

Clinical governance in action

Using statistical process control (SPC) chart techniques to support data quality and information proficiency: the underpinning structure of high-quality health care

David S Simpson MSc PhD
Head of Primary Care Informatics, Teesside Primary Care Informatics (TPCI)

Tony Roberts BA (Hons) MA MSc
Head of Clinical Effectiveness, North Tees Primary Care Trust (NTPCT) and Honorary Research Fellow, South Tees Hospitals Trust and the University of Durham

Chris Walker
Informatics Advisor

Kevin D Cooper
Informatics Advisor

Fiona O'Brien
Informatics Support Officer

Teesside Primary Care Informatics (TPCI), Middlesbrough, UK

ABSTRACT

Healthcare process measurements are routinely completed in the NHS. Statistical Process Control (SPC) techniques, when applied to such data, can be used as a basis for quality improvement in healthcare processes in much the same way as they have been effectively applied to manufacturing. WA Shewhart recognised in the 1920s that a process can contain two types of variation – that due to random causes and that due to assignable causes (i.e. random (common) or assignable (special)). WE Deming later derived the expressions 'common cause variation' and 'special cause variation' – common cause variation is an inherent part of all processes; that is, it is ever present. Special cause variation is that due to things that really weren't part of the way the process was designed, and which somehow artificially find their way into it. Since a reason for its presence can be identified, its effect on the process is usually infrequent and can often be eliminated, but the effect on outcomes can be huge. Healthcare data should be used to guide quality improvements, and the role of SPC in this process is to identify assignable (special) causes and understand its origin (it should be prevented if bad and spread if good). This graphically informative approach of presenting health data is an alternative method to conventional things such as performance league tables for presenting outcomes, because tables with only common cause variation tend to encourage unwarranted tampering, may lead to local special cause variation being ignored, tend to encourage the 'blame culture', and are not linked directly to improvement activity.

Keywords: data quality, healthcare data, informatics, information quality, quality improvements, statistical process control, variation

Introduction

The National Programme for IT (NPfIT) represents a huge change in how the NHS works and it is crucial that organisations are prepared for this. Implementing new General Medical Services (nGMS) indicators and ensuring National Service Framework (NSF) targets are maintained are also key priorities. Overlaying a major change programme onto such areas is challenging. When the change also requires new competencies, such as informatics, to be generated for significantly large numbers of staff, that challenge assumes enormous proportions. Clearly the whole informatics agenda is changing within the NHS and it is important to ensure these changes are dealt with. This is an evolving discipline that deals with the collection, recording, storage, retrieval, analysis, communication and optimal use of data, information and knowledge, but without standardisation and consistency in its approach it will fail to maximise investments.

Background

Teesside Primary Care Informatics (TPCI) is hosted by North Tees Primary Care Trust (PCT) and is tasked with supporting 59 general practices – serving a population of approximately 400 000 people – and three PCTs: Hartlepool (HPCT), Langbaurgh (LPCT) and North Tees (NTPCT). The service provides equal access to high-quality, resource-effective support designed to facilitate ongoing improvements in data quality and standardisation, and to develop information proficiency skills at both an individual and practice level. To enable this, it provides support on all aspects of clinical computer systems and facilitates the collection and analysis of a range of differing data through the development and use of appropriate tools. Thus, in conjunction with a formal data collection process, data are consistently and seamlessly extracted, analysed and formally presented back to practices. This holistic approach (with data and information quality at the core) has enabled practices to make significant improvements and, most importantly, to bring those practices with the least resources up to a more equitable standard. The core work of Teesside PCI supports both local and national initiatives by promoting data and information skills of the highest standard. This supports the development of healthcare processes, and the promotion of an information culture, which is the foundation of integrated care records, the national data spine, and high-quality patient care. Implicit within this approach is the belief that developing a more structured methodology to care within the complex world of health care requires better management of information and is underpinned by data quality.

