A practical method for monitoring general practice mortality in the UK:
findings from a pilot study in a health board of Northern Ireland

Mohammed A Mohammed, Kathryn Booth, David Marshall, Máire Brolly, Tom Marshall, Kar-Keung Cheng, Martin Hayes and Sandy Fitzpatrick

ABSTRACT

Background
The Baker report into Dr Harold Shipman's murders recommended monitoring mortality in general practice, but there is currently no practical method available to implement this.

Aim
To monitor mortality rates in response to the Baker report and to use the data to improve quality of care.

Design of study
Prospective mortality monitoring study.

Setting
Eastern Health and Social Services Board, Northern Ireland.

Method
Linked quarterly mortality data from 1994–2001 were compiled for 114 general practices in Eastern Health and Social Services Board in Northern Ireland. Cross-sectional control charts compared crude and adjusted mortality rates across all the practices. Longitudinal control charts analysed quarterly mortality rates over 28 quarters within each practice. Practices were sent their own control charts and invited to feedback workshops. Special cause variation in mortality was investigated as follows: checks on data, case-mix, practice structures, processes of care and finally individual carers.

Results
Age, sex and deprivation adjusted cross-sectional control charts identified 18 practices as showing special cause variation in their mortality (11 high and 7 low). Assignable causes were found for all high special cause practices: large numbers of nursing home patients (six practices), very high levels of deprivation and high morbidity not captured by our case-mix adjustment (five practices). For three of seven low special cause practices, case-mix adjustment underestimated affluence and overestimated morbidity levels. Feedback indicated widespread support for the principle of monitoring, but concerns about the public disclosure of mortality data.

Conclusions
We have successfully developed and piloted a general practice mortality monitoring system with the support and participation of local stakeholders. This used control charts for analysis and followed a scientific strategy for investigating special cause variation.

Keywords
control charts; general practice; mortality; monitoring.

INTRODUCTION

Following Dr Harold Shipman's murder of at least 215 of his own patients, the Baker report recommended routine monitoring of general practice mortality rates.[1] This recommendation is challenging for a number of reasons.[2] Routine mortality monitoring needs high quality mortality data linked to general practices.[3] There is debate about how easy it is to distinguish unusual variation in general practice mortality from chance variation.[4,5] Finally, it is not clear what action should be taken in the event of an unusual variation in mortality.

The Baker report did not recommend a statistical method for monitoring. Standard statistical tests such as χ^2 test or *t*-tests might be used or more sophisticated approaches such as multilevel modelling.[6] Control charts based on Shewhart's theory of variation are another alternative.[7,8] This theory classifies variation into two categories according to the action required to reduce it. Common cause variation is unlikely to have an

MA Mohammed, PhD, senior research fellow; **T Marshall**, MSc, PhD, MRGP, MFPHM, senior lecturer; **KK Cheng**, PhD, MRCGP, FFPHM, professor of epidemiology, Department of Public Health & Epidemiology, University of Birmingham. **K Booth**, MB ChB, BAO, GP medical adviser; **M Hayes**, MSc, GMS contracts manager, Eastern Health & Social Services Board, Belfast. **D Marshall**, BSc, PhD, principal statistician, demography and methodology branch; **M Brolly**, BSc, deputy principal statistician, Northern Ireland Longitudinal Study, Northern Ireland Statistics and Research Agency, Belfast. **S Fitzpatrick**, BSc, principal statistician, Central Services Agency, Belfast.

Address for correspondence
Kathryn Booth, Eastern Health and Social Services Board, 12–22 Linenhall Street, Belfast BT2 8BS.
E-mail: KBooth@ehssb.n-i.nhs.uk

assignable cause. It is expected in any process because it is intrinsic to every process and it affects all in that process. To reduce common cause variation requires action on the underlying process. Special cause variation is likely to have an assignable cause. It does not affect all in that process. Special cause variation needs to be investigated to identify the assignable cause and appropriate action taken. If it results in less favourable outcomes, appropriate action may be to remove the assignable cause. If it results in more favourable outcomes knowledge of the assignable cause is used to improve the process as a whole.

