Technical Briefing 2

Published December 2007 - Reprinted August 2009

ASSOCIATION OF PUBLIC HEALTH OBSERVATORIES

Statistical process control methods in public health intelligence

"All things have their own measure; ultimately there are fixed limits outside which something must be wrong"

Horace, Satires. 35BC

Purpose

This is the second in a series of technical briefings produced by the Association of Public Health Observatories, designed to support public health practitioners and analysts and to promote the use of public health intelligence in decision making.

In this briefing we look at the uses of statistical process control (SPC) methods and associated visual display tools (control charts and funnel plots) in public health intelligence. The briefing provides a summary overview. Updates and more material, including methods and tools to support our Technical Briefing series will be made available through our website at http://www.apho.org.uk

Contents

Author

Julian Flowers

Acknowledgements

Mohammed Mohammed
David Spiegelhalter
Jake Abbas
Diane Stockton
Paul Fryers
Anne Cunningham

Background

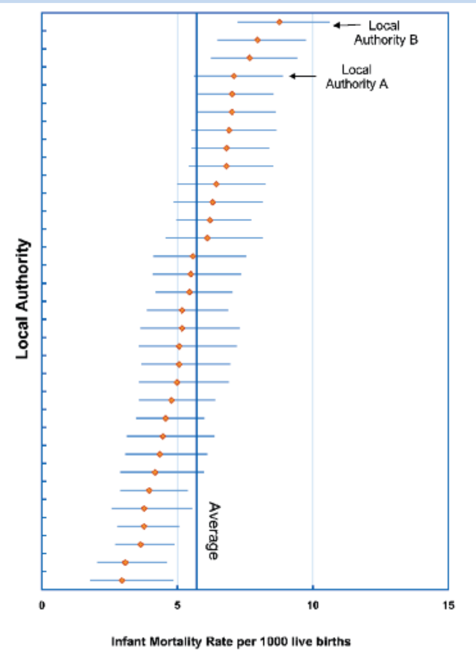
Public health practice is highly dependent on the effective use of information, and commonly makes comparisons between areas, groups or institutions. Methods based on ranking, such as league tables or percentiles, have a number of flaws. The main problem with ranking is the implicit assumption that apparent differences between organisations are the results of better or poorer performance. Simply because institutions may produce different values for an indicator, and we naturally tend to rank these values, does not mean that we are observing variation in performance. All systems within which institutions operate, no matter how stable, will produce variable outcomes.

The questions we need to answer are: 'Is the observed variation more or less than we would normally expect?'; 'Are there genuine outliers?'; 'Are there exceptionally good performers?'; 'What reasons might there be for excess variation', and so on. Alternative methods based on understanding variation may be more appropriate, and SPC techniques can be very helpful. SPC methods are particularly useful when, as is often the case in public health monitoring, there are small numbers of events. They have also been shown to improve targeting of performance management compared with the use of league tables.[1]

Figure 1 gives an example of displaying such information comparing a conventional 'caterpillar' plot with a funnel plot of the same data. The funnel plot approach makes it easier to identify which data points indicate areas that may be worthy of further investigation.

SPC methods based on control charts have been in use for more than 80 years, particularly in industrial quality control, but increasingly in recent years in health and health care. Awareness of the power of these methods has been highlighted by the Shipman and Bristol Royal Infirmary enquiries.[2;3] They are designed to study variability of performance over time or between institutions or areas and are powerful tools for population health surveillance and monitoring.[4] In public health practice, control charts are very useful in showing the variation that exists in health outcomes or performance between groups, areas or institutions - often a starting-point for needs assessment, targeting services and epidemiological understanding.

Most charts can be drawn in standard spreadsheet or statistical packages, and freely available tools have been developed which automate the process for common measures used in public health practice (see, for example, http://www.erpho.org.uk/topics/tools/ and http://www.indicators.scot.nhs.uk/SPC/SPC.html).