Statistical process control (SPC)

Healthcare process measurements are routinely completed in the NHS. Statistical Process Control (SPC) techniques, when applied to such data, can be used as a basis for quality improvement in healthcare processes in much the same way as they have been effectively applied to manufacturing.[1] WA Shewhart, while working for Bell Telephone Laboratories in the 1920s, recognised that a process can contain two types of variation – that due to random causes and that due to assignable causes (i.e. random (common) or assignable (special)).[2] WE Deming later derived the expressions 'common cause variation' (to mean the variation due to random causes) and 'special cause variation' (to mean the variation due to assignable causes).[3]

The first thing to understand is that variation exists in all processes – i.e. things change. Take for example how we manually write a series of the letter 'a'. Is it always the same? Some are better than others and the resulting letters can be placed in a rank order such as a league table. The question is, what went wrong with the bad ones and what was right with the good ones? Nothing was planned, just some came out well and some didn't. Thus ranking can be unhelpful because the row of 'a's were created by the same person, in the same place, at the same time, with the same pen, the same paper, etc. Yet variation is ever present and the 'a's are not exactly the same. However, it is absurd to ask what was done differently. The variation is due to the complexity intrinsic to the process of writing and is caused by factors that are inherent in the system over time – i.e. they affect all outcomes – and in SPC jargon this is called common (unassignable) cause variation. If there is only common cause variation in the process, it is said to be in control and stable. That is, it is predictable within limits, meaning that the probability of any future outcome falling within the limits can be stated (i.e. it is intrinsic to the process). Conversely, if the process contains special cause variation it may be unstable and unpredictable. It may also highlight an improvement in the process that can be shared. The task is to identify the cause and understand its origin, thus preventing it if bad and spreading it if good.

Common cause variation is an inherent part of all processes; that is, it is ever present. The effect of this type of variation is usually minimal and results from the regular rhythm of the process. An example of this is patients visiting general practice surgeries; since everyone is different, so is their outcome. Special cause variation is that due to things that are unique to the original process design, and which somehow 'artificially' find their way into it. Since a reason for its presence can be identified, its effect on the process is sometimes infrequent, but the effect on outcomes can be huge. As an example, take the case of a clinician, and assume that this clinician exhibits poor practice when compared to his colleagues. Immediately one can deduce that the observable variation within the process might be attributed to poor methods. This variation would happen infrequently but would have a large effect on outcomes and, as such, is often artificially a part of the process.

SPC tools

Two of the most popular SPC tools in regular use are the time chart and the confidence chart. If comparing a single individual, unit or department over different periods, a time chart may be helpful, whereas, if comparing different individuals, units or practices, etc, over a single time period, a confidence chart is best. Both charts are relatively easy to interpret, since there are only a few basic rules to remember in order to identify the variation type without the need to worry too much about the underlying statistical theory.

Time charts

A time chart consists of a time-ordered sequence of data (see Figure 1), with a centre line drawn horizontally through the chart. The chart enables the monitoring of the process level and identification of the type of variation in processes over time.

Confidence charts

A confidence chart is an ordered sequence of data (see Figure 2), with a centre line, calculated using the mean of the data, drawn horizontally through the chart. In addition, upper and lower control limits are added to the chart, the effect being to draw a 'trombonogram' which highlights common and special cause data variation. As with the time chart, a confidence chart enables the monitoring of process levels and identification of the type of variation in a process over time. The extra power of this chart over its counterpart for detecting special cause variation in healthcare data comes via the additional rules associated with the control limits. In general, these limits are locally agreed and are commonly set at 3 × sigma (the same as standard deviation, but calculated using formulae that take account of the order of the data). So, broadly speaking, confidence charts are essentially simple graphical tools that enable the monitoring of current process performance and are designed to identify which type of variation exists within a process.

TPCI SPC Builder

TPCI, as part of its commitment to the facilitation of healthcare informatics and in an effort to illustrate and maximise the underlying value in health data, has worked in conjunction and with the advice of the head of clinical effectiveness (evidence-based medicine specialist also linked to a local acute trust and a regional university) from one of the PCTs it supports and developed a mechanism to ease the provision of data in a more effective and understandable manner.