Control charts distinguish between common cause and special cause variation. They have three lines, a central line (the average) and upper and lower control limits set at 3σ from the central line.[9] Data points appearing outside the control limits, or certain unusual patterns indicate special cause variation.[9]

Shewhart control charts have proved powerful at identifying special cause variation in high profile cases such as Bristol and Shipman.[10] They have been used for analysis of variation in clinical outcomes by different service providers.[11,12] They are also less likely to lead to over-investigation than traditional analyses.[13] We sought to explore the feasibility of routine monitoring of death rates in general practice using a system designed around control charts. The immediate requirement was to meet the recommendations of the Shipman Inquiry.[14] However, monitoring and systematic investigation of special causes, linked to an educational process provides a potentially useful model for a general approach to quality improvement in primary care. We report findings from a 2-year pilot project (February 2002–April 2004) based in Northern Ireland's Eastern Health and Social Services Board.

METHOD

Northern Ireland Mortality Data

The population denominator is a list — held by the Central Services Agency (CSA) — of all patients registered with a GP in Northern Ireland: the Central Health Index (CHI). The numerator is the number of deaths. These are notified to the CSA on a weekly basis, mainly from the General Register Office (GRO) and routinely linked back to the CHI list. Nearly all (approximately 98%) GRO records are traceable. We used the CHI to determine the number of deceased and live patients for each general practice.

Additional information on each patient included age, sex and a Jarman deprivation score. The deprivation score was calculated for each enumeration district from 1991 Census data[15] and assigned to each patient through their residential postcode. A Jarman deprivation profile was also

How this fits in

Monitoring mortality rates in general practice was recommended in the Baker report. To date no practical method of mortality monitoring has been implemented. Monitoring of mortality rates in general practice is feasible and practical. Implementation of a monitoring system requires participation and education of stakeholders; a practical tool to distinguish signals arising from unusual variation in mortality rates from noise; and an agreed process for follow-up of signals. In this study unusual mortality rates were the result of unmeasured case-mix factors (demography and deprivation) affecting practices. Stakeholders identified concerns about the public disclosure of mortality data and the overall aim of monitoring.

assigned to each general practice.[16] Detailed data on all deaths based on the death certificate, such as cause and place of death, have been compiled at practice level by linking the GRO deaths data with the CSA mortality data. This data can assist in the determination of patient case-mix.

General practice population sizes were available on a quarterly basis via the General Medical Services (GMS) payment system. This, along with the number of deaths of patients registered with a general practice, enabled historical mortality data, in the form of quarterly rates, to be determined. No detailed historical information on age, sex and deprivation was available.

Analysis

The Eastern Health and Social Services board includes 147 general practices. Practices undergoing mergers and splits were excluded and we analysed mortality rates from 114 anonymised general practices, which cover 667 000 of Northern Ireland's population of approximately 1.7 million.

We produced two types of cross-sectional control charts, one comparing crude mortality rates for all 5 years (1996–2000) from each of the 114 practices to the Northern Ireland mortality rate. The other compared age, sex and deprivation adjusted mortality rates. The control limits for this analysis are based on the Northern Ireland data. Practice list sizes were omitted from these two control charts to protect individual general practices' anonymity.

We produced longitudinal (time–series) control charts for each individual general practice. These showed the difference between the Northern Ireland quarterly mortality per 1000 population and the crude practice quarterly mortality per 1000 practice population over 28 quarters, (third quarter of 1994 through to the second quarter of 2001). These control charts are XmR (moving range) charts, with control limits for each practice derived from the difference between practice quarterly mortality rates and the Northern Ireland quarterly mortality rate.

Longitudinal control charts were used in the process of investigating special cause variation identified in the cross-sectional charts. For illustrative purposes only, control charts from four practices — A (low mortality), B (high mortality), C (low mortality) and D (high mortality) — are included. These indicate the kind of patterns seen in these charts.

Participation and education

We set up eight educational/feedback workshops for GPs. These were intended to educate participants in the underlying principles, tools and techniques of this pilot project. Invitations were also sent to patient representatives and coroners. The workshops were designed to explain the project, the use of control charts, to invite participants to raise concerns and to influence the design and development of the project. Prior to the workshops each general practice was sent their own mortality data and two control charts — a risk-adjusted cross-sectional chart highlighting their own practice and their own longitudinal chart. We report specific comments on these workshops in the results.

Systematic investigation of special cause variation

Practices showing special cause variation are likely to have an assignable cause for that variation. We sought to identify 'credible' assignable causes of variation, using a systematic strategy adapted from industry.[17] This strategy is represented as a pyramid (Figure 1). When we see special cause variation, we first check the data, then patient case-mix, then practice structure and resources, then the processes of care being used by the practice and finally individual carers.

Our search for a credible assignable cause had two phases. In an exploratory desktop phase we used existing data to identify possible causes of special cause variation. Subject to GP consent, we then discussed mortality data during an on-site meeting with GPs in their practices.