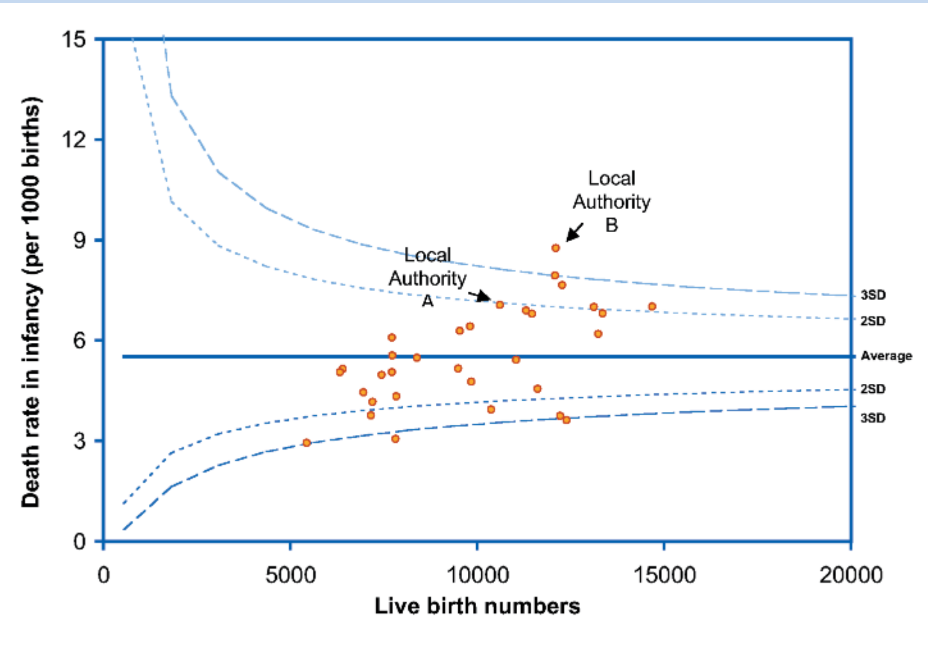


Figure 1. Traditional "caterpillar" plot with 95% confidence intervals vs funnel plot of the same data. In the funnel plot it can be seen that point A lies within the inner control limits so is not an outlier, whereas point B lies above the upper limit so is a case of special cause variation.

Principles of SPC techniques

There are a few key principles:

1. In any process or system, variation is to be expected. By use of simple statistical techniques we can define the limits of variation beyond which data points are deemed worthy of investigation.

2. These limits are known as control limits. Variation within these limits is often called common-cause or process variation; variation outside these limits is often called special-cause or extra-process variation.[2;5-7]

3. Common-cause variation is that which can be expected to occur in a stable process or system - one which is 'under control'. Special-cause variation may derive from systematic or unexpected deviation from the norm and may highlight an area or an observation which is worthy of further investigation.

4. A useful estimate of expected 'performance' of a system is often the group average, and the best estimate of expected variation around the group average is ± 3 standard deviations (SDs) (roughly equivalent to 99.8% confidence intervals). This degree of variation has both empirical and theoretical justification.

5. These limits (control limits) can be readily derived and depend on the nature of the data being used to assess the process.

Types of control chart

All control charts are plots of the underlying data with lines indicating the mean, median or target value and control limits superimposed. The common types are based on simple statistical distributions: the Poisson distribution for counts, rates and ratios; the binomial distribution for proportions; and the normal distribution for continuous data (see Figure 2):

1. Conventional control charts - usually the indicator of interest is plotted on the y-axis, against time or the unit of analysis on the x-axis. Control charts can be plotted with small numbers of data points although their power is increased with more data.

2. Funnel plots - a type of chart where the indicator of interest is plotted against the denominator or sample size - this gives it the characteristic funnel shape.[8]

The choice of chart will depend on the phenomenon being studied and the type of data being examined. Conventional control charts have been developed for count data, proportions and continuous variables. Funnel plots have been developed for proportions, directly standardised rates, indirectly standardised rates and ratios and rate ratios. Spiegelhalter describes how they can be constructed for a range of statistics.[9]

