

Following up mortality 'outliers'

A review of the programme for taking action where data suggest there may be serious concerns about the safety of patients

Acknowledgements

We would like to thank the following for their valuable contributions to the work described in this report:

Professor Sir Brian Jarman, Dr Foster Research Unit, Imperial College London.

Dr Paul Aylin, Dr Foster Research Unit, Imperial College London.

Professor David Spiegelhalter, MRC Biostatistics Unit, Cambridge University.

Contents

1. Aim of the report

2. Background

What is an 'outlier' and why should we be concerned?
Being proactive – looking for outliers
Understanding our impact

3. Our approach

Step 1: Using information to trigger 'alerts'
Step 2: Preliminary analysis of the alerts
Step 3: Decision panel – do we proceed?
Step 4: Engagement with the NHS trust
Step 5: Secondary analysis

4. Results after one year

Establishment of the programme
Categorisation and review of the first year's cases
Feedback from the trusts involved

5. Conclusions

6. Recommendations

Appendix: Identifying outliers using cumulative sum (CUSUM) methodology

1. Aim of the report

This report sets out, after the first year of operation, the results of a joint programme of work between the Healthcare Commission's Investigations Team and analysts in Informatics, to follow up concerns where NHS trusts appear to have higher than expected rates of mortality.

The main aim has been to protect the safety of patients by identifying – and acting on – potential problems at an early stage. We also want to encourage learning throughout the NHS and show how the Healthcare Commission can use sensitive, complex data to bring about improvements in healthcare.

2. Background

What is an ‘outlier’ and why should we be concerned?

The Healthcare Commission’s overriding purpose is to promote improvement in the safety and quality of healthcare services provided to patients. To achieve this, we have been given explicit powers to follow up concerns and referrals, and investigate potentially serious failings in the provision of healthcare.

We recognise that modern healthcare is complex, in that it embraces a huge range of treatments, patients and different ways of delivering care. As a regulator, our remit extends across hundreds of different organisations, thousands of clinical professionals, and millions of contacts with individual patients.

Clearly, we cannot closely monitor all the details across this breadth of activity, but we have to be vigilant to where there might be a problem. To do this, we have to exploit the information that is available to us – and choose very carefully where we, as regulators, engage with individual providers of care.

A mortality ‘outlier’ arises from the analysis of information that is routinely available. It identifies where, within an organisation, the numbers of patients who have died from a particular illness or condition is significantly higher than we would expect. There are many reasons why rates of mortality may be high: we cannot and do not automatically assume that there is a problem with the quality of care. However, we use these signals to prompt further analysis. In the majority of cases, we find a plausible explanation. In some cases, we cannot explain the signal and seek further information from the trust or other sources.

We have an established process for following up outliers. This includes communicating with the organisations concerned, to establish the accuracy of the data used in our analysis and to understand whether there may be known problems with the quality of data within the organisation. For most cases, an outlier is considered to exist if differences between observed and expected outcomes are great enough for it to trigger using a statistical process control technique. We carefully and individually consider each case before we decide whether we need to follow up with the trust concerned.

Statistical process control

Statistical process control (SPC) is a set of methods that use control charts to identify significant deviations from a predefined standard. It originated in the manufacturing industry and is now regularly applied to the monitoring of healthcare.

We use a specific SPC technique called CUSUM (cumulative sum), which detects significant deviations from expected outcomes. In our analysis, expected outcomes are either derived from standardised mortality rates or from the underlying risk attributed to each patient. If the plotting of data crosses a fixed 'control limit', then a significant run of poor outcomes is detected and an alert is signalled. Even if the underlying risk for each patient in a hospital is average, the observed rate over time will vary by chance. Therefore, limits have to be set to guard against too many 'false alarms' occurring as a result of random variation.

Being proactive – looking for outliers

Clearly, it is important to follow up outliers when they come to our attention. We also believe that we should proactively identify outliers, on the basis that each of these may represent a risk to the safety of patients.

We therefore started this new programme of work to build on our experience and, using the wealth of data available to us, proactively search for and follow up statistical outliers, with the aim of identifying potential failure at the earliest possible stage. This requires a particularly thoughtful and sensitive approach, as data alone is unlikely to provide a complete insight into the complexities of healthcare. However, with the right handling, this is also a powerful way to highlight problems – and ensure they are addressed – before they build up into the sort of systemic failings described in our published investigation reports.

In developing this programme, we wanted to test out a new approach to regulation, which exploited information to the full and intervened only when necessary. We have chosen high mortality rates, as we see these as a strong trigger for further enquiries, but we could have looked at other triggers such as rates of readmission to hospital following surgery and, indeed, these are part of our plans for the future.

Understanding our impact

We realise that this approach to regulation is new in this context and we need to assess its effectiveness. This report is part of that process of assessment. It summarises the first 12 months of activity from August 2007, although some preparatory work had been undertaken before that date.

Our criteria for success were that:

1. We were able to establish a stable programme that maintained a satisfactory balance between:
 - The numbers of signals/outliers we identified
 - The depth and quality of analyses we were reasonably able to undertake
 - The timescales and levels of request for further information from trusts
 - The resources available to undertake analysis and follow up with trusts.
2. Our work should help in some way to improve the quality of care provided. Such impacts could appear in a number of guises, either directly or (more likely) indirectly, in terms of improving governance systems in the NHS.
3. This programme of work should help to identify potential problems earlier than they would otherwise have been detected.
4. The programme did not create an unnecessary burden on healthcare providers and yet provided reassurance to commissioners and the public.

3. Our approach

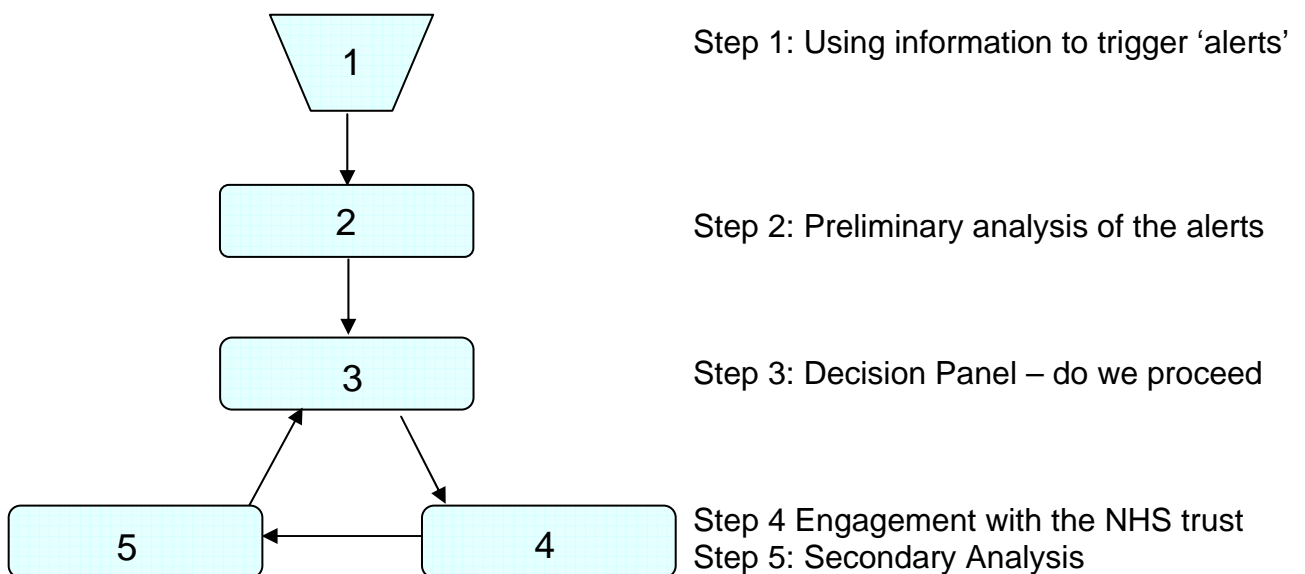
We were keen to establish a set of principles to guide our approach. In particular, we wanted to strike the right balance between being overly cautious – and risk failing to follow up an issue of genuine concern – and an over-reaction that may cause unnecessary alarm and start the frustrating pursuit of a large number of data signals that ultimately reveal nothing.