All 114 practices were offered an on-site meeting to discuss their mortality data in further detail. Practices with special cause variation on the age–sex–deprivation adjusted cross-sectional control chart were prioritised for practice on-site meetings, since this was the focus of the analysis. At each on-site meeting, the practice's own mortality data — including a list of all deceased patients — and their control charts with one or more GPs was discussed. They were invited to question the integrity of data and report any data verification they may have undertaken. Only when practices acknowledged that our data were correct, did we proceed to the next step.

Using a semi-structured interview tool and aided by longitudinal control charts for each of three age bands (>75 year olds, 65–75 year olds, <65 years), GPs were asked to suggest possible explanations for their special cause variation on the adjusted cross-sectional control chart. These hypotheses were recorded and it was investigated whether these hypotheses concurred with those generated by the desktop phase or whether any additional analyses could be undertaken in support or otherwise of these hypotheses. If a GPs' suggested explanation appeared to be credible (face-validity) and there was supporting evidence from our independent desktop analysis (using additional data sets where applicable), we concluded that a credible assignable cause had been found. Where there was no such agreement (despite further desktop analyses) we concluded that a credible assignable cause could not be found.

Two additional data sets were used in testing possible hypotheses. The first was the Noble index:[18] a multiple deprivation index, derived in 2001 and based largely on 1999 administrative data. The Jarman index is a measure of GP workload than of social deprivation and was based on the 1991 census rather than more recent data. The Noble index therefore replaced the Jarman index in a sensitivity analysis of case-mix adjustment. The second was the prevalence of 'Limiting long-term illness'; derived from the 2001 census.[15] It is a self-assessment of limiting long-term illness health problems or disability limiting daily activities or work and it includes problems due to old age.

Figure 1. 'Pyramid' model of investigation to find a credible cause for high mortality.
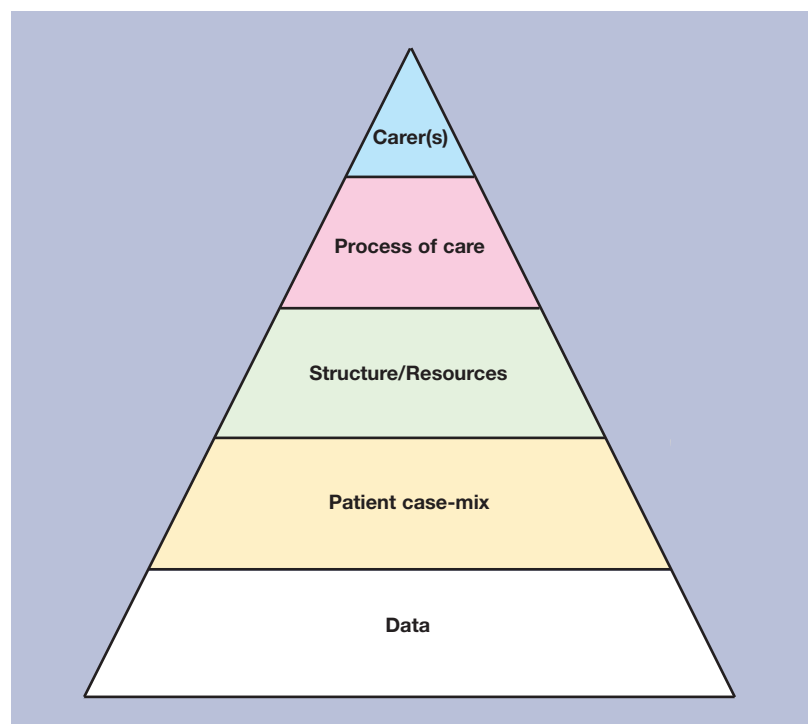
Where appropriate, we report the practices' rank in each case (Noble index and Limiting long-term illness) based on 566 wards (Noble quintiles) and 582 wards (Limiting long-term illness quintiles). Under the Noble index, the highest ranked wards (quintile 1) are the most deprived.

RESULTS

Cross-sectional control charts

Almost half of the general practices show special cause variation on the control chart of crude mortality rates from 1996–2000. When mortality rates are adjusted for age, sex and deprivation, 18 practices show special cause variation. (Figure 2) The practice (code D) furthest away from the upper control limit is used to illustrate the systematic investigation of special cause variation using the pyramid model.

Longitudinal control charts

Longitudinal control charts — of the quarterly difference between the practice crude and Northern Ireland crude mortality rate – are illustrated for four general practices.