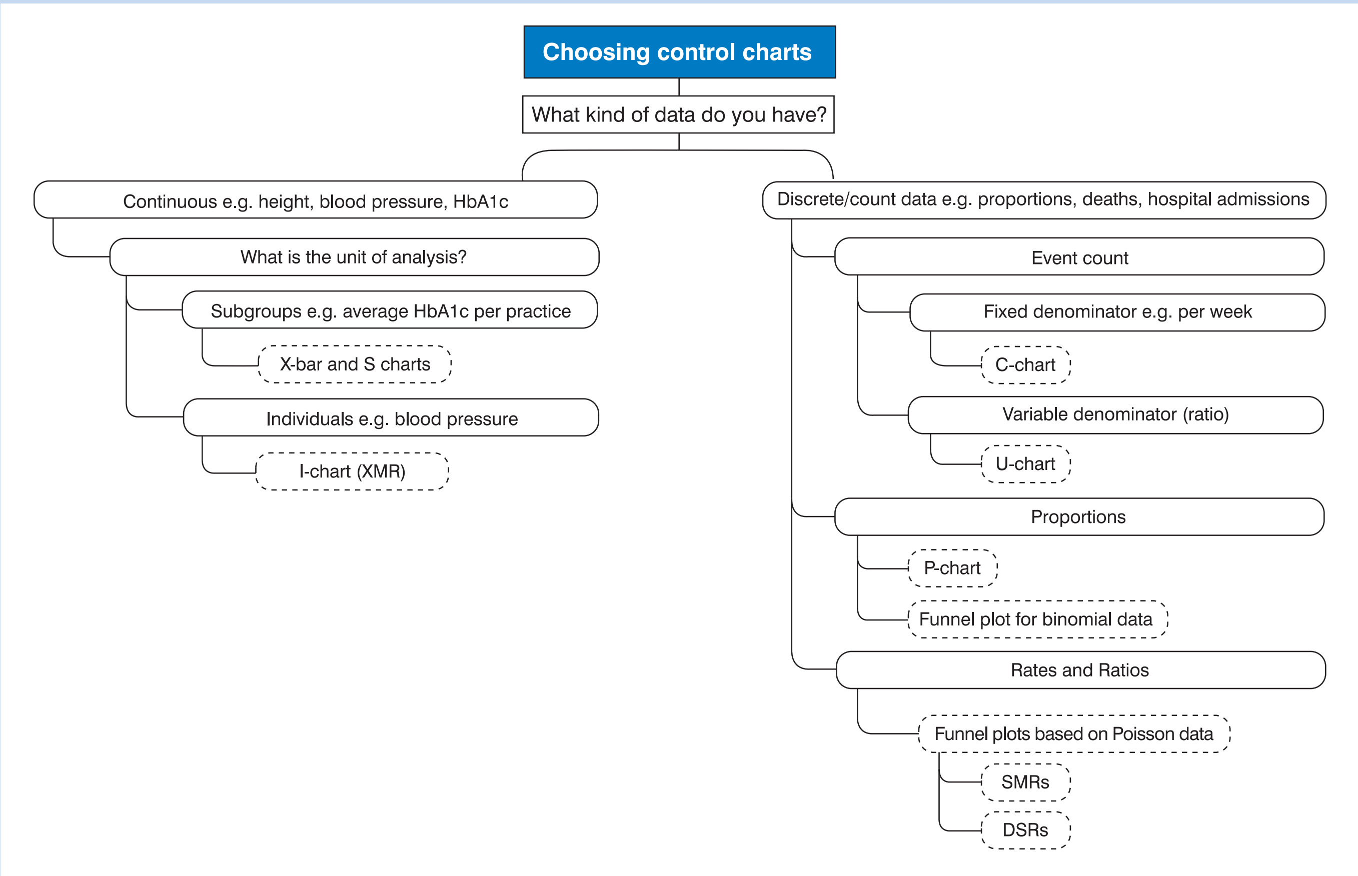


Figure 2. Typology of control charts which can be used in public health practice

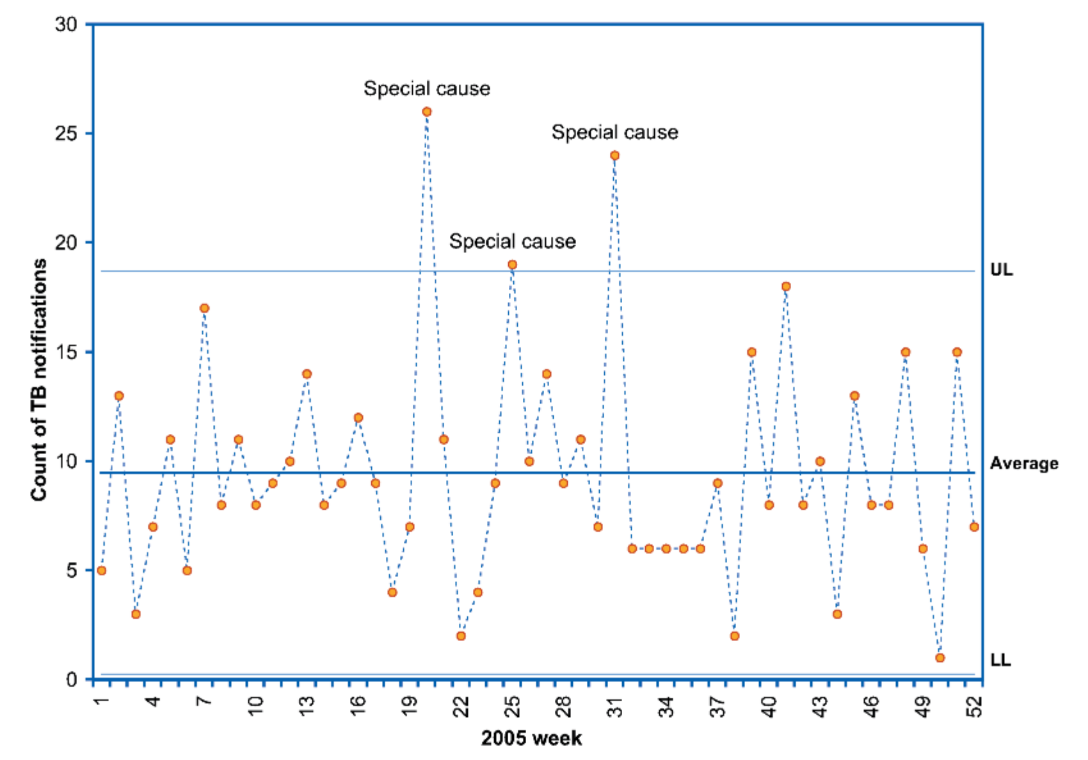


Figure 3. C-chart of TB notifications by week. Source: Disease notifications data; HPA

Examples of using control charts and SPC methods in public health practice

Control charts can help us to present and interpret our information more intelligently. They can be used to detect unusual or outlying patterns, e.g. poor performance, outbreaks or unusual patterns of disease; in health profiling and assessing levels of performance; to decide whether or not targets are being met; and in assessing health inequalities. In this section we provide some examples of how charts can be used in different situations.

Identifying outbreaks

Data from notifications or enhanced surveillance systems are often used to give early warning of unusual patterns of disease which might indicate an outbreak. Observations are tracked over time. In Figure 3 the weekly number of cases of tuberculosis in 2005 in one region of England is plotted with the long-term average as the centre line. This is known as a C-chart. The control limits are calculated as $\pm 3^*\sqrt{\text{mean}}$. The chart shows a well-controlled pattern consistent with an average number of notifications of nine per week, with the exception of three special-cause variations in weeks 20, 25 and 31 which may be worth investigating further.

Assessing performance against targets or desired levels

Control charts can be used to assess performance against targets. Figure 4 plots the proportion of weekly A&E attendances which were seen within the four-hour wait target, across a former strategic health authority (SHA). This form of control chart is known as a P-chart. During the period shown, there were changes in the national target from 90% to 95% and then to 98%, which are plotted on the chart.

The chart suggests that:

- During the initial period, performance was consistent with the 90% target but was not well controlled, with several special-cause variations.
- The average performance rate then increased but did not meet the new target of 95%.
- In the latter period, performance improved and exceeded the 98% target, consistent with a 'controlled system' with an average achieved percentage of 98.4%.

Methods like these can also be used to monitor commissioning activity against planned levels and to detect true or unexpected variances or shifts in activity.

