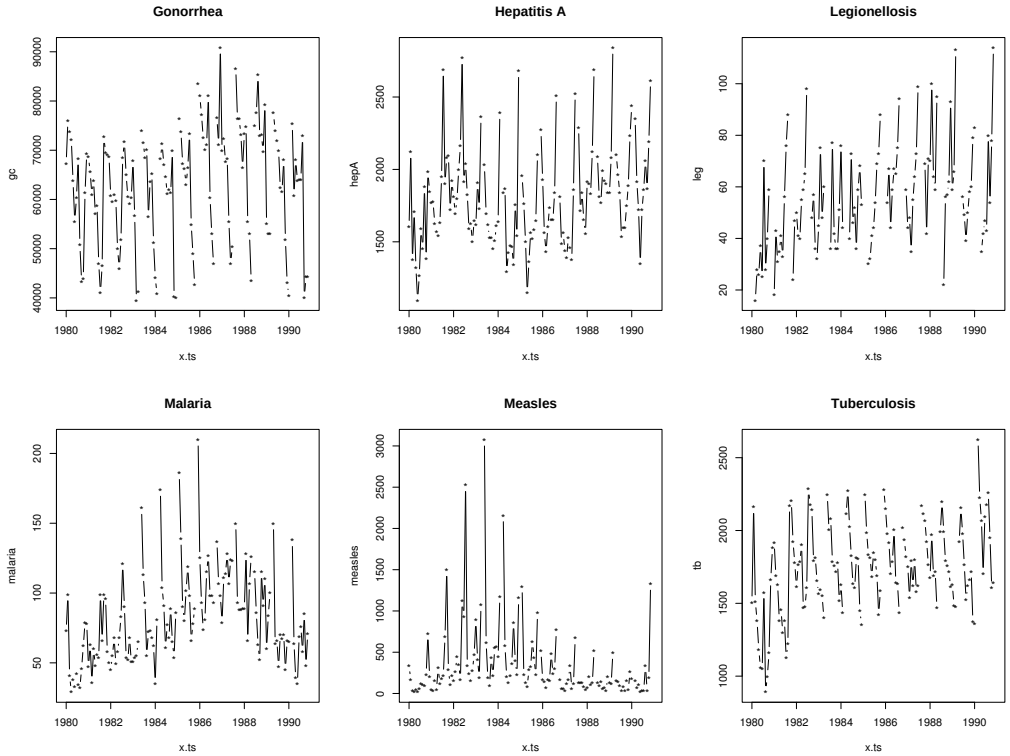
Number of cases of notifiable disease in 4-week period
13 periods per year, 1980–1990 (11 years)
Next count comes in