From the outset, we agreed that our approach should:

- Focus primarily on safety – following up only where the data suggests that there may be a significant risk of harm to patients.
- Be intelligent and proportionate – making use of existing information to best effect, rather than placing an undue burden on the NHS.
- Provide organisations with the opportunity to explain their situation and identify their own solutions where necessary.
- Help to drive improvements in clinical care and encourage the NHS to understand the importance of monitoring the safety of its services.
- Make best use of expert clinical and statistical advice, including information from other organisations where available.
- Ensure that every potential outlier was pursued until we were satisfied that all concerns had been properly resolved.

Figure 1 summarises our process, which is discussed in more detail below.

Figure 1: How the process works



Step 1: Using information to trigger ‘alerts’

As a regulator, we need systems that are able to respond when things might be going wrong within the health system. There are many different signs or signals that can help identify such problems. So for example, we respond to concerns from members of the public, professionals and whistle-blowers working within health services, or from government.

We can also use statistical analysis to look for the early warning signs – symptoms of a problem. In our current work, we already have access to a wide range of health-related intelligence, either from national sources or arising from our own assessments and analyses. During our work, we expect to come across information that appears to indicate poor performance that may have a serious impact on the care of patients.

Information about patients in hospitals

The Healthcare Commission uses data from a variety of sources to inform the different aspects of its work. One of the main sources of data used by the Commission is Hospital Episode Statistics (HES), a national dataset compiled by the NHS Information Centre and originating from the patient administration systems in individual hospitals or trusts. It includes a wide range of information about patterns and outcomes of treatment for inpatients in NHS hospitals throughout England. HES is used by many organisations to analyse a wide range of healthcare data for the NHS.

However, while the HES database collects very valuable information on outcomes for patients, it was not designed to collect detailed clinical information, nor does it allow for outcomes to be adjusted for a comprehensive number of risk factors. For example, in one investigation into a cardiothoracic surgical service, it was clear that there had been some concern from clinicians and specialists about the use of HES for monitoring outcomes and rates of mortality following surgery. The HES data for the period under investigation clearly showed the trust to be a significant clinical outlier for the national benchmark procedure: the coronary artery bypass graft. However, when analysing the more detailed clinical data contained in the specialist Central Cardiac Audit Database (CCAD), the trust’s results were not as concerning. However, the use of the CCAD presented its own challenges. The very detailed information required by the CCAD from the cardiac units resulted in a lack of consistency in the submission and quality of data provided. This made detailed analysis and comparison between organisations difficult.

Data such as that contained in HES and CCAD are a valuable source of information, but it is important to use them in context and to bear in mind their limitations. Any concerns identified from data analysis should be supported by other information.

By organising routine information in the right way, we are able to look for early warning signs. This proactive approach has the advantage that:

- It is systematic, in that it considers a defined set of organisations rather than selectively looking at a few.
- It looks across a wide range of patient groups and care settings.
- It can be continually updated to retain a constant vigilance.

Alerts are triggered when series of outcomes for an individual trust cross a pre-defined threshold. These thresholds are established so that, when they are crossed, there is sufficient statistical evidence to suggest that mortality is greater than expected. It is important to note that we focus our attention on identifying concerns among specific clinical groups of patients rather than trying to analyse mortality rates across an entire hospital.

Alerts for the mortality outliers programme are currently generated in four ways:

1. We analyse HES data because we have received concerns about the care of a particular patient group within a trust. The signal may relate to a trust that is not the one for which the concern was initially raised.
2. We proactively scan HES data for mortality outliers across all 610 Healthcare Resource Groups (HRGs) and admission type (elective or emergency). Expected mortality is calculated by indirect standardisation accounting for differences in age and gender. Emergency and elective admissions are considered separately. The details of our approach are described in the appendix to this report.
3. We receive an automatically generated alert from the Dr Foster unit at Imperial College, which suggests that a trust's observed rate of mortality for a particular of diagnosis or procedure is persistently higher than the expected rate (see box below).
4. As part of the process of assessing outliers from Imperial College, we analyse the relevant HES data and this may, in turn, reveal further outliers for the same groups of patients. These may not have been detected by Imperial College because of differences between the HES data and the Secondary User Service (SUS) data (see below), or because our methodologies differ.

Working with other organisations

Joint work with the Dr Foster Unit at Imperial College

The Dr Foster Unit at Imperial College London (an academic research unit, separate from the commercial company Dr Foster Intelligence) routinely analyses its own 'cleaned' version of the Secondary User Service (SUS) data for all NHS acute trusts, relating to 50 groups of primary diagnoses and 85 groups of procedures.

In the course of its routine analysis, the unit identifies trusts with significantly higher than expected mortality. An 'alert' is automatically generated when observed mortality for any of these groups of diagnoses or procedures is persistently higher than the expected rate after allowing for variable patient risk. Expected rates are risk-adjusted controlling for a range of patient-level factors:

- Age
- Sex
- Admission method (elective or non-elective)
- Socioeconomic deprivation
- Diagnosis or procedure subgroup
- Co-morbidity based on Charlson Score (capped at 6)
- Number of emergency admissions in the past 12 months
- Presence of episode under specialty of palliative care within spell
- Financial year of discharge
- Month of admission (for some respiratory diagnoses).

Following discussion with representatives of the unit, we have agreed a process for the sharing, on a regular basis, of information about potential outliers. The unit sends notification of each automatically generated alert to the trust concerned, and copies the correspondence to the Healthcare Commission. Each trust is therefore aware of the Commission's potential involvement from the outset. Alerts have been sent in this way since July 2007. Typically, they relate to information available about three months prior to the date of receiving the alert.

Reference

Bottle A, Aylin P, "Intelligent information: a national system for monitoring clinical performance", *Health Services Research* 2008;43:10-31.

We have made efforts to develop links with other bodies wherever there is a possibility of gaining access to other, more specialist data sources that could help us make more informed decisions. For example, the organisation UK Transplant undertakes monitoring and audit of a specific range of transplant

procedures, and holds more meaningful, and more up-to-date, information about mortality rates following transplant surgery, than would be available through HES.

Similarly, we have worked with CEMACH (the Confidential Enquiry into Maternal and Child Health) to make use of its advice and specialist data on, for example, maternal and perinatal mortality. Trust-level data has only been used after we have obtained the agreement of the trust and consultants involved.

We have also developed constructive working relationships with the Society for Cardiothoracic Surgery in Great Britain & Ireland, and the Department of Health, in relation to using information from the Central Cardiac Audit Database (CCAD), and also worked with the Myocardial Ischaemia National Audit Project (MINAP), in relation to using data from its extensive audit programme.

Step 2: Preliminary analysis of the alerts

All alerts that suggest there may be a high rate of mortality are formally logged as part of the programme, and details are immediately copied to the relevant regional staff team with a standard letter asking if the region has any relevant local intelligence that may be able to help with our consideration (a two-week period is specified for the submission of local intelligence).

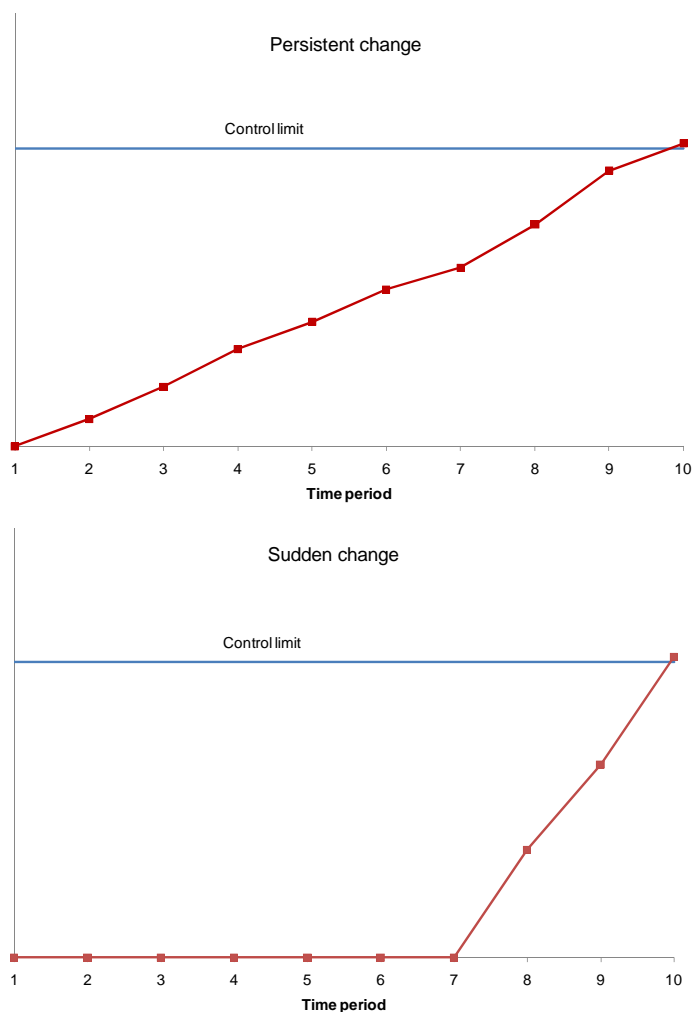
In the meantime, analysts at the Healthcare Commission consider the alerts in relation to a number of factors, such as the strength and reliability of the original signal, the statistical likelihood that it is a false alarm, and whether there is anything unusual about the trust's data when compared to its peers. Any additional intelligence that we may have relating to that trust or the particular diagnosis/procedure will also be taken into account, such as any known data quality problems, trust specialisation or recent changes in the organisation that might affect the provision of care.