The mortality rate of general practice A was below the Northern Ireland mortality rate in all 28 quarters and is consistent with common cause variation (Figure 3). Its mortality rate is within the control limits on the age–sex–deprivation adjusted control chart. (Figure 2).

The mortality rate of general practice B shows special cause variation in quarter two (Figure 4). The signal from this quarter is so loud that it is likely to be masking another signal in quarter 10. This lies outside the upper control limit if quarter two is excluded from the calculations. Furthermore, there is evidence of a step down in the mortality rate at quarter 11. General practice B appears within the control limits in Figure 2.

General practice C has been consistently below the Northern Ireland mortality rate, but there is evidence of a step-up in mortality in quarter 20 (Figure 5). General practice C is below the lower control limit in the control chart of crude mortality rates, but is within the control limits in Figure 2.

General practice D appears well above the upper control limit in the control chart of crude mortality rates and remains so after adjustment (Figure 2). A longitudinal control chart for this practice shows that mortality is decreasing (Figure 6) . Preliminary follow-up discussions with the practice highlighted the high proportion of nursing home patients on their list. This proportion subsequently declined as a result of a change in practice policy designed to reduce the numbers of nursing home patients on their list. A subsequent analysis of the crude death rates in those aged over 75 years old confirms this trend (Figure 7).
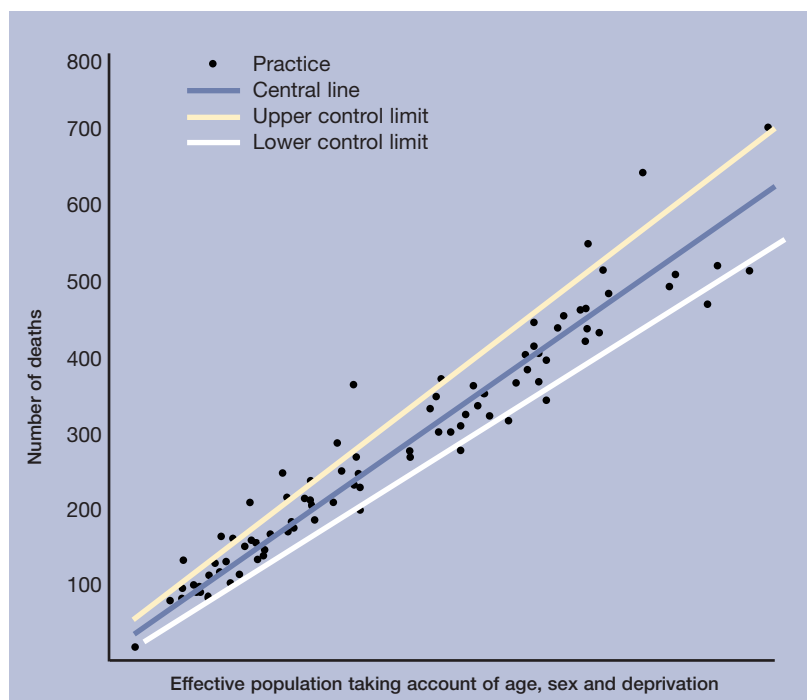

Figure 2. Cross-sectional control chart showing age–sex-deprivation adjusted mortality rates for 114 general practices in the Eastern Health and Social Services Board, Northern Ireland (1996–2000).
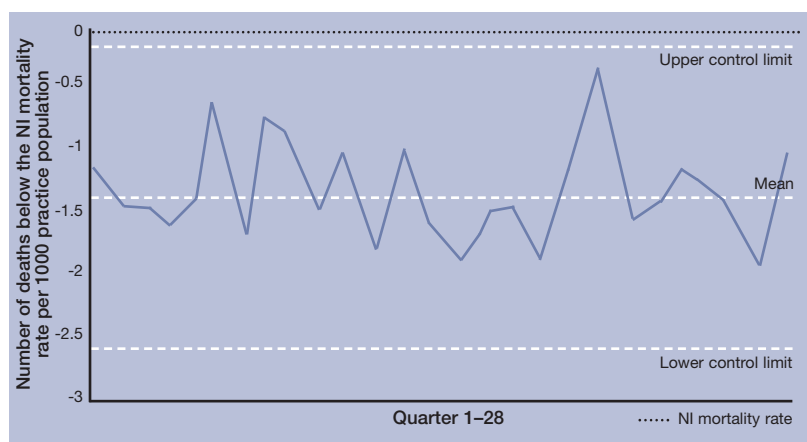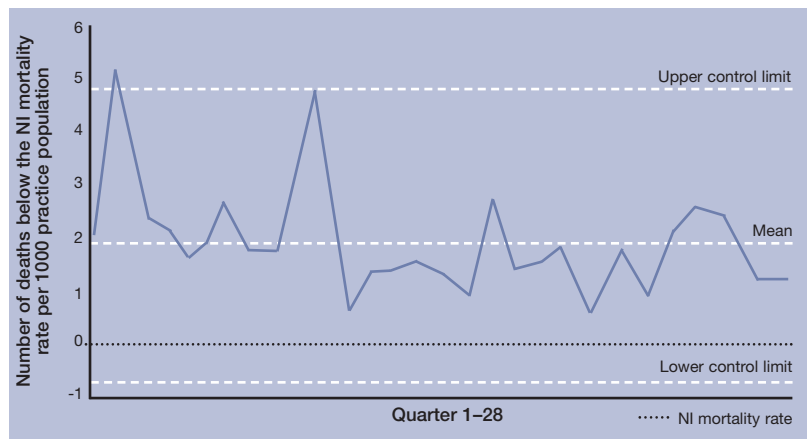