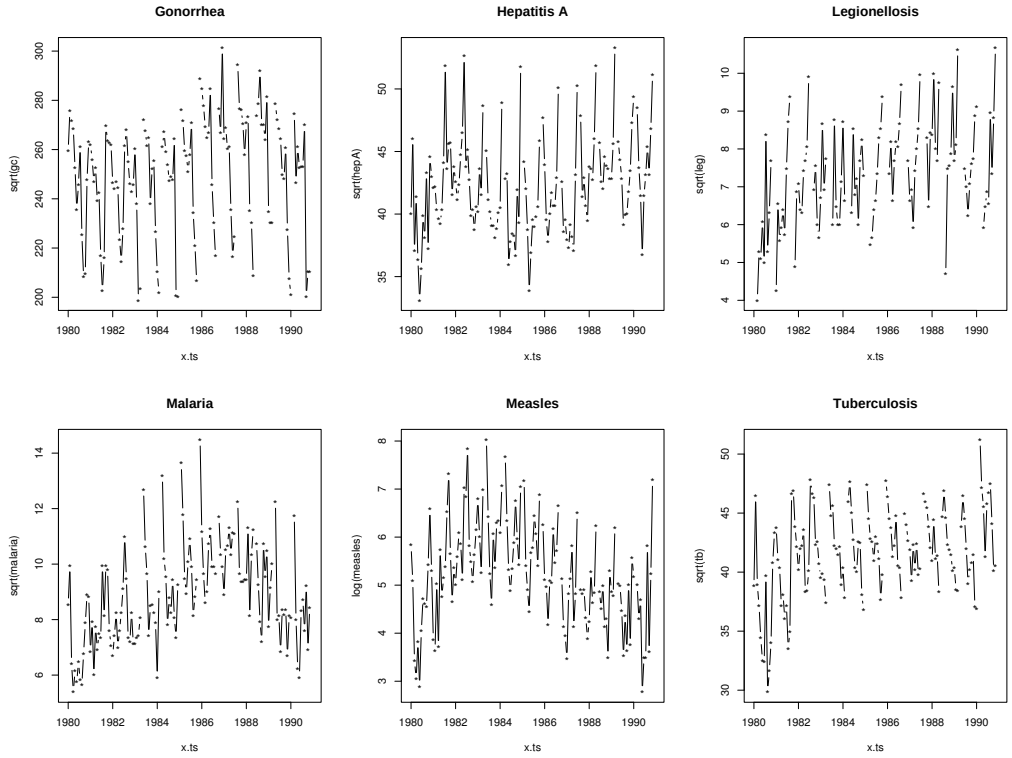
Question: Is there cause for alarm?

Sample data: Legionellosis

	1	2	3	4	5	6	7	8	9	10	11	12	13
82	16	18	24	48	36	40	30	54	59	69	22	56	35
83	28	43	47	57	77	73	32	67	44	42	56	49	42
84	26	31	50	36	42	46	41	44	48	71	57	39	47
85	37	35	41	32	36	52	44	59	35	70	62	50	43
86	25	41	40	45	36	36	54	67	55	100	93	58	80
87	70	33	55	75	51	59	69	65	64	64	59	60	54
88	28	56	59	48	76	68	73	75	73	59	66	79	78
89	40	76	65	60	44	53	88	94	99	95	113	83	114
90	59	88	98										

Is "98" unusual?


Gonorrhea
gc
40000
50000
60000
70000
80000
90000
1980
1982
1984
1986
1988
1990
x.ts
Hepatitis A
hepA
1500
2000
2500
1980
1982
1984
1986
1988
1990
x.ts
Legionellosis
leg
20
40
60
80
100
1980
1982
1984
1986
1988
1990
x.ts
Malaria
malaria
50
100
150
200
1980
1982
1984
1986
1988
1990
x.ts
Measles
measles
0
500
1000
1500
2000
2500
3000
1980
1982
1984
1986
1988
1990
x.ts
Tuberculosis
tb
1000
1500
2000
2500
1980
1982
1984
1986
1988
1990
x.ts


Gonorrhea
sqrt(gc)
x.ts
Hepatitis A
sqrt(hepA)
x.ts
Legionellosis
sqrt(leg)
x.ts
Malaria
sqrt(malaria)
x.ts
Measles
log(measles)
x.ts
Tuberculosis
sqrt(tb)
x.ts

Problems for Disease Surveillance

SPC techniques assume:

1. independent measurements
2. equally-likely units are randomly selected
3. sources of variation are largely enumerable
4. changes can be isolated and addressed

Disease surveillance data involve:

1. highly correlated data across time periods
2. missing people that have zero chance of selection
3. unknown sources of variation
4. changing effects of variation (e.g., HIV)
5. Changes in measuring instruments (survey, reporting)
6. Changes in data availability (HIPAA)
7. Changes in definition of disease (e.g., CD4 count)
8. Effect of treatment on incidence data (e.g., AZT)

How well do these tools work on such data?