The different stages of analysis are as follows:

a. Drawing any conclusions we can from the observed outcomes for this specific group of patients

A trust signals as an alert by fulfilling certain statistical criteria comparing observed and expected outcomes. The nature of the alert may be important: for example, despite the signal, there may be a high likelihood that it is a false alarm. Observed outcomes may have been persistently higher than expected over a long period or there may have been a recent change in the pattern of outcomes (see figure 2). Organisational factors, case mix issues or consistent data quality issues are more likely to show up as a persistent trend. Apart from highlighting a deteriorating quality of care, apparent sudden increases in mortality may reflect changes in the organisation or changes in coding practice. These, in turn, may be easier to focus upon.

Figure 2: Different patterns of outcomes



b. Identifying any similar concerns from other data sources

A trust may signal as an alert using one data source, but this may not be verified from another available source. If our primary source is recognised as potentially unreliable for a particular group of patients, then it has been important to identify and gain intelligence from other sources, such as CCAD for cardiology and cardiac surgery or CEMACH for maternal and perinatal mortality, as previously mentioned.

c. Making an initial assessment of the patient group

These assessments are first impressions to steer the initial analysis and are based on the experience and knowledge that we may have built up about the clinical codes that define a particular group of patients. For example, we may be able to reach relatively quick judgements about the reliability of data by referring back to previous analyses.

d. Searching qualitative intelligence

These searches make use of intelligence already gathered by the Commission in relation to the trust and the nature of any recent

engagements, particularly by field staff. We also look for basic information about the organisation and the services they may provide for the patient group concerned. Sometimes, more detailed searches are required that make further use of external intelligence: for example, published trust board minutes or local press cuttings. Such information can be highly important in directing the analysis or focusing any engagement with the trust.

e. Assessing the quality of data

Most of the issues around the quality of data centre on clinical coding practice and how it varies between trusts. Some primary diagnoses do not map well onto the ICD10 codes that are used by HES. If a trust has a tendency to code higher risk patients to a specific code than the country as a whole, it would have a greater chance of appearing as an outlier. Also, a trust that is poor at coding co-morbidities may have a disproportionate number of patients within a 'without co-morbidities and complications' coding group. Other trusts may over-use 'general' (that is, non-specific) codes, which could in turn reflect poorly written patient notes.

f. Detailed assessment of case-mix or other organisational issues

Although certain risk factors are accounted for in the methods used to identify outliers, it is not by any means comprehensive and a trust may appear as an outlier because of insufficient risk adjustment. This may especially be the case with some trusts that act as specialist referral centres for particular conditions. Outcomes for patients within a particular trust are also linked to factors such as length of stay: a trust that has problems discharging patients to community care is likely to have a higher in-hospital mortality. Appropriate benchmarking and soft intelligence from other sources are important parts of this assessment.

g. Analysing related groups of patients

Whether or not it would be appropriate to analyse further groups depends on what has been found out so far with regards to the patient group, patient risk and data quality. At this stage, we should be clearer about what we aim to achieve by such analyses, for example confirmation of data quality or case-mix issues, a focus for engagement with the trust.

Some of the key issues are whether we are observing the consequence of poor quality of care among a specific subgroup or, conversely, the consequence of wider failures in the system. If the outlier relates to a poorly defined group, we may be able to identify related groups of patients that are more homogeneous or receive similar care.

h. Specific patterns from records of individual patients

There may be indications that poor outcomes could be due to a particular aspect of the care process that can be identified in individual patient-level records in the HES data.

i. Cross-checking against relevant process and outcome measures

The Healthcare Commission has access to many different process and outcome measures, for example, emergency readmission rates, infection

rates, pre-operative length of stay and levels of service provision. Some of these are used in core standards assessments, and information from these measures may strengthen our concerns or provide a focus for engagement.

Not all the stages presented in this summary will be carried out in all cases. For example, some cases can be closed internally after very little analysis. In others, the nature of our engagement with the trust becomes clear early on. Also, the order in which these activities are carried out can vary according to the case. The evidence and decisions are all documented.

Resources for internal analysis

The internal analysis phase, Step 2, has proven to be one of the most resource intensive parts of the process in the first year. It is also the area where we have made most improvements. Though the amount of analytical attention is not the same for all cases, we have reached a stage where an initial report for the decision panel is typically produced in less than two person days.

Significant amounts of time have been devoted to developing a series of tools and methods, as shown in table 1.

Table 1: Tools and methods that have been developed

Time series analyser	Automated Excel tool to generate time series graphs and runs SPC across all trusts
Funnel plot generator	Excel tool to produce cross-sectional standardised mortality comparisons that account for expected variables in the sample population
Drill down program	Generates breakdown tables to do subset analyses/cross-tabulations by a range of variables, eg diagnoses; procedures; age; admission method; year
HES requester	An automated query tool that generates series of standardised indicators from HES
Co-morbidity extractor	A tool that allows for individual patient level data to be extracted from HES
PINGO	A program to undertake scanning of HES data, on a quarterly basis, to identify outlying values. This uses the time series analyser across all HRGs and all trusts
Outliers management database	Bespoke database to record and monitor progress individual cases
CUSUM simulator	Used to estimate the likelihood of a CUSUM signal occurring by chance alone over a specified period (false discovery rates)

Step 3: Decision panel – do we proceed?

A multidisciplinary decision panel considers every case that arises as part of the outliers programme. The panel includes staff from the Investigations Team, analysts and the Commission's senior medical advisor. It is generally able to consider cases within two to four weeks of receipt. The preliminary analysis is presented in the form of a summary report, enabling each case to be discussed and considered in detail.

The panel makes a decision on the most appropriate way forward in each case. Examples of decisions include:

Further internal analysis: Following discussion, analysts may be asked to examine a specific issue in more detail and asked to return to the panel at a later date.

Seek advice or information from another agency: This situation has arisen when we are aware of specialist information or expertise that may be relevant to the issue at hand. For example, and as mentioned in the previous section, rather than using HES data in relation to certain forms of cardiac surgery, we have sought more detailed information from the Central Cardiac Audit Database.

Do not pursue further at this stage: In about half the cases (see later), the decision is taken not to pursue the case any further. If the trust is already aware of the concerns, we will always write to them with an explanation of our decision and share our analysis, inviting comments. We also copy the correspondence and analysis to the Commission's regional team, for information.

Pursue the matter with the trust: For about half the cases, there has been a decision to engage with the trust, usually through a standard letter outlining the information we have, together with a summary of the analyses we have undertaken to date. In some cases, the panel will wish to focus on particular issues arising during discussion of the analysis. For example, apparent differences in mortality rates between different sites within the same trust.

Step 4: Engagement with the NHS trust

Wherever we have outstanding concerns, we contact the trust for further information and proceed much as with any case referred for investigation. Our approach is tailored to the situation and the evidence we have, adapting the focus of our questions and the level of information requested in each case. This is informed by our preliminary analysis of the data, which we will also share with the trust where we think this will be helpful. All formal information requests are carefully checked to ensure that they are clear and accurate, and the trust is given a specific timespan in which to provide a response. We also ensure that the trust can contact us by telephone or email if they need further clarification or wish to discuss any aspect of our request.

We try to strike a balance between sharing our information and analysis with the trust, and yet not prejudging any conclusion the trust may reach. We also aim to acknowledge any problems or inherent weaknesses in the data, as shown in this extract from one of our letters:

“We acknowledge that it is difficult to make comparisons between trusts for this particular group of primary diagnoses. However, we would like some further information regarding the raised mortality at your trust noted this letter, particularly focusing on the period between 06/07 Q4 to 07/08 Q2.”