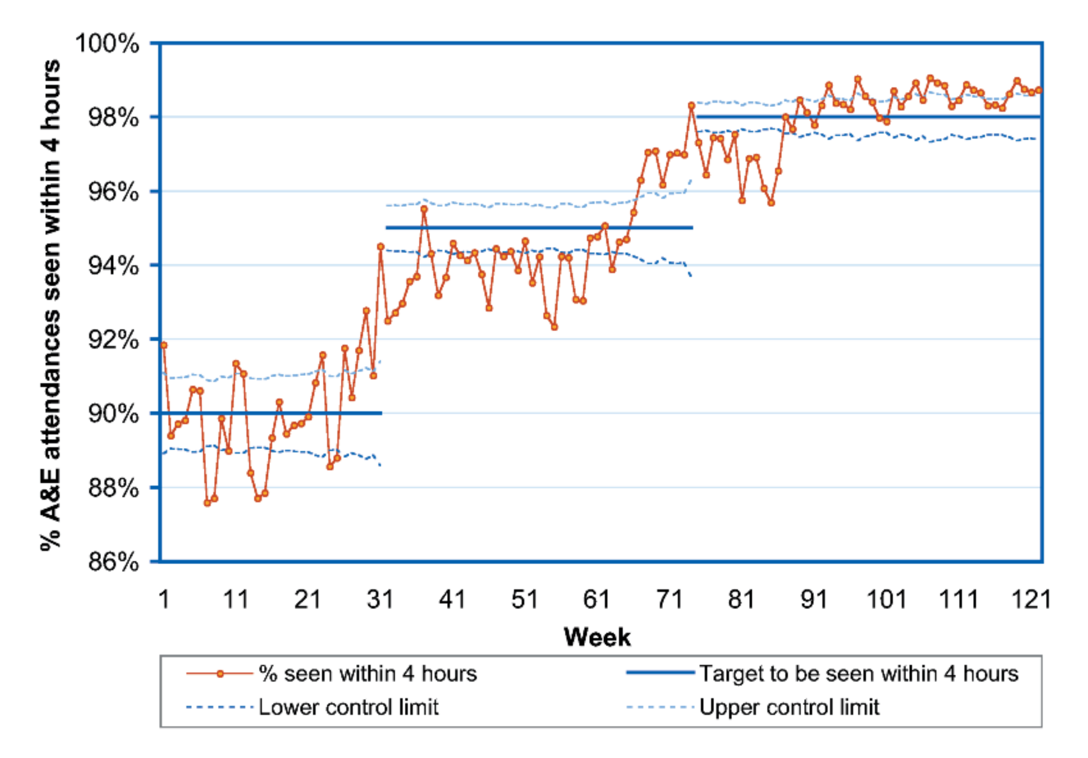


Figure 4. Control chart showing weekly percentage of A&E attendances seen within 4 hours compared with national targets

Identifying unusual patterns of disease

Public health practitioners are often called upon to investigate apparent unusual patterns or levels of disease occurrence - the initial task is to decide if such an occurrence is genuinely unusual. Funnel plots may be helpful here. Although they may be more familiar to some in detecting publication bias in meta-analysis, they are a powerful tool for the display of public health data. Figure 5 overleaf shows the variation in all-age breast cancer mortality between local authorities (LAs) in England in 2001-2002. It shows that virtually all the variation in death rates is within the limits of expected chance variation, with the exception of two areas both lower than the national average. (Note: there are 350 LAs in this example - we would expect by chance only two outliers per 1000 observations.)

Identifying health inequalities

The funnel plot for breast cancer mortality in Figure 5 shows that a high proportion of LA areas have death rates which are contained within the funnels; that is, the extent of variation can be reasonably explained by chance and the rates are consistent with the average value. Compare this with the plot for lung cancer (Figure 6 overleaf) which shows a high proportion of points outside the funnels. Figure 6 demonstrates greater special-cause variation - fewer LAs can be considered as having death rates equivalent to the national average, i.e. there is greater inequality between LAs.

Performance monitoring frameworks

SPC methods can be used to generate typical 'traffic lights' for performance reports - they have been used, for example, in comparing regional performance. However, for most health indicators there are many reasons for variations in outcomes; it should not be assumed that being outside the control limits necessarily implies poor performance.

Issues with control charts

In the lung cancer example, there are many LA areas which lie outside the control limits exhibiting special-cause variation. Such an abundance of points outside the control limits is sometimes known as overdispersion.[10-12] It arises when there are large numbers of events, and case-mix or other risk factors (e.g. deprivation) are not accounted for.

In this example, the overdispersion is probably due to the strong relationship between lung cancer and deprivation. The absence of overdispersion from the breast cancer plot reflects a much weaker relationship with deprivation.

There are two schools of thought as to how to handle overdispersion. In performance management practice, we are trying to identify differences that can be fairly attributed to differences in organisational performance. In this case it is usual to adjust the control limits or the data to eliminate potential sources of variation, such as case-mix and demography. This has the effect of creating a 'level playing field'.[11-13]