6. Modeling Seasonality

Need to take account of seasonal trend

One simple way:

- Estimate typical "Period 1" effect (average of previous "Period 1" observations)
- Subtract typical "Period 1" effect from all observations that occurred in Period 1

But:

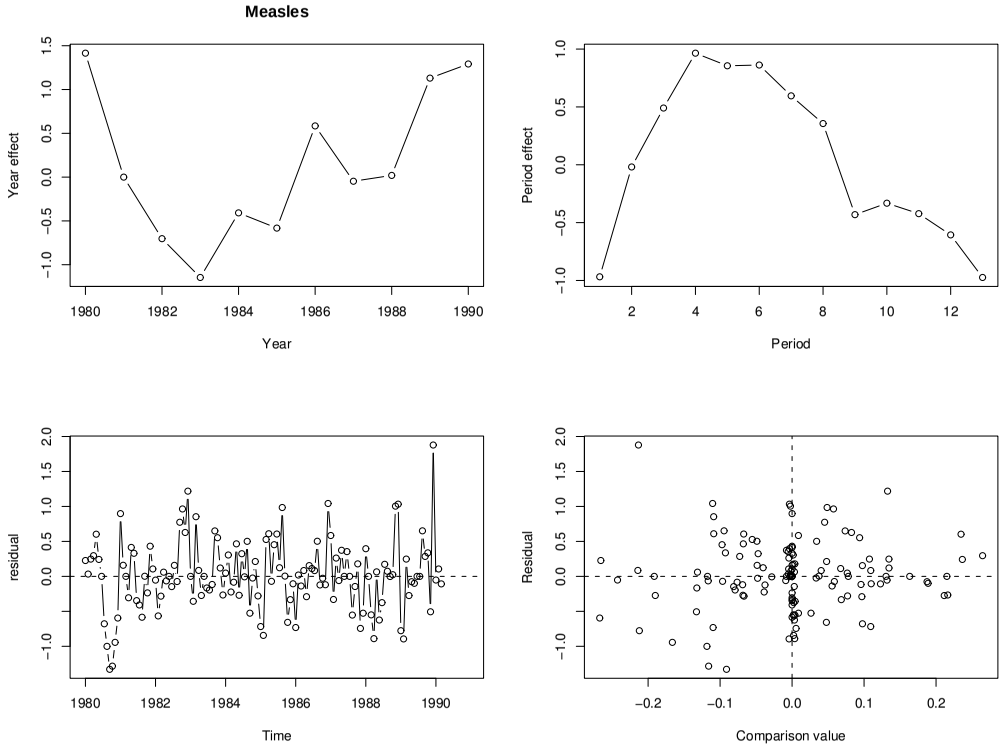
- Years are changing also
- Time points are not independent; highly correlated (even after subtracting year/period effects)
- Usually we do not know "typical year" effect or "typical period" effect
- Abrupt changes