In some instances, we have decided to meet formally with senior managers from the trust, if the issues are particularly complex or require detailed discussion. As with any case managed by the Investigations Team, we assess potential risks to the safety of patients at different stages and do not hesitate to escalate any matter where urgent action may need to be taken.

Step 5: Secondary analysis

The responses from trusts are further analysed and then formally considered by the decision panel.

At this stage, we are keen to ensure that the trust has provided adequate evidence to allay any possible concerns we might have. In general we look for:

- Evidence that the trust has given serious consideration to the questions raised, and addressed all of the matters raised.
- Robust evidence in support of the trust’s arguments. Arguments made by the trust must be reasonable and, ideally, can be verified using statistical techniques. Anecdotal views are unlikely to be sufficient.
- An indication of whether the trust was already aware of the issue and already taking some form of action.
- Evidence that the trust is making reasonable clinical judgements, as viewed by our specialist advisors.
- Assurance that the trust is monitoring its own rates of mortality (and related information) and undertaking reviews where necessary.

We consider the response of the trust and any actions that may already be in hand, taking further expert advice as necessary. In some cases, we will wish to make formal recommendations for improvement, or escalate the concerns, possibly resulting in an investigation. We have recognised that, as part of this process, it is essential for the Investigations Team and the analysts to work in close partnership, as many of the decisions require careful interpretation of the data as well as an analysis of clinical or management actions.

4. Results after one year

Establishment of the programme

Within the Healthcare Commission, the programme encompasses teams from Informatics and Investigations. The processes have evolved over the year. One of our early concerns was that it would be difficult to ensure the right balance between the sensitivity to respond to signals within the resources available, the danger being that this type of analysis can consume a massive amount of resources.

We have been able to establish a stable programme, maintaining a satisfactory balance between:

- The numbers of signals/outliers we identified.
- The depth and quality of analyses we were able to undertake.
- The timescales and level of request for further information from trusts.
- The resources available to undertake analysis – and follow up with trusts.

Categorisation and review of the first year's cases

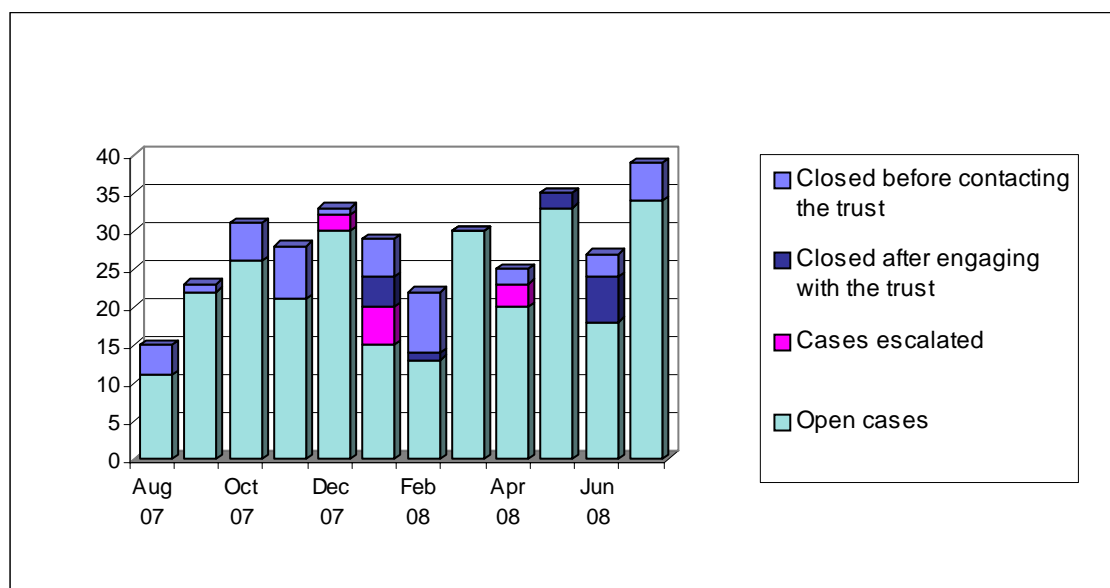
During the time covered by this report, August 2007 to July 2008, we have considered 85 alerts. Table 2 below shows the source of the alerts and the decisions made about each case when it was first considered. The rationale for 'closing' our interest for 43 of the alerts is shown in the top section of the table, with the most frequent instances being either a weak signal or further knowledge about the patient case-mix or organisation. Where trusts were aware of the alert, having received a letter about it from the Dr Foster Unit at Imperial College, a further letter was sent to them from the Healthcare Commission informing them of the Commission's decision to close the matter. In such cases, trusts also received a copy of the summary analysis undertaken by the Commission.

In 42 of the cases, we pursued the alert with the trust concerned. The reasons for pursuing these cases are summarised in the second section of the table.

Table 2: Summary of decisions made by panel (first 85 cases)

INITIAL DECISION by PANEL		<u>Imperial College</u>	<u>Self Generated</u>	<u>Internal</u>
Close	43	19	23	1
Close - other party involved	1		1	
Close - weak signal	19	8	10	1
Close - data artefact - poor coding	2	2		
Close - data artefact - poor coding - weak signal	1	1		
Close - data artefact – case mix/ organisational issues.	15	5	10	
Close - data artefact - case mix/ organisational issues - weak signal	3	1	2	
Close - unconfirmed by other data source	2	2		
Pursue	42	26	15	<u>1</u>
Pursue - suspect data artefact but require confirmation/ explanation	10	5	4	1
Pursue - unable to explain as an artefact	23	15	8	
Pursue - trust already aware	1	1		
Pursue - more general, other concerns/ multiple alerts	8	5	3	
TOTAL	85	45	38	2

Figure 3: Number of outlier cases by month, August 2007 to July 2008



Note: 'Cases escalated' refers to outliers that have led to a wider analysis of mortality, or an investigation, at the trust concerned. During the period July 2007 to July 2008, 1 wider analysis and 1 investigation have resulted from recurring outliers at two different trusts.

Outlier cases per month	Aug 07	Sep 07	Oct 07	Nov 07	Dec 07	Jan 08	Feb 08	Mar 08	Apr 08	May 08	Jun 08	Jul 08
Closed before contacting the trust	4	1	5	7	1	5	8	0	2	0	3	5
Closed after engaging with the trust	0	0	0	0	0	4	1	0	0	2	6	0
Cases escalated	0	0	0	0	2	5	0	0	3	0	0	0
Open cases	11	22	26	21	30	15	13	30	20	33	18	34
Total cases considered	15	23	31	28	33	29	22	30	25	35	27	39

Cases closed following preliminary analysis

Perhaps the most important part of our learning in the first year has been in understanding how to analyse information at the first, preliminary analysis stage.

The most common reasons to close a case at this stage are:

The signal is “weak”: The panel decisions about the strength of the signal were based on a combination of additional statistical information (in addition to the original trigger). For example there may be little evidence that mortality is a current concern: the CUSUM plot may have been close to signalling for some time. If such cases are followed up with the trusts concerned, the likelihood of finding anything untoward needs to be balanced against the constituent effort.

The group of patients is vulnerable to inconsistent coding: We rely on clinical coding to identify groups of patients, whether by primary diagnosis or HRG. However, some types of patient are not consistently coded between different trusts, or coding practice can change within a trust over time. For example, non-ST elevated myocardial infarction can be hard to diagnose and can be coded in some trusts as unstable angina, while in others it is coded as a myocardial infarction. Some trusts may over-use codes relating to ‘unspecified’ types of condition that, if specified, would end up in a different group of codes. This means that, if we take such groups at face value, we are not comparing like with like. If the conditions are not life-threatening and no systemic problems within the trust can be identified, then these are good grounds for closing a case.