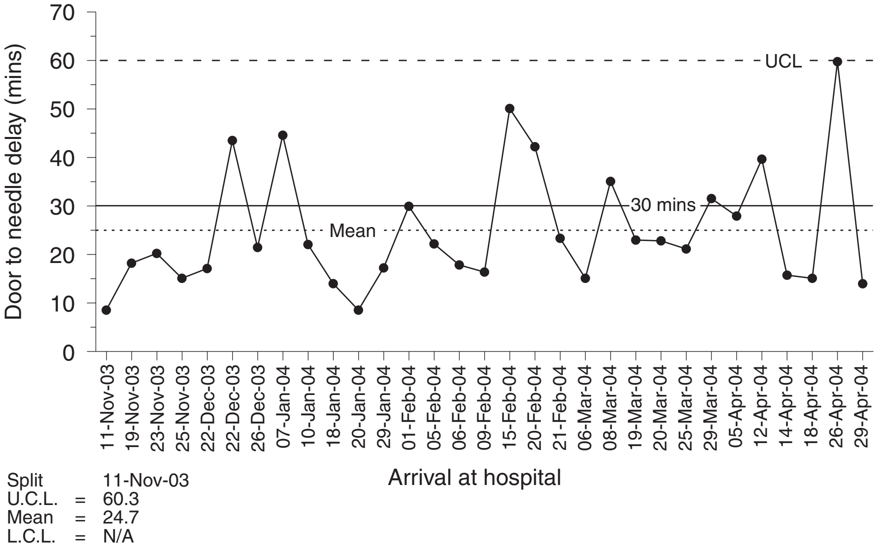

Figure 1 Time chart example highlighting delays to thrombolysis in minutes

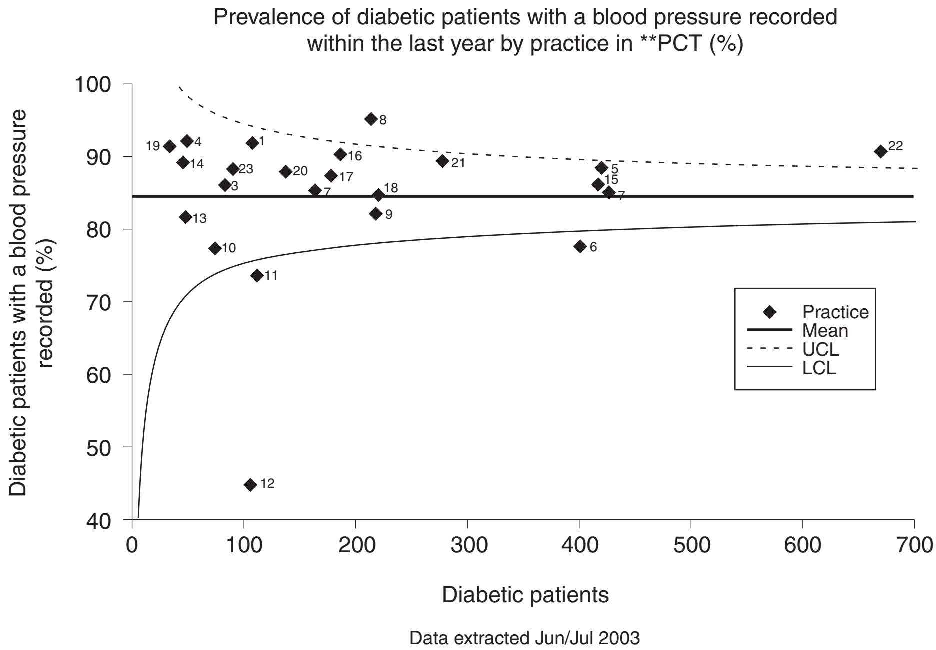

Figure 2 Confidence chart example highlighting upper and lower confidence limits. UCL, upper confidence limit, LCL, lower confidence limit