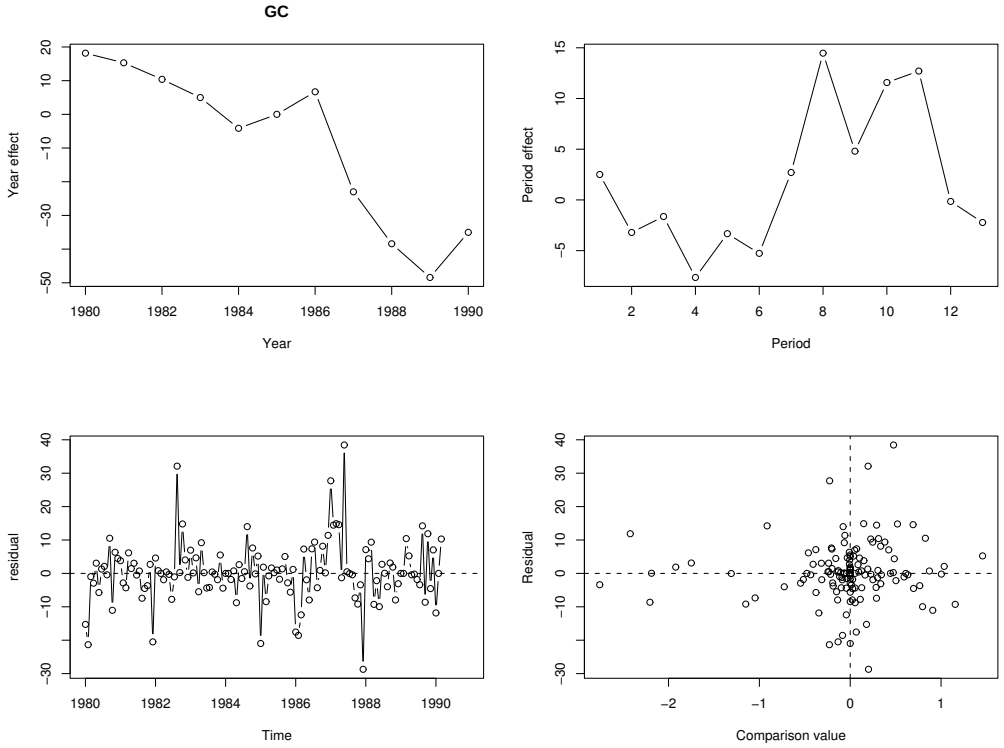
Therefore, usual "control chart" procedures may not apply.