The box below shows a summary from an internal report on a signal related to raised mortality rates for the group “Spondylosis, intervertebral disc disorders, or other back problems”:

SUMMARY

- Alert raised by xxxxxx, signalling in October 2007
- HES data currently available only up until June 2007 – i.e. does not cover the period in which the signal occurred
- Not a funnel plot outlier in the periods July 2006 – June 2007 ($z=1.66$) or July 2005 – June 2007 ($z=1.65$)
- Not an outlier using CUSUM analysis for the period July 2005 – June 2007
- However, deaths are higher than expected in most of the quarters of HES available
- Almost all deaths are emergency admissions
- Raised mortality rates are among those aged 75+
- Deaths fall mainly under M54.9 – Dorsalgia, unspecified (*upper back pain*)
- Analysis of individual patient-level data shows that, although admitted with a spinal or musculoskeletal diagnosis, many of these patients were diagnosed with other significant problems by the discharge episode, and that the majority of those who died had one or more major co-morbidities recorded

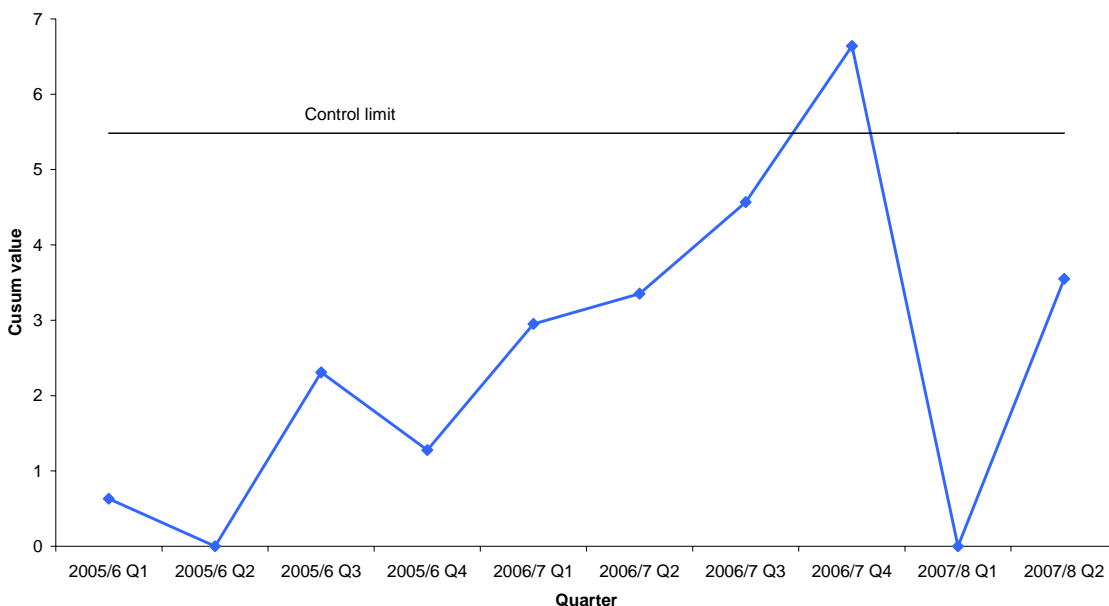
- This is not a group of patients that has high mortality rates associated with it, and it appears likely that the raised mortality among these patients may be as a result of co-existing problems

RECOMMENDATION

On the basis that this is a heterogeneous group of patients, with various additional serious problems unrelated to back pain, the classification of these patients is likely to be vulnerable to diagnostic variations between trusts and is therefore not a reliable group for comparison. We therefore do not recommend pursuing this case with the trust.

Issues with the case mix or organisation: These are cases where the internal analysis revealed some aspects of the case mix or the organisation – which was not considered in the risk adjustment – but was deemed to be a plausible explanation for any observed variance. The CUSUM plot and box below show a summary from an internal report where the decision to close was linked with insufficient risk adjustment:

Figure 4: HRG A03: Intracranial procedures except trauma - category 3. Emergency admissions. CUSUM plot



SUMMARY

- Compared with most acute trusts in England, a large proportion of patients within this HRG code are receiving operations following a brain haemorrhage.
- xxxxxxxx is a major centre for neurological care in the xxxxxxxx area.

- Nationally, emergency cerebrovascular patients have a higher in-hospital mortality rate than the other patients that are included within this HRG code.
- xxxxx has a high proportion of cerebrovascular admissions within this HRG code compared with most acute trusts in England.
- For patients in the wider group of cerebrovascular disease as defined by ICD10 codes, there are no concerns with standardised mortality.

RECOMMENDED ACTION

This alert appears to be an artefact of the mix of patients in this HRG group at xxxxx. Therefore, we recommend that this case is not pursued.

Cases pursued following preliminary analysis

Table 3 below shows the current status of the alerts, at the time of writing this report, which the Healthcare Commission decided to follow up.

- Seven alerts related to one trust where an investigation has since been undertaken, and will shortly be published.
- Twelve alerts have been closed following engagement with 11 trusts (one trust had two alerts), where they were able to provide sufficient information enabling the Commission to be satisfied for the reasons for the alert and any necessary corrective action. Further details and information about these alerts can be found in the next section.
- At the time of writing, ongoing engagement was being pursued with trusts in relation to the remaining 23 alerts.

Table 3: Current status of pursued cases

		<u>Closed</u>	<u>Referred to Investigations</u>	<u>Trust Engaged</u>
Pursue - suspect data artefact but require confirmation/ explanation	10	4		6
Pursue - unable to explain as an artefact	23	8	2	13
Pursue - trust already aware	1			1
Pursue - more general, other concerns/ multiple alerts	8		5	3
	42	12	7	23

Impact on the trusts where we followed up

As part of the evaluation of the programme, we wrote to the 11 trusts regarding the 12 alerts that have been followed up through the process (that is, to the point where we were satisfied that we could close our interest in the matters concerned). The trusts were asked for their agreement that we could include a summary of the specific issues relating to each trust with regard to the alert(s) that had been observed for them and any changes that had implemented as a result. The trusts were also asked three specific questions regarding the process, which are discussed further in the next section.

Table 4 summarises the information we have received about the responses in individual trusts that we followed up. In almost all cases the trust undertook some specific follow up or investigation in response to our queries. Most commonly these identified either problems in data recording or in diagnostic coding. Some organisation also identified some steps such as training that would be needed to reduce the chances of future similar problems.

Table 4: Summary of individual alerts that have been closed.

Trust A: *Cardiac Pacemaker System Introduced Through Vein*

The trust commissioned an analysis of cases as well as a review of five case notes under this HRG. This concluded that the deaths rates were higher than the national average because the Trust admits a much higher proportion of high-risk temporary pacing patients than nationally and refers permanent pacing patients to other trusts. This was verified by a reanalysis of their own data by the Dr Foster Unit at Imperial College. Current nationally available codes do not allow the trust to accurately record the actual activity undertaken which therefore does not enable comparable comparisons between trusts to be made.

Trust B: (2 alerts) *Pleurisy pneumothorax and pulmonary collapse*

The trust conducted two investigations which made the following recommendations:

Trust-wide

- Work to be undertaken with clinical teams to improve the recording of summaries for individual patients, and the coding team's application of clinical information to assignment codes.

Within cardiothoracic specialties

- The hospital should continue its governance process whereby deaths are reviewed by a consultant from a different specialty and presented at governance meetings.
- The thoracic medical team should consider establishing a system in which post-operative thoracic deaths can be reviewed at a joint medical/surgical thoracic meeting.
- A mechanism needs to be in place for accurately coding patients with pleural effusions reflecting their likely aetiology using clinical documents and data from the intensive care unit.

- To enable accurate coding, local clinical governance teams should monitor that medical teams are using a discharge tool to write discharge summaries on patients who die.

Trust C: *Coronary atherosclerosis and other heart disease*

In its response to the Commission, the trust stated that in the absence of a definitive diagnosis the coding department appears to default to the use of unstable Angina in those patients where a provisional diagnosis of Acute Coronary Syndrome had been written in the notes. They will address this through further review, clarification of diagnostic practice, recording of final diagnosis in the medical record and revising the current coding practice. It will also be looking at its current definition of myocardial infarction and introducing revisions in light of later evidence.

Trust D: *Aortic, peripheral and visceral artery aneurysms*

The trust admitted that there were clinical coding problems. In response to the specific mortality issues identified, it stated that patients allocated to one of the two vascular 'on-call' hospitals in the trust were wrongly identified as transfers into the trust and coded with admission source "81" as if they had been transferred from another trust. Many patients recorded as having died after being admitted as an emergency under this HRG had had significant medical co-morbidities, which on their own could be the cause of death. The decrease in elective admissions identified by the Commission could be explained by several factors including reduced number of critical care beds (increase usage of critical care beds by emergency cases), increased workload from other areas and seasonal variation. An Endovascular aortic aneurysm repair programme started in late 2007. In terms of the raised mortality rate for elective cases in 2006/07 this could be explained by patients being wrongly coded under HRG Q02 after being admitted with a primary diagnosis code of I713 and vice-versa.

Actions:

- There will be monthly audit of all abdominal aortic aneurysms recorded on the trust's patient administration system.
- The trust is in the process of creating a 'vascular registry', an in-house database of all major vascular activity within the Trust, which will provide a more robust source of data for future audit.