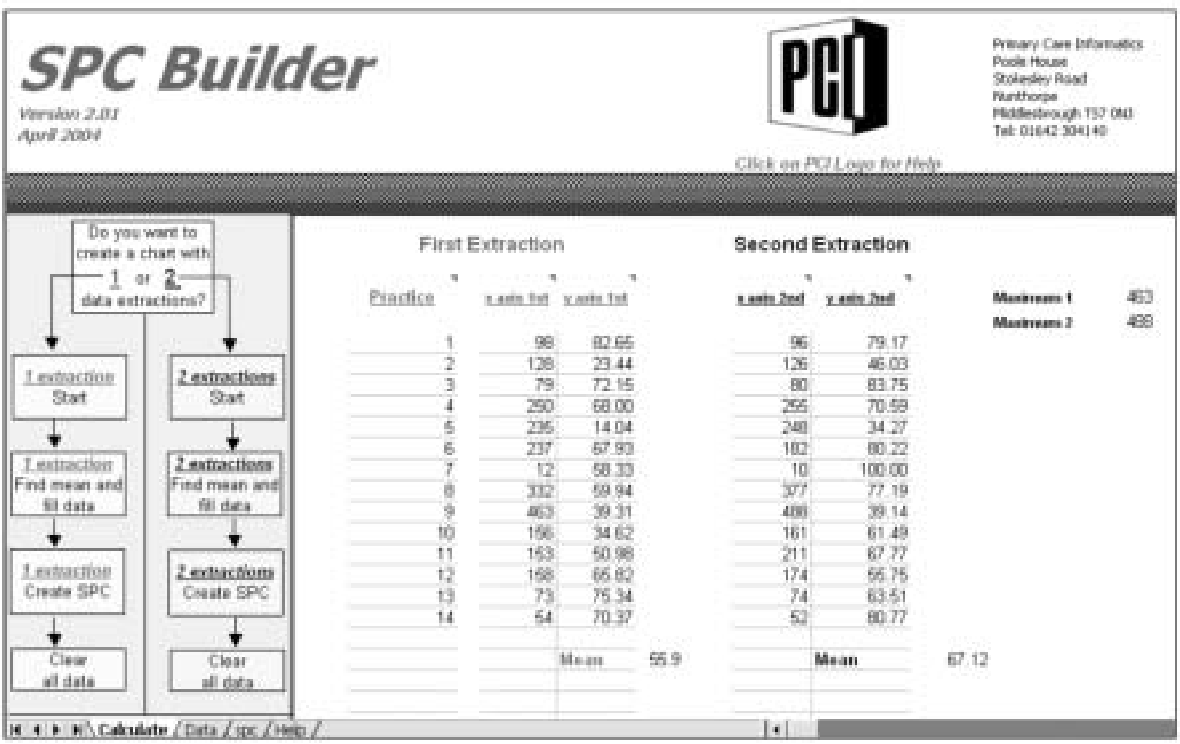

Figure 3 Example of SPC Builder supporting comparative data analysis

In short, this is achieved through the manipulation of SPC confidence charts that highlight data and information quality issues and support improvements in clinical working practice amongst other things. To ease the process, a piece of software has been developed in-house – making it easier to gain clinical acceptance by delivering speedier and more transparent information benefits – and has been made freely available to all interested parties both locally and nationally (see Figure 3). Furthermore, the software has been enhanced by enabling the user-driven option of either single or comparative (the most recent data extract 'overlaying' the previous one – see Figure 4) charts. For clarity and consistency, all TPCI data extracts are now supplied

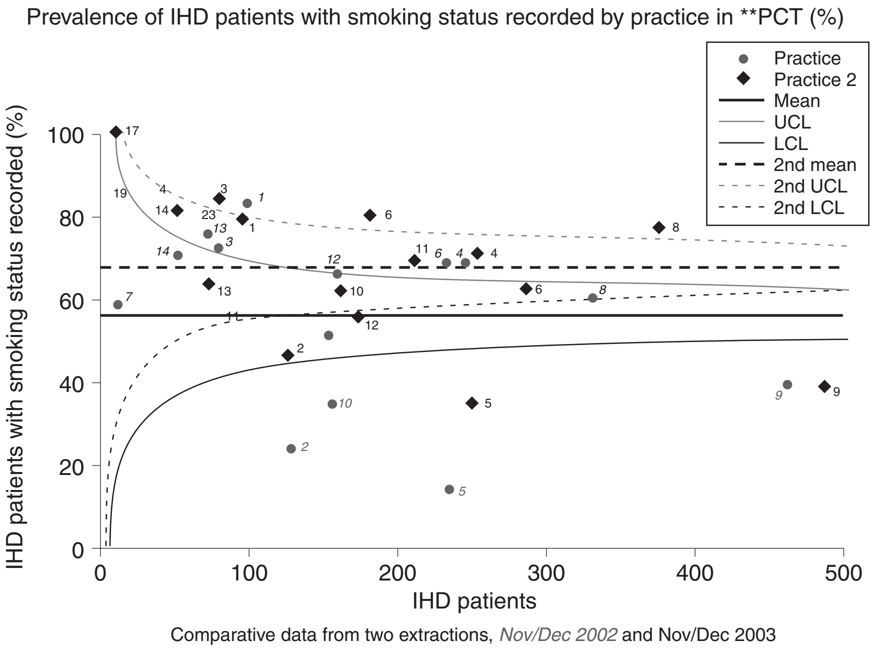

Figure 4 Example of SPC Builder highlighting comparative data analysis of IHD patients' smoking

back (as far as possible) in this visually amplified manner (i.e. SPC), and this includes formal feedback (engendering user engagement), the development of action plans, and SPC training sessions (with ongoing evaluation measuring true effectiveness) as appropriate.