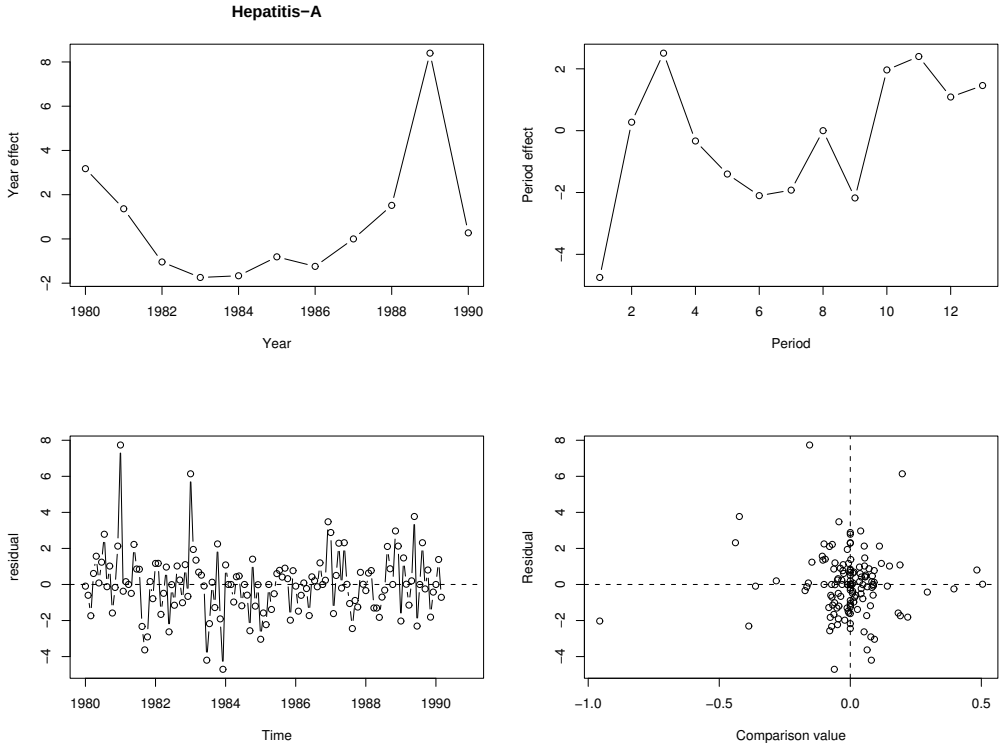
6. Some modeling

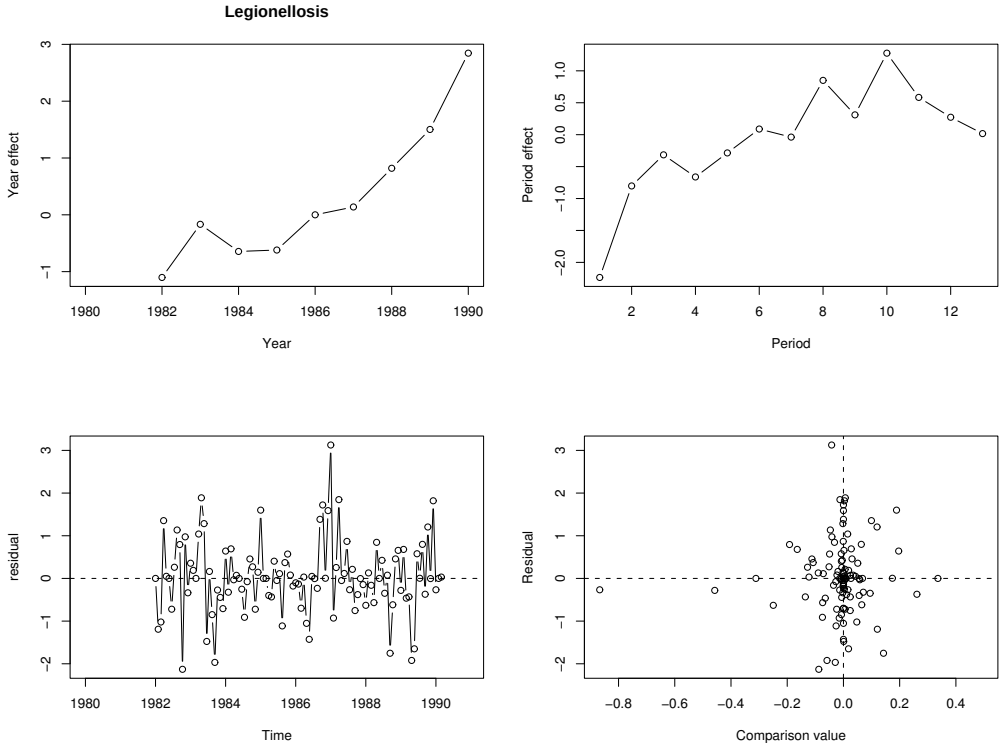
Methods depend on type of data:

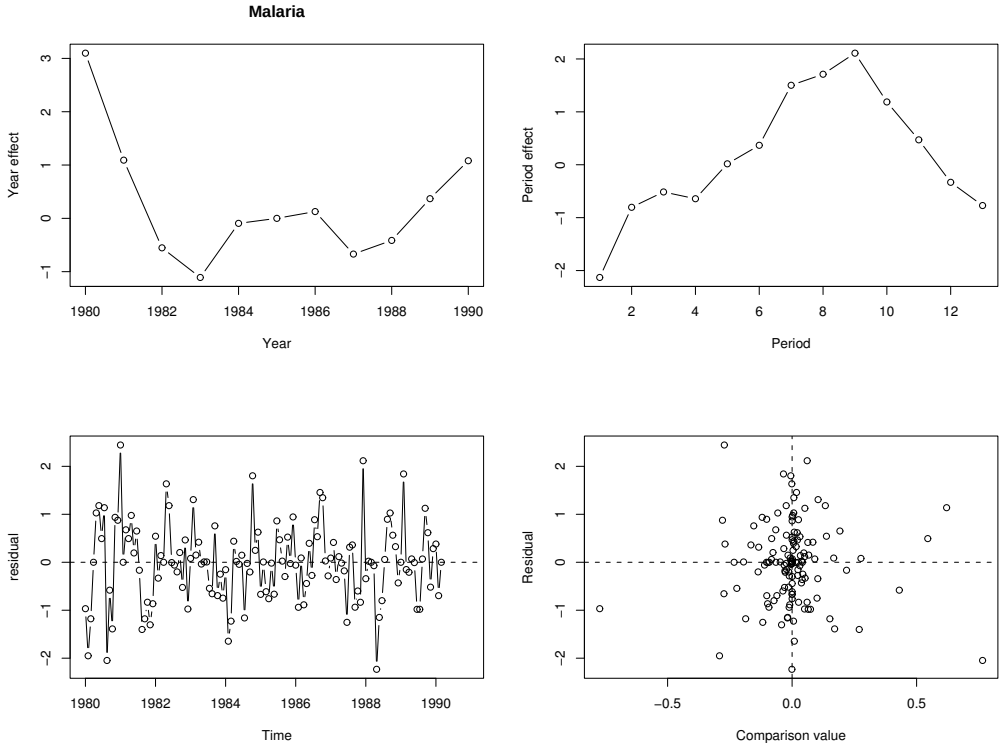
- Time period: Monthly, weekly, daily, hourly
- Seasonal effects?
- smallish counts (0, 1, 2, ...)
- medium-sized counts (dozens)
- large counts: hundreds, thousands
- Typically transform via square roots or logs to remove dependence of uncertainty on magnitude of the count