Trust E: *Septicaemia*

The trust commissioned a report which concluded that most anomalies that account for the outlier are due to recording practices, with co-morbidities being omitted and the coding of "septicaemia" being inappropriately applied.

Actions:

- The Medical Director has written to all consultants to stress the need for precision in clinical documentation, to ensure maximum accuracy of clinical coding
- The trust's guidelines on Septicaemia have been reviewed and revised by the trust's infection control team.

- All decisions to code a patient episode as “septicaemia” are monitored by the trust’s information department, and are subject to verification by the responsible consultant, on a continuous basis. The Trust’s Clinical Audit and Effectiveness Committee reviews the application of the code at its bi monthly meetings.
- Episodes are no longer coded as septicaemia unless there is a positive blood culture result in the patient’s notes.

Trust F: *Heart valve disorders*

The trust commissioned internal analysis and produced a report, which concluded that there was coding inconsistency with a lack of clarity over the recording of primary and secondary causes of death.

Actions:

- The cardiology team were to present relevant cases at directorate mortality review and audit days, which will look at issues around defining cause of death. They will also review heart valve disorder mortalities on an annual basis.
- The directorate will identify any training and support issues arising from this review.

Trust G: *Perinatal mortality*

The trust’s clinical midwifery manager informed the Commission that the trust’s data on perinatal mortality and stillbirths, which showed a much lower figure and is taken from the maternity unit’s computer system, CEMACH forms and NHS numbers for babies, did not match the information supplied by the Commission, as non-viable foetuses were being entered on the trust’s patient administration system, in addition to stillbirths. The trust also stated that there was an issue in the discrepancy between the trust’s and the PCT’s figures for perinatal mortality.

Regarding the under-reporting of caesarean, forceps and ventouse deliveries, the trust stated that this issue appeared to arise from an interface problem between the trust’s patient administration system and the maternity unit’s computer system.

Trust H: *Coronary atherosclerosis and other heart disease*

The trust’s response was based on local audit work undertaken into each case. The opening of a catheter laboratory in February 2006 at one of the trust’s sites accounted for the increase in activity for coronary angiograms and led to an increase in the diagnosis of atherosclerosis and more emergency referrals to the cardiologists with an associated primary diagnosis. It also accounted for the number of elective admissions to the trust with the laboratory conducting three sessions a week on behalf of another trust.

Additionally, an audit was conducted of 23 patients who were coded as having died from unstable angina, which although accurate was based on minor rises in cardiac enzymes (TNI) without other features of unstable angina being

evident. This conclusion is supported by evidence that the cause of death in more than 50% of the cases was related to another diagnosis that was more relevant.

Actions:

- A review of clinical coding of primary diagnosis in this area has been initiated to establish whether the clinical coding should have been different. Following this review any relevant issues will be communicated across the organisation.
- Communication between the clinical coders and the clinicians needs to be improved to ensure the robustness of the clinical coding.
- The trust will investigate ways of improving the clinical coding so that it accurately reflects the primary cause of death.

Trust I: *Fracture of neck of femur*

The trust replied that it was aware of the outlier and was conducting an internal enquiry. This enquiry concluded that there appeared to be no common clinical cause for the deaths, and that the case mix contained a large number of patients with a high number of contributing co-morbidities, which may have been underestimated by the coding team.

Actions:

- Audit discharge summary completion
- Ensure that the fractured neck of the femur patient pathway is applied to all patients regardless of which ward they are situated on and review the emergency care pathway for fracture of neck of femur
- Continue to monitor and audit fracture of neck of femur information, including gender mix.
- Consider revising the coding process for orthopaedic patients.

Trust J: *Coronary atherosclerosis and other heart disease*

In a letter answering our questions about aspects of this outlier, the trust said that it would be making a range of changes to the way they undertake coding including ensuring that coders have access to all medical notes of the patients before the data submission deadlines. They also stated that medical staff will be reminded of the importance of completing death certificates accurately and that they will be investing in their coding team.

Trust K: *Intracranial injury*

The trust's associate medical director directed a case note review of cases, which concluded that there were no issues arising from these cases and that the senior clinicians dealt with them appropriately and expeditiously.

In every case, we note details of the actions taken by each trust and pass relevant information on to the Commission's regional staff. This would be crucially important should we identify any recurrence of the original alert (such a situation has not arisen as yet).

Feedback from the trusts involved

As part of our evaluation, we sent a postal questionnaire to the 11 trusts where alerts had been observed and pursued, and the cases were subsequently closed because we had no further concerns. These 11 trusts were asked the following questions:

1. With regards to our contact with you, was it useful to you in terms of identifying matters that you had not previously been aware of?

Five trusts replied that they thought our contact was useful in identifying matters that the trust had not previously been aware of. One trust did not answer this question and the remaining five said that it was not useful in these terms (in part because they were already aware of the issue).

2. Has it led to any changes other than those that you may have already informed us about? If it did, can you explain briefly what they are?

Six trusts said that our contact had led to changes; these include improvements to data quality and related training, and new clinical forums to promote clinical debate. Five trusts said that our contact had not led to any *further* changes to those that they had already told us about (see table 4 above).

3. Do you think that the Healthcare Commission (and the forthcoming Care Quality Commission) should continue its work of identifying and considering providers of healthcare services where data suggests there might be a performance issue?

Almost all the trusts, ten, replied that our work should continue. One of the trusts, which did agree that the work should continue, did advocate for more coordination between the regulators and organisations involved in generating and communicating alerts with trusts. The single trust that did not agree that the work should continue thought that, as trusts now have systems to assess their performance, they should evaluate such matters first, themselves, before involving the Commission.

5. Conclusions

We have established a new programme of work to identify and follow up concerns about apparently high rates of mortality in the NHS. Over the 12 month period considered by this report, we have considered 85 alerts, relating to 56 acute trusts (one out of three acute trusts in England).

After analysing the data, we concluded that 42 of these alerts (relating to 29 – or 17% – of acute trusts in England), needed to be pursued by the Commission. We followed up where there was no clear explanation of the apparently high mortality rates, to obtain further information, or explanation, from the trust so that we could reach a judgement about the need for further action. The level of our concerns has varied across these cases, and this has guided our approach. Such was our concern with one trust – which generated several different alerts – that we launched a formal investigation. We will be publishing our findings from that investigation shortly.

We have started this work cautiously, wishing to avoid creating unnecessary alarm and recognising that we ourselves are still learning about the best approach to take in this important but sensitive area. Overall, we have seen a very positive response to our enquiries from trusts, who have generally followed up concerns quickly and thoroughly. Rather than being an additional burden, the trusts themselves have indicated overwhelmingly that they see value in the regulator continuing this work. We have also had very positive feedback from other national bodies that we have worked with, and confirmation that the Healthcare Commission is breaking new ground, internationally, in this area.

There is clear evidence that our follow-up action has led directly to improvements in clinical practice, and in the use of clinical data in NHS trusts. We expect this to have led to direct improvements in the care provided to patients, but acknowledge that further evaluation is necessary to understand the extent to which this may have happened.

Analysis of the cases considered in this report shows that there are clear variations across the NHS in the quality of coding clinical outcomes, and variations in the extent to which statistical information is used to monitor the quality of local services and inform decisions at a senior level within NHS trusts. This is of particular concern in a modern, information-driven health service where the interpretation and use of data is a fundamental means of improving clinical care.

One important aspect of this work concerns the extent to which this information should be routinely published. We believe it wholly appropriate in principle that information about mortality rates should be made more explicitly available to the public. However, we recognise the challenges in ensuring that data on mortality rates are accurate and informed, and that they need to be expressed in such a way that does not cause unnecessary alarm amongst patients nor lead to unhelpfully risk averse behaviour amongst clinicians. These challenges, whilst they are significant, should not prevent us from working with others to make information on mortality rates more widely available, building on the experience

of our joint approach with the Society for Cardiothoracic Surgery, to make such information available via the Commission's website.

Our work has been focused on specific aspects of care and on specific groups of patients. It has required detailed, often complex analysis before we can understand what the data are telling us, and whether we need to take further action. Our decisions have not been based on a single quantitative indicator but on a range of different pieces of information and intelligence. The outliers themselves are questions to prompt further enquiry – not judgements on the quality of care. With this in mind, although the publication of 'high level' hospital-wide mortality rates is superficially appealing, we do not believe that this is the best way of providing meaningful information to the public, in order to help them to make informed choices about the quality of clinical care.