The approach has been positively accepted by all users and promotes both clinical engagement and dialogue. One example of the successful implementation of this mechanism is when special cause variation was identified in a practice involved in the PCT wide data extraction around coronary heart disease (CHD). SPC charts highlighted a fundamental problem (essentially) concerning an over recording of ischaemic heart disease (IHD). Once this was tracked back, actioned and remedied as appropriate, the data extract in the specific practice was repeated and 'new' charts were generated. This had the effect of positively highlighting the impact of a team-based intervention and bringing the data 'up' and more equitable across the entire PCT. This has immense benefits because, amongst other things, it graphically illustrated to key personnel that the work undertaken has been beneficial and supports high-quality healthcare.

TPCI SPC Builder: the process

The conventional way of constructing an SPC is time-consuming and complicated. For example, the mean of a series of data is used as a basis to firstly calculate 3 × sigma and then further formulae are used to calculate the upper and lower control limits. When all these calculations have been recorded, Microsoft Excel can be prompted to create a chart and the 'wizard' tool will guide the user through the choice of settings for the chart. SPC Builder enables the user to simply and automatically create the same chart, but without the need to understand the statistical processes, and in a fraction of the time. The user is offered the choice of creating a single or comparative SPC (see Figure 3) and is then led through the building process by a series of three command buttons. Using raw data that the user is prompted to enter, the command buttons activate macros which seamlessly allow the statistical calculations to run in the background, creating an individualised SPC. A fourth command button clears all data, calculations and SPC in preparation for SPC Builder's future use.

As a practical example of building an SPC, consider the prevalence of diabetic patients with a record of cholesterol measured within the last 12 months. After clicking 'Start' in SPC Builder, a pop-up box would ask how many practices are to be included, the user would then enter a number. A list appears in a column with labels for Practice 1 through to the number entered. A further pop-up box would then prompt the user to enter or copy and paste values for the x axis (population, i.e. diabetes) and the y axis (prevalence, i.e. the percentage of diabetic patients who have a record of cholesterol). The user is also prompted to alter the practice identifiers if required. Clicking on the second and third buttons will allow the calculations to be made in the background and the SPC to be created.

Healthcare information proficiency

The development detailed here is an innovative (and therefore somewhat controversial) approach to analysis using statistical process control theory. Clearly, if performance data is constructed as a league table, it can be shown that the rank order changes, particularly in the middle of the league table, when potentially confounding variables such as age, sex, health, employment status, etc, are used to adjust the estimate.[4] This is what statistical process control theory tells us to expect and counsels against doing. Instead, a simple scatter plot of percentage against the denominator should be plotted. The degree of scatter can then be visually compared to the binomial distribution for a sample with an equal average and limits calculated at $3 \times$ sigma points to identify special cause variation in any outliers. If the analysis is repeated using a percentage figure adjusted using the potentially confounding variables and any outlier was within the limits, the likely cause of the special cause variation is identified. If any outliers remained after adjustment, special causes (perhaps actual differences in performance) need to be sought. If all variation is in fact within limits, either on crude analysis or after adjustment, only common cause variation exists. This approach has been used successfully to look at such things as general practice variation around clinical interventions, nGMS indicators, NSF measures, prescribing rates, referral patterns, and hospital inpatient and outpatient utilisation rates, etc. This helps organisations to understand where they stand in relation to their peers (even when it only tells them that some aspect of their data is poor quality). It is also invaluable for understanding patterns of healthcare provision and linking to such things as deprivation and access to services.

Conclusion

Medicine is recognised as both art and science. Informatics solutions must similarly allow for flexibility and choice. To aid this, application solutions should be intuitive and the transition from current organisational, technological and human situations requires evolution not revolution. Informatics can, and will, deliver support to more effective health maintenance but only if its use is enthusiastically grasped by professionals. This will only happen if all involved understand and are committed to the outcomes, and this is best supported by the development of information proficiency.

Data and information quality are recognised as crucial to the provision of health care and are major strands in supporting NSFs, nGMS and the NPfIT. Thus, high-quality information is vital to safe and effective health care and for this reason accurate, timely data is crucial. Models of care are changing and broadening to rely upon interdisciplinary care pathways and clinical networks that go beyond organisational boundaries. Information is increasingly used in new and innovative ways, in a range of different settings, underpinning the continuity and quality of care. These developments create pressure to improve the quality and standardisation of existing processes, but also lead to new and greater availability demands.

Over the past 30 years, attempts have been made to introduce universal electronic medical records. Through NPfIT there is now a centrally driven concerted effort to have an electronic record for every patient. Among the benefits this will offer is the potential for a fundamental improvement in both data and information quality. However, without agreed standards and approaches, computerisation can significantly increase the risk of poor quality data leading to misinterpretation and adverse events. Thus, the skills needed to effectively realise the benefit from improved information quality must be developed and consistently maintained.