Figure 3. Longitudinal control chart of general practice A.

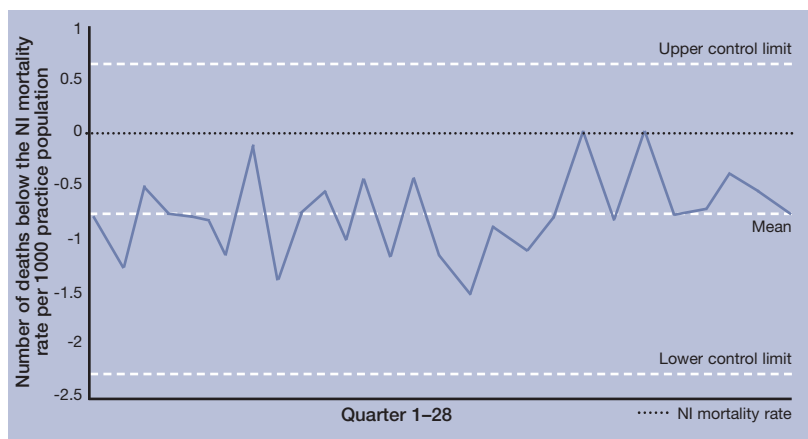*Figure 4. Longitudinal control chart of general practice B.*

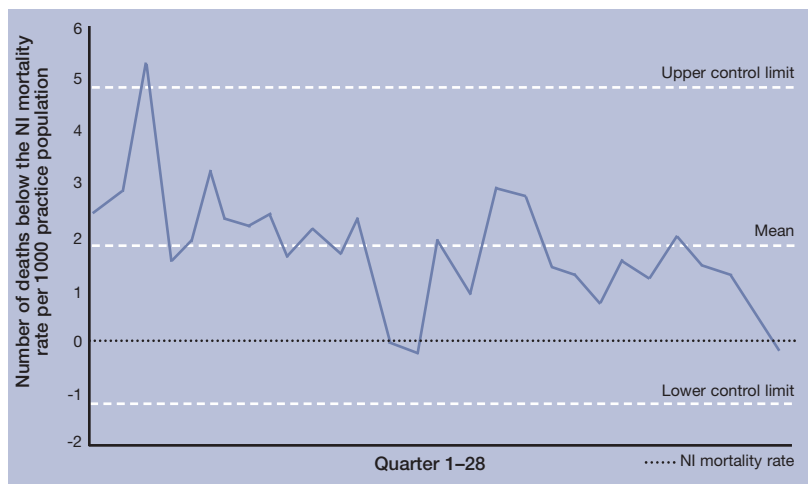**Figure 5. Longitudinal control chart of general practice C.**