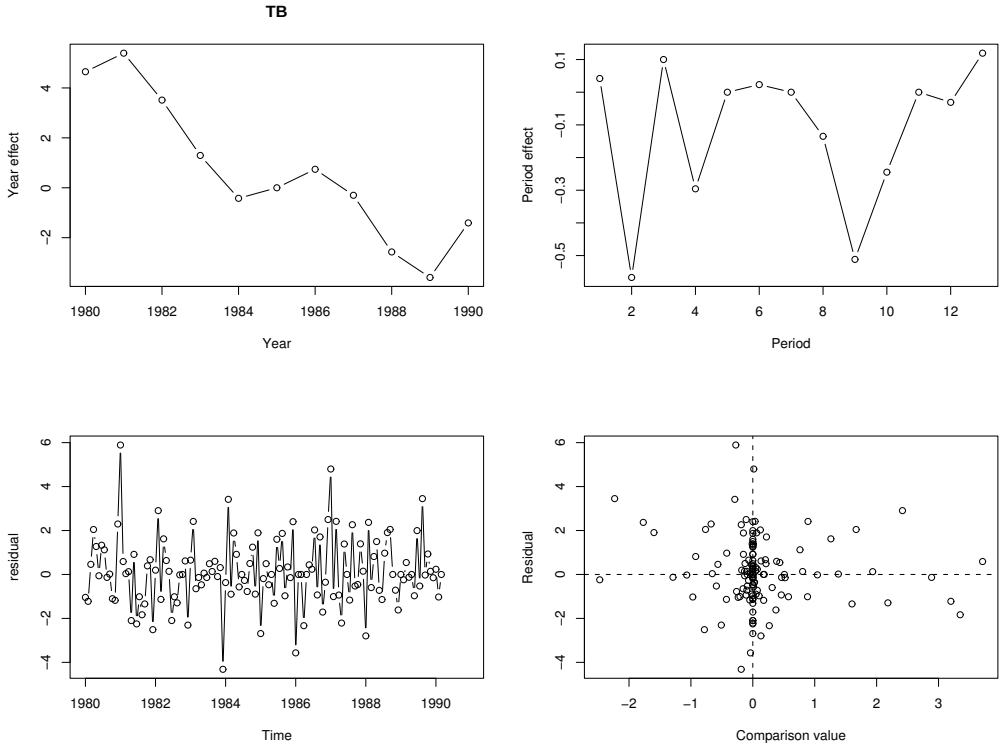

Measles
Year effect
Year
Period effect
Period
residual
Time
Residual
Comparison value


GC
Year effect
Year
Period effect
Period
residual
Time
Residual
Comparison value


Hepatitis−A
Year effect
Year
Period effect
Period
residual
Time
Residual
Comparison value