Instead, we recommend that the detailed mortality data already collected by NHS trusts should be used more explicitly to prompt questions and direct the collection of intelligence that forms the overall assessment of each trust's performance. Where there is confidence that data can be used for meaningful comparison, it should be made publicly available, along with other aspects of the overall assessment. This would provide useful information to patients which, when placed in its proper context, would help to inform their choices. It would also focus the attention of NHS leaders, especially those in trusts, on the importance of ensuring that their data are accurate and used to improve the quality of care to patients.

We believe that our work to date has progressed well, is wholly consonant with current national policy on regulation, and that there is good reason for it to continue. We have, however, merely scratched the surface of how data of this kind can be used intelligently by an information-led, risk-based regulator. Our progress over 12 months in this limited, focused, programme of work suggests that there is great potential in further developing its application to different care sectors. We have already begun to look at incident data from independent health care, in order to identify and follow up outliers, in a similar way that we have been doing in the NHS.

Looking to the future we need to consider applications in primary care and social care settings, as well as refining our approach in relation to, for example, mental health services. So far, we have deliberately focused our attention on mortality rates but there is a wealth of other data, already available, which could provide a more sophisticated view of where potential problems may exist, but may be as yet undetected. For example, we might look in detail at re-admission rates, excessive length of stay, infection rates or other data, which is specific to different specialties or care sectors.

6. Recommendations

We highlight four particular recommendations arising from this report:

Recommendation 1: Based on the progress made so far, and the support from trusts themselves, this programme should continue in relation to NHS acute services and further consideration should be given to developing its application to different care sectors.

Recommendation 2: The programme should be extended beyond the consideration of mortality rates, to include other potentially serious outliers derived from data on re-admission rates, excessive length of stay, infection rates and other data specific to different specialties or care sectors.

Recommendation 3: The detailed mortality data already collected by NHS trusts should be used more explicitly to prompt questions and direct the collection of intelligence that forms the overall assessment of each trust's performance. Where there is confidence that data can be used for meaningful comparison, it should be made publicly available, along with other aspects of the overall assessment of NHS trusts.

Recommendation 4: This report should be widely disseminated across the NHS to highlight the importance to trusts of ensuring that their data are as accurate, complete and up-to-date as possible.

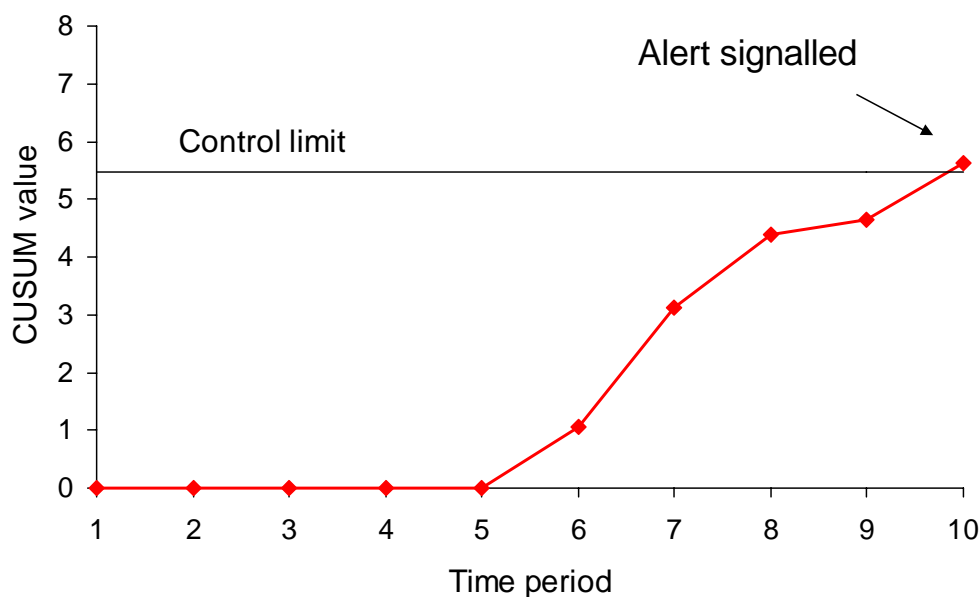
Appendix: Identifying outliers using cumulative sum (CUSUM) methodology

The CUSUM

The CUSUM is a statistical process control (SPC) technique that is used to detect persistent deviations from a reference value. Like most SPC methods, it was developed in manufacturing industry, but is now widely used in health outcomes monitoring.

The CUSUM is a sequential hypothesis testing technique by which evidence in favour of outcomes occurring at the expected rate (the null hypothesis, H_0) is continually weighed up against evidence that a change has occurred (the alternative hypothesis, H_1).¹ CUSUM control charts are plots of the cumulative log likelihood ratio between these two hypotheses. They are also constrained not to fall below zero. So, for example, if a CUSUM is designed to detect series of outcomes that are worse than expected, it cannot build up credit for series of good outcomes. If the CUSUM exceeds a predefined threshold, or control limit, then the hypothesis of a change (H_1) is accepted in favour of the null (H_0) and this constitutes an 'alert' or 'signal'. After each alert, the CUSUM is reset to zero so that if any changes subsequently occur there is time for them to take effect. Alternatively, if poor outcomes persist then a further signal is likely to occur at a later time. Control limits have to be set to guard against too many 'false alarms' occurring as a result of random variation, but not be set at too a high a value that it becomes very difficult to detect any differences in mortality. A CUSUM is illustrated in Figure A1.

Figure A1: CUSUM control chart



Mathematically, if C_t denotes the CUSUM value and w_t the log likelihood ratio at time t then:

$$C_0 = 0$$
$$C_t = \max\{C_{t-1} + w_t, 0\}$$

The values w_t are called the CUSUM weights.

Indirect standardisation

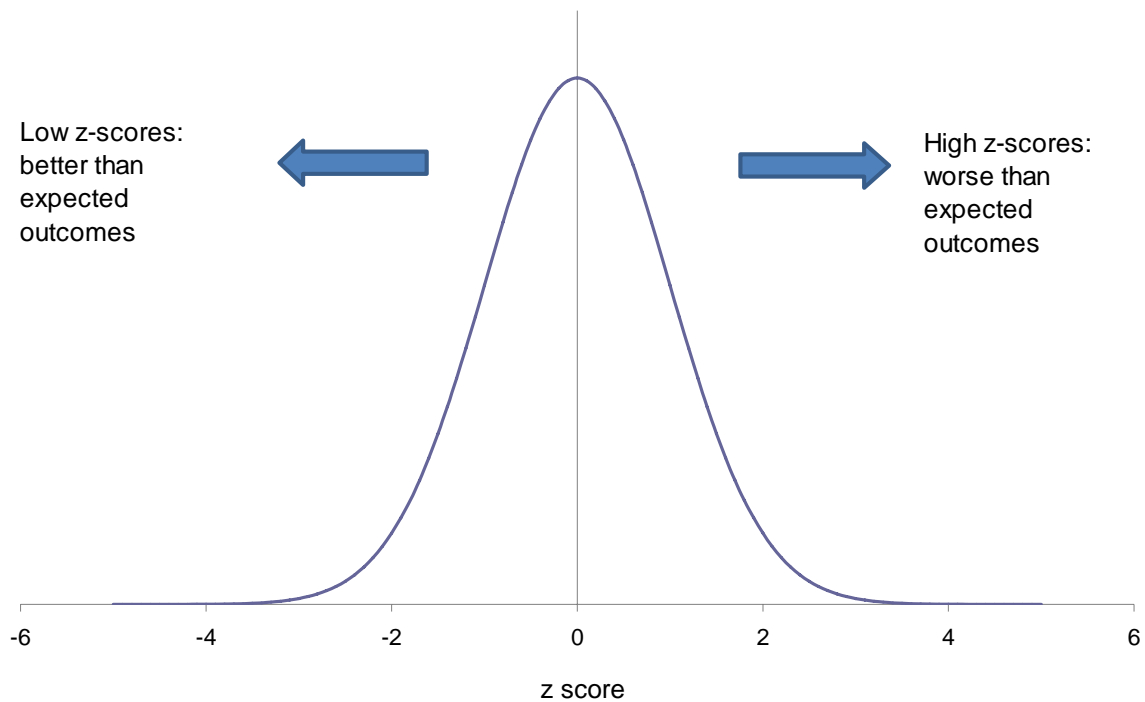
The mortality data we monitor is indirectly standardised so that a trust's outcomes are compared against an expected value that allows for differences in patient ages and gender.² We further standardise by time period (calendar quarter) so that comparisons are always local to each period and effects of seasonality and national trends are avoided.