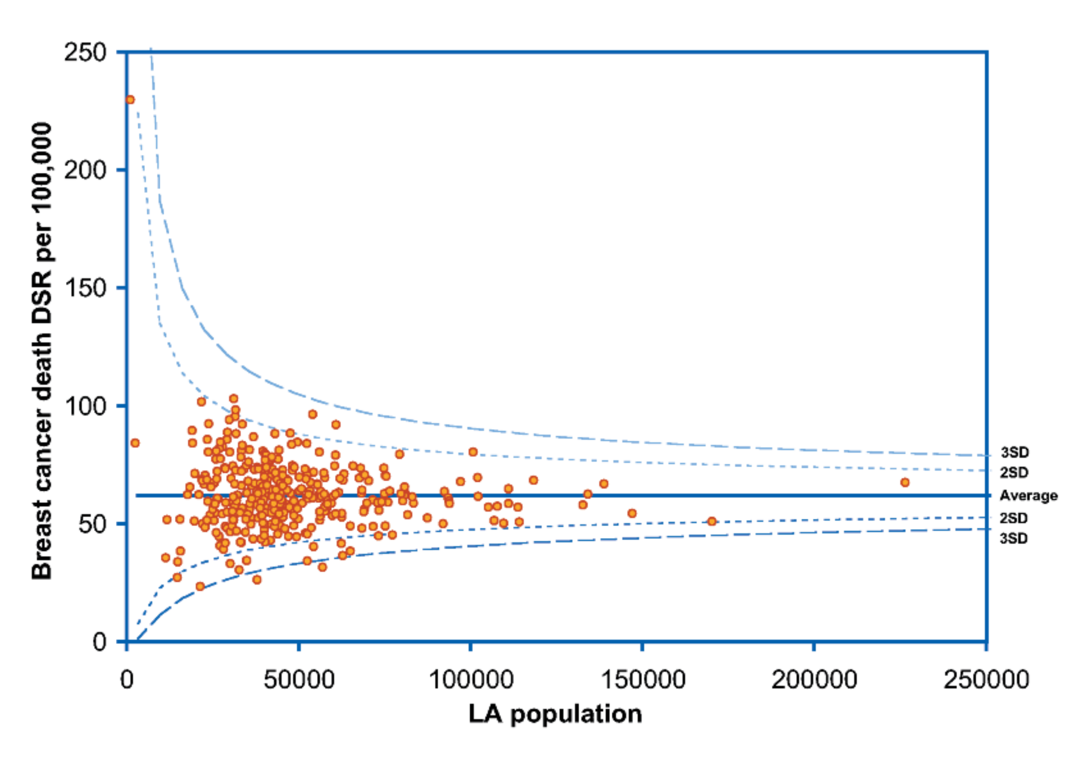


Figure 5. Funnel plot of variation in breast cancer mortality across England

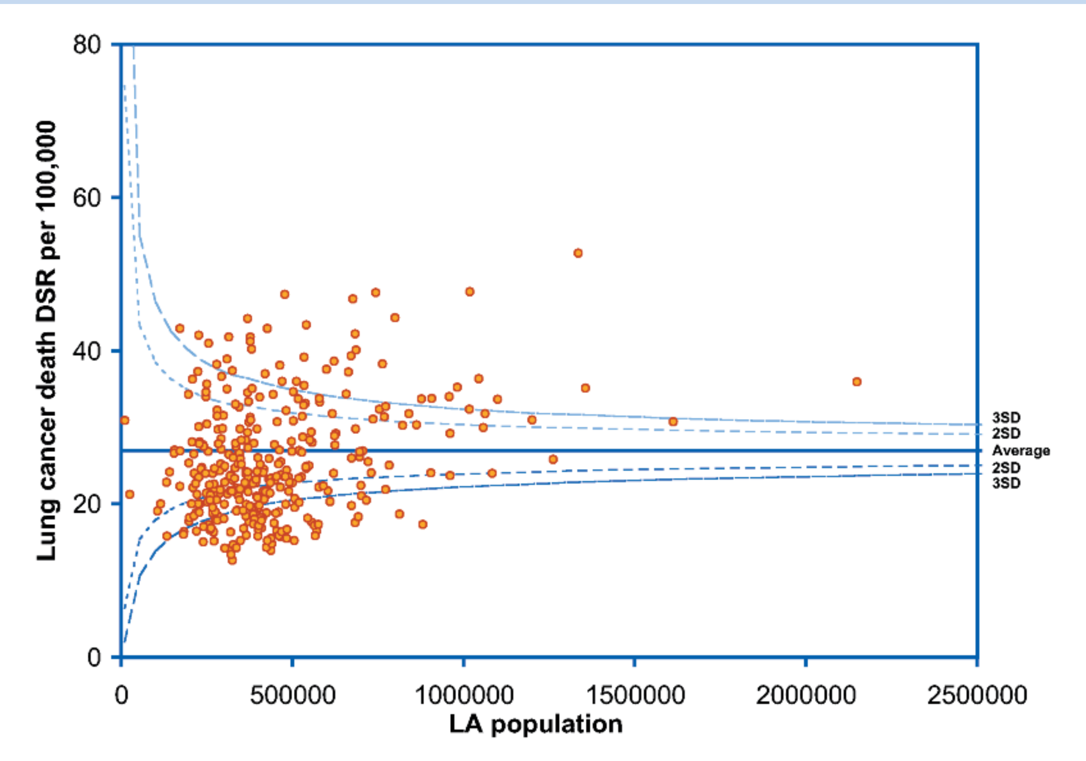


Figure 6. Funnel plot of variation in lung cancer mortality across England

In public health practice, on the other hand, we are likely to be interested in such sources of variation for their own sake. Rather than eliminate them, we want to draw attention to them and understand the reasons behind them, by means of further analysis if necessary. Hence we tend not to alter control limits, and prefer to display the variation in Figure 6 as it actually is.

There are several more powerful techniques being developed, particularly for assessing individual clinical performance, and change over time, which could be applied more widely as tools for their use become more accessible.[13;14]

Dealing with special-cause variation and 'out of control' systems

Identifying sources of variation

Variation is inherent in all 'systems'. It may arise either by chance or by assignable or special causes as we described - for example:

- Data issues such as differences in counting, coding or measurement or problems with information systems often give rise to special-cause variation.
- Demographic differences, deprivation levels and ethnic diversity frequently cause significant variations in population health indicators. Case-mix and socio-economic variation can cause differences in health outcomes. Expert statistical help may sometimes be needed. It may sometimes be difficult to pinpoint the exact reasons for observed variation.
- If special-cause variation cannot be accounted for by population differences, it may be necessary to undertake more detailed investigation of underlying processes.

Acting on findings

How to deal with variation will depend on findings:

1. If the system is under control (no special-cause variation) and operating at an appropriate level, no action is necessary.
2. If the system is operating at an appropriate level but there is special-cause variation, then investigating the cause may be appropriate.
3. If the system is under control but operating at an inappropriate level, it may be necessary to change the whole system.