In the past, those in possession of data might have opted for inaction or called for better data. Recent high profile cases have contributed to conditions where the tendency for action will be more frequent. It is perhaps optimistic to suggest that use of control charts could prevent the recurrence of tragic and unfortunate episodes such as Bristol or Shipman. What is clear is that analysing data with an understanding of common cause and special cause variation provides the NHS with a basis to act. Thus, healthcare data should be used to guide quality improvements and the role of SPC in this process is to identify assignable (special) causes and understand its origin (it should be prevented if bad and spread if good). Mohammad and Adab have constructed confidence charts of the kind detailed here for healthcare data, and the NHS Modernisation Agency/collaborative services and the Commission for Healthcare Audit and Inspection (CHAI) favour its promotion as the data presentation method of choice.[5,6] This graphically informative approach of presenting health data is an alternative method to conventional things such as performance league tables for presenting outcomes, because tables with only common cause variation tend to encourage unwarranted tampering, may lead to local special cause variation being ignored, tend to encourage the 'blame culture', and are not linked directly to improvement activity.

The benefits from the cohesive and structured approach adopted and detailed are essentially different from other initiatives because it works alongside and in support of healthcare professionals, is completely transferable, supports value for money (through the maximisation of data) and embraces the wider information quality issue. The result is that data quality is supported to 'seamlessly' fall out of the process, as users become more confident with the skills developed around information proficiency. The management of change is thus supported, encouraging continuous development, empowering both individuals and organisations to take control, and ensuring a dynamic informatics structure and patient centre rationale.

The next step would be to develop the software further and, among other things, increase its usability and make its use interactively available via the TPCI's website,[7] so users could (easily) upload data, generate SPC charts, and download the results as and when required. This would encourage a whole-systems approach to both healthcare data and information, and help to realise the benefits of both the systems and resources (i.e. people) required to utilise them. Additionally, it would be useful to continue to seek the value in healthcare data and use the mechanism to map other data items together such as deprivation and service provision, disease prevalence and prescribing patterns, etc – thus supporting genuine holistic health care. Furthermore, the forthcoming creation of the 'Secondary Uses Service' – which will enable access to consistent, patient-based data to support performance improvement and assessment, clinical audit and governance, monitoring and benchmarking, surveillance, research and planning – will fail to achieve its potential if the importance of information quality is understated, and tools (such as SPC) must be utilised to support the ongoing development of information proficiency and to highlight all the benefits.

Note: The software detailed in this paper – TPCI SPC Builder – is freely available to all interested parties, both locally and nationally, either by direct download from the TPCI website: www.PrimaryCareInformatics.co.uk or on request using the correspondence details provided.

REFERENCES

1 Tekkis PP, McCulloch P, Steger AC, Benjamin IS and Polonieck JD. Mortality control charts for comparing performance of surgical units: validation study using hospital mortality data. *British Medical Journal* 200;326:786–8.
2 Shewhart WA. *Economic Control of Quality of Manufactured Product*. New York: D Van Nostrand Company, 1931. (Reprinted by Milwaukee: ASQ Quality Press, 1980.)
3 Deming WE. *Out of the Crisis*. Massachusetts: Massachusetts Institute of Technology, 1986.
4 Lilford R, Mohammed MA, Spiegelhalter D and Thomson R. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *The Lancet* 2004;363:1147–54.
5 Mohammed MA, Cheng KK, Rouse A and Marshall T. Bristol, Shipman, and clinical governance: Shewhart's forgotten lessons. *The Lancet* 2002;357:463–7.
6 Adab P, Rouse AM, Mohammed MA and Marshall T. Performance league tables: the NHS deserves better. *British Medical Journal* 2002;324:95–8.
7 www.PrimaryCareInformatics.co.uk

CONFLICTS OF INTEREST

None.

ADDRESS FOR CORRESPONDENCE

Dr David S Simpson, Head of Teesside Primary Care Informatics, Teesside Primary Care Informatics, Poole House, Stokesley Road, Nunthorpe, Middlesbrough TS7 0NJ, UK. Tel: +44 (0)1642 304167/304140; fax: +44 (0)1642 304162; email: David.Simpson@tees-shs.nhs.uk; TPCI website: www.PrimaryCareInformatics.co.uk

Received 15 October 2004
Accepted 24 November 2004