Legionellosis
Year effect
Year
Period effect
Period
residual
Time
Residual
Comparison value


Malaria
Year effect
Year
Period effect
Period
residual
Time
Residual
Comparison value


TB
Year effect
Year
Period effect
Period
residual
Time
Residual
Comparison value

Kafadar and Stroup, *Statistics in Medicine*, 1992

Stroup et al., *American Journal of Epidemiology*, 1993

Current count: 98

Counts from same period, ±1 period, for previous 5 years:

	Original			Square Root		
current		98			9.90	
2003	76	65	60	8.72	8.06	7.75
2002	56	59	48	7.48	7.68	6.93
2001	55	75	51	5.74	7.42	8.66
2000	40	45	36	6.40	6.32	6.71
1999	41	32	36	5.92	6.40	5.66

Mean (median) of 15 historical counts = 50.9 (48)

Standard deviation of 15 counts = 13.8

Ratio of current to historical mean (median) = 1.93, 2.04

Square root: Mean (median) of 15 past counts = 7.06 (6.93)

Standard deviation of 15 counts = 1.00

Ratio of current to historical mean (median) = 1.40, 1.43

Estimated standard deviation = 0.165

"2-SD-interval": $1.40 \pm 2(0.165) = (1.07, 1.73)$

Interval does not cover 1.00 $\Rightarrow$ "98" may be considered "high"

7. Other Methods

Tests for data exhibiting no seasonality:

- "Clusters" in time: Many events in few adjacent time periods
- "Rare events": Many periods with no cases
- Most methods are based on binomial or Poisson counts
- Tukey's "statistical strength" (1992 unpublished, of course):

$$Q_t = \sqrt{4(obs_t) + 2} - \sqrt{4(exp_t) + 1} \sim N(0, 1)$$

rounded to nearest integer (Freeman and Tukey 1950)

7. Final Thoughts

Routine application of SPC tools (e.g., control charts) will require thoughtful modeling, to account for autocorrelation, seasonality, accountable changes (e.g., changes in reporting)

Some simple tools might be useful in the meantime

Churchill Eisenhart: "practical power of a procedure" = product of mathematical power and probability that the procedure will be used: "A compact procedure may well be used so much more often as to more than compensate for its loss of mathematical power" (Tukey 1959)

Much work remains to be done

Some References

1. Vardeman, Stephen B.; Jobe, J. Marcus (1999), *Statistical Quality Assurance Methods for Engineers*, Wiley, NY.

2. Thompson,James R.; Koronacki,Jacek (2001), *Statistical Process Control: The Deming Paradigm and Beyond, Second Edition*, Chapman & Hall/CRC, NY.

3. *Statistical Quality Control Handbook* (1984), Western Electric.

4. Crowder, Stephen V. (1987), A simple method for studying run-length distributions of exponentially-weighted moving average charts, *Technometrics* 29(4), 401–408.

5. Healy, John D. (1987), A note on multivariate CUSUM procedures, *Technometrics* 29(4), 409–412.

6. Alwan, L.C.; Roberts, H.V. (1988), Time series modeling for statistical process control, *Journal of Business and Economic Statistics* 6:87–95.

7. Wardell, Don G.; Moskowitz, H.; Plante, Robert D. (1994), Run length distributions of special-cause control charts for correlated processes (with discussion), *Technometrics* 36(1), 3–27.

8. Beneke, M.; Leemis, L.M.; Schlegel, R.E.; Foote, B.L. (1988), Spectral analysis in quality control: A control chart based on the periodogram, *Technometrics* 30(2), 63–70.

9. Spurrier, John D.; Thombs, L.A. (1990), Control charts for detecting cyclical behavior, *Technometrics* 32(2), 163–171.

10. 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.

11. Many other articles in *Technometrics* (e.g., February 2000 issue)