4. In some situations, a system may be neither in-control nor operating at an appropriate level, in which case tackling special-cause variation before altering the process may be necessary (there may be no 'system' in place!).

Reference list

(1) Marshall T, Mohammed MA, Rouse A. A randomized controlled trial of league tables and control charts as aids to health service decision-making. Int J Qual Health Care 2004 Aug;16(4):309-15.

(2) Mohammed MA, Cheng KK, Rouse A, Marshall T. Bristol, Shipman, and clinical governance: Shewhart's forgotten lessons. Lancet 2001 Feb 10;357(9254): 463-7.

(3) Mohammed MA, Cheng KK, Rouse A, Marshall T, Duffy J. Was Bristol an outlier? Lancet 2001 Dec 15;358(9298):2083-4.

(4) Woodall WH. The use of control charts in health-care and public health surveillance. Journal of Quality Technology 2006;38:89-104.

(5) Marshall T, Mohammed MA. Understanding variation in quality improvement: the treatment of sore throats in primary care. Fam Pract 2003 Feb;20(1):69-73.

(6) Marshall T, Mohammed MA, Lim HT. Understanding variation for clinical governance: an illustration using the diagnosis and treatment of sore throat. Br J Gen Pract 2002 Apr;52(477):277-83.

(7) Mohammed MA, Cheng KK, Rouse A, Marshall T. Use of Shewhart's technique. Lancet 2001 Aug 11;358(9280):512.

(8) Spiegelhalter D. Funnel plots for institutional comparison. Qual Saf Health Care 2002 Dec;11(4): 390-1.

(9) Spiegelhalter DJ. Funnel plots for comparing institutional performance. Stat Med 2005 Apr 30;24(8):1185-202.

(10) Spiegelhalter DJ. Problems in assessing rates of infection with methicillin resistant Staphylococcus aureus. BMJ 2005 Oct 29;331(7523):1013-15.

(11) Spiegelhalter DJ. Handling over-dispersion of performance indicators. Qual Saf Health Care 2005 Oct;14(5):347-51.

(12) Mohammed MA, Laney D. Overdispersion in health care performance data: Laney's approach. Qual Saf Health Care 2006 Oct;15(5):383-4.

(13) Grigg OA, Farewell VT, Spiegelhalter DJ. Use of risk-adjusted CUSUM and RSPRT charts for monitoring in medical contexts. Stat Methods Med Res 2003 Mar;12(2):147-70.