Calculating z-scores

Outcomes are converted to z-scores so that we have a common scale for analysis across the range of measures we use. The z-score is equivalent to a standard normal variate that measures the number of standard deviations away from a reference value, preceded by a plus or minus depending on whether it is respectively above or below the reference value. High z-scores indicate worse

outcomes and low z-scores better outcomes. Z-scores correspond to p-values in that a one-tailed p-value of 0.01 is equal to a z-score of 2.3 and a p-value of 0.001 approximately matches a score of 3.0. In our CUSUM analysis we test for series of persistently high z-scores.

Figure A2: z-scores and their relationship to the normal distribution



If O represents observed mortality for patients assigned a particular HRG within a trust, and E the expected value we calculate z-scores from the square root of the standardised mortality ratio, $\sqrt{(O/E)}$. If we also assume a Poisson distribution of outcomes, this ratio has expected value of one and a standard deviation approximately equal to $1/(2\sqrt{E})$. We can therefore derive an approximate z-score for this measure by the formula:

$$z = 2(\sqrt{O} - \sqrt{E})$$

The square root transformation is used to improve normality and has an added advantage of stabilising the variance (being inversely proportional to the expected value).

Statistical model

These z-scores are very likely to be over-dispersed, i.e. their true variances are greater than one, which may be a consequence of insufficient benchmarking or the presence of common-cause factors that render the Poisson model inadequate.³ To allow for over-dispersion the z-score distribution is modelled with a hierarchical structure: if z_{kt} represents the z-score for trust k at time t ,

$$\begin{aligned}z_{kt}|\theta_k &\sim N(\theta_k, \sigma^2) \\ \theta_k &\sim N(0, \tau^2)\end{aligned}$$

i.e. a trust's z-scores are normally distributed about a local mean for that trust, θ_k , with standard deviation, σ . These trust mean values are themselves normally distributed about zero with standard deviation, τ .

Estimating variances

Estimates for the local means θ_k are derived from the z-scores over a specified time period ($t = 1$ to T , say):

$$\hat{\theta}_k = \sum_{t=1}^T z_{kt} / T$$

To estimate the variances we Winsorise the data by shrinking the most extreme z-scores and using the formulae:

$$\begin{aligned}\sigma^2 &= \frac{\sum_{k=1}^N \sum_{t=1}^T (z_{kt}^{(w)} - \hat{\theta}_k)^2}{N(T-1)} \\ \tau^2 &= \frac{\sum_{t=1}^T \sum_{k=1}^N (z_{kt}^{(w)} - \bar{z}_t)^2}{(N-1)T} - \sigma^2\end{aligned}$$

where N denotes the number of trusts, $z_{kt}^{(w)}$ are the Winsorised z-scores and \bar{z}_t is the mean Winsorised z-score over period t .

Hypothesis tests

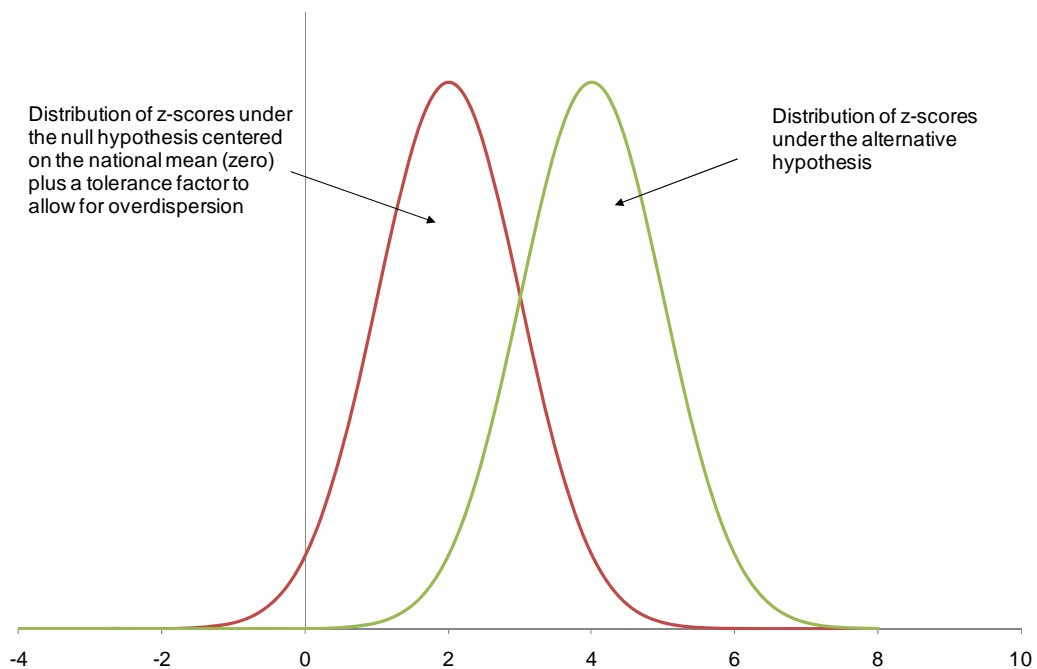
To set a null hypothesis we set a value for the local mean that is in the upper part of its probability distribution, i.e.

$$H_0 : \theta_k = \gamma_1 \tau$$

where γ_1 can be interpreted as a tolerance factor for the mean. We are thus allowing the expected value for a trust to be greater than zero (see Figure A3). This is tested against the alternative hypothesis:

$$H_1 : \theta_k = \gamma_1 \tau + \gamma_2 \sigma$$

Figure A3: Assumed distribution of z-scores for testing higher than expected mortality



We can transform our z-scores into standard normal variates under the null hypothesis by calculating:

$$z_{kt}^* = \frac{z_{kt} - \gamma_1 \tau}{\sigma}$$

and our hypothesis test becomes:

$$H_0 : \theta_k^* = 0$$

$$H_1 : \theta_k^* = \delta$$

where $\delta = \gamma_2$

The CUSUM for standard normal data

With standard normal data and a hypothesis test as above, the CUSUM weights have values:¹

$$w_t = \delta z_{kt}^* - \frac{\delta^2}{2}$$

Deriving control limits and other stopping rules from steady state p-values

For such a CUSUM it is possible to estimate steady-state p-values:⁴

$$p = 1, \quad \text{if } x = 0$$

$$= \lambda_1 e^{-x} - \lambda_2 e^{-\kappa x}, \quad \text{if } 0 < x \leq x'$$

$$= \gamma e^{-x}, \quad \text{if } x > x'$$

where:

$$\lambda_1 = e^{-0.735(\delta-0.139)} + 1.098(0.073\delta + 0.031) - 0.074$$

$$\lambda_2 = 0.073\delta + 0.031$$

$$\kappa = e^{-1.30(\delta-1.31)} + 1.04$$

$$\gamma = 0.986(0.56^\delta) + 0.008\delta$$

$$x' = 0.170\delta^2 + 1.053\delta - 0.02$$

Knowing the p-value is necessary for setting a stopping rule based on the false discovery rate (FDR)⁵ and also enables constant limits to be set that correspond to pre-specified tails of the null distribution of the CUSUM. We typically use the

limit that corresponds to the upper 0.1% tail of the adjusted z-scores, z_{kt}^* (a CUSUM z-score equal to 3).

Scanning HES mortality data using the CUSUM

We have developed a computer program that calculates CUSUMs within each NHS acute trust by individual HRG and admission method (emergency and elective). There are 610 HRGs (version 3.5) and approximately 170 NHS acute trusts, therefore, since we also consider two methods of admission, the program is calculating approximately 200,000 CUSUMs. Any signals that occur are flagged and with our chosen threshold we are generating approximately 30 new signals each quarter.

References

1. Hawkins D, Olwell D. Cumulative sum control charts and charting for quality improvement. Springer: New York, 1997.
2. Armitage P, Berry G, Matthews JNS. Statistical methods in medical research. 4th edition. Blackwell: Oxford, 2002. pp. 663-667.
3. Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Stat Med* 2005;24:1185-1202.
4. Grigg OA, Spiegelhalter DJ. The null steady-state distribution of the CUSUM statistic. Technical report. Medical Research Council Biostatistics Unit, Cambridge. http://www.mrc-su.cam.ac.uk/BSUsite/Publications/Preprints/cuspval_ogds.pdf.
5. Jones HE, Ohlssen DI, Spiegelhalter DJ. Use of the false discovery rate when comparing multiple healthcare providers. *J Clin Epidemiol* 2008;61:232-240.