(14) Spiegelhalter D, Grigg O, Kinsman R, Treasure T. Risk-adjusted sequential probability ratio tests: applications to Bristol, Shipman and adult cardiac surgery. Int J Qual Health Care 2003 Feb;15(1):7-13.

Further reading

Carey RG. Improving healthcare with control charts: basic and advanced SPC methods and case studies. Milwaukee, Wisconsin: American Society for Quality Press; 2003.

Glossary

Binomial distribution: Applies to the probability distribution of discrete data with only two possibilities - e.g. alive or dead, male or female. It is used to generate control limits for percentages or proportions.
C-chart: Control chart suitable for plotting counts of (adverse) events, where the opportunity for them to occur can be assumed equal from one time period to the next.
Common-cause variation: Random variation which is inherent and to be expected in a system and does not require any further explanation.
Continuous data: Variables which can take any value within an uninterrupted range - not just whole numbers. E.g. temperature, weight. (Opposite of discrete data.)
Control chart: Chart which plots the variable of interest against time, or for different institutions at the same point in time, with control limits superimposed.
Control limits: Upper and lower limits within which variation can be considered 'common-cause'. Any variation beyond these limits is 'special-cause'.
Controlled system: A system exhibiting only common-cause variation.
Direct standardisation: Method of working out the overall number or rate of deaths, etc, that would result locally if the population followed a standard (hypothetical) age profile. Allows comparison between local populations of different age structure. E.g. DSR (Directly Standardised Rate).
Discrete data: Variables which can only take specific distinct values - often whole numbers. E.g. number of operations, number of errors. (Opposite of continuous data.)
Funnel plot: Control chart showing data for different institutions at the same point in time, arranged so that their control limits become narrower from left to right.
I-chart (or XMR): Control chart presenting a continuous variable for 'individuals'. However, in practice the 'individual' is often a single day, week, month etc.
Indirect standardisation: Method of working out the overall number of deaths, etc, that would result locally if the standard (e.g. national) rate for each age-group prevailed. Allows local populations of different age structure to be compared with the standard. E.g. SMR (Standardised Mortality Ratio).
Normal distribution: A familiar bell-shaped curve, which is a good representation of the distribution of many naturally-occurring variables.
Outlier: An observation which is so unusual as to signal special-cause variation.
Overdispersion: A situation where an excessive number of observations are outliers.
P-chart: Control chart plotting the proportion of observations meeting some criterion. E.g. proportion of each week's blood samples which test positive (where number taken varies from week to week).
Poisson distribution: Statistical distribution which applies to discrete data concerning the number of events (e.g. accidents) in a fixed space of time.
Special-cause variation: Variation beyond that which is inherent in a system under control, implying that the process has changed.
Standard deviation (SD): Statistical measure of the amount of variation around the mean.
Statistical Process Control (SPC): Methodology for monitoring and improving systems, originating in industry and characterised by the use of control charts.
Traffic lights: Use of red, amber and green ratings to classify performance.
U-chart: Control chart for plotting the rate of events per time-period, where the opportunity for these to occur is not constant. E.g. if bed occupancy varies from month to month, patient falls would be divided by the number of patient days before plotting.
X-bar and S chart: Control chart presenting a continuous variable for subgroups of more than one observation (e.g. turnaround time for a daily sample of blood tests). Actually two charts - one of subgroup mean (X-bar), plus one of within-group standard deviation (S).

About the Association of Public Health Observatories (APHO)

The Association of Public Health Observatories (APHO) represents and co-ordinates a network of 12 public health observatories (PHOs) working across the five nations of England, Scotland, Wales, Northern Ireland and the Republic of Ireland.

APHO facilitates joint working across the PHOs to produce information, data and intelligence on people's health and health care for practitioners, policy makers and the public.

APHO is the largest concentration of public health intelligence expertise in the UK and Republic of Ireland, with over 150 public health intelligence professionals.

APHO helps commissioners to ensure that they get the information they need and our websites provide a regular stream of products and tools, training and technical support.

We work with partners to improve the quality and accessibility of the data and intelligence available to decision-makers.

We are constantly developing and learning new and better ways of analysing health intelligence data. We use these new methods to improve the quality of our own work, and share them with others.

Updates and more material, including methods and tools to support our Technical Briefing series are available through our website at http://www.apho.org.uk

For further information contact:
Association of Public Health Observatories
Innovation Centre, York Science Park,
Heslington, York, YO10 5DG

Telephone: 01904 567658
http://www.apho.org.uk