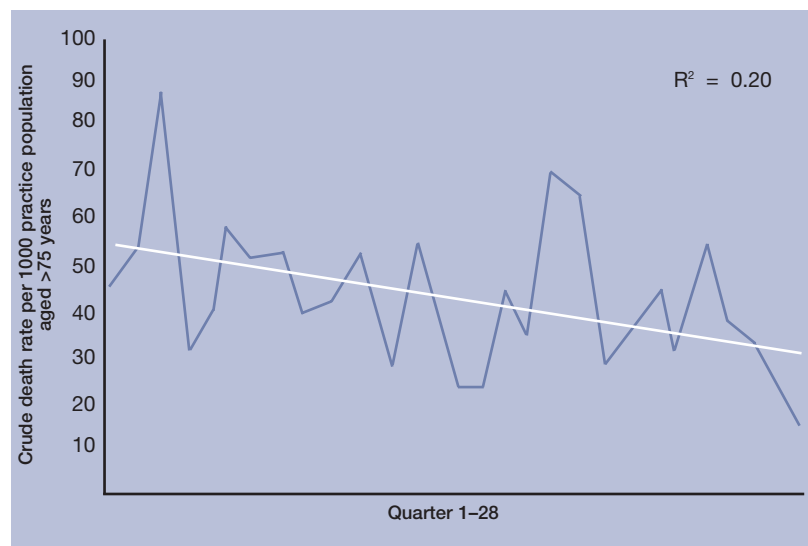**Figure 6. Longitudinal control chart of general practice D.**

Figure 7. Crude death rate in patients aged >75 years at practice D with a least squares line.

GP feedback

Feedback from the workshops identified a number of themes. Participants expressed anxiety about scrutiny of death rates and the need to protect clinicians from the adverse consequences of monitoring. There was a fear that high mortality rates would be interpreted as malpractice unless proven otherwise. Some GPs expressed relief that their practice data was within control limits. Participants also expressed concern about the public disclosure of such data.

Participants expressed general agreement on our use of the 'pyramid' model of investigation. A number of GPs requested a breakdown of their practice mortality data in order to validate it against their own practice database. Several participants requested additional assistance in interpreting their own data, perhaps with additional analyses from our existing data set. A recurrent closing theme in all our workshops was the simple question — what is the aim?

Explanations for special cause practices

In all, 44 site visits were undertaken, including 14 of 18 special cause practices. Four low special cause practices did not consent to a visit and we made no further attempts to identify assignable causes for their low mortality rates. All other practices were given access to their mortality file (list of their deceased patients) prior to the site visit. All 14 practices judged the data to be reliable as compared with their own records and memories. Furthermore, all 44 practices we visited (30 showing only common-cause variation) found the mortality data useful and requested that such data be made available to them annually.

Table 1 shows the most likely assignable causes identified for each practice together with supporting evidence from key data items. Six of the 11 practices with high special cause variation had an unusually high proportion of patients dying in nursing homes. Adjustment by place of death had not been considered prior to the analysis.

In the three low special cause practices the Jarman index appeared to underestimate affluence. Re-analysis using the Noble index supported this hypothesis. For one low special cause practice, affluence was combined with very low levels of self-reported limiting illnesses.

DISCUSSION
Summary of main findings

Our monitoring system distinguished special cause variation in both cross-sectional and longitudinal mortality rates from background chance variation. Special cause variation invites us to engage in the scientific method.[19] Our pyramid model provided a scientific method of investigation in partnership with the practice. This begins with the hypothesis that there is an assignable cause and systematically attempts to identify it. A check on the data makes improvement in data quality integral to, and not a prerequisite of, the

monitoring process. In all practices that we visited, our data were found to be accurate and reliable. A check on case-mix identifies whether some practices serve unusually high-risk or low-risk populations.

In our study both high and low special cause variation could be credibly assigned to unmeasured case-mix factors: a large nursing home effect, unmeasured excess or low levels of morbidity, unusually high or low levels of deprivation. Since practice age, sex and deprivation profiles provided sufficient assignable causes for mortality variations, we did not need to proceed higher in our pyramid model. If necessary, further investigation would focus on the levels of practice resources. We would then investigate the process of care such as differences in implementation or practice organisation. Finally we would focus on the individuals involved to identify factors associated with the individual.

Our findings confirm the view that case-mix adjustment is not a perfect science and has its own associated risks.[20] The first practical result of our findings is to modify the case-mix adjustment methods. We agree that it is incorrect to interpret residual variation after adequate case-mix adjustment as attributable to either quality of care or the healthcare provider.[21]

Strengths and limitations of the study

The Baker report recommends monitoring at GP level rather than at practice level.[1] GP level data were available to us, however attribution of patients to individual GPs is not reliable and the new GMS contract encourages practice level rather than individual responsibility. We therefore did not deem individual GPs an appropriate unit of analyses in our pilot study. An optimum time interval for mortality monitoring needs to be determined. Our use of 5-year mortality data is arbitrary and the choice of quarterly time slices for the longitudinal charts is determined by the frequency of practice population downloads from the GMS payment system. In our view, 3–5 years of data is sufficient for cross-sectional charts. Monthly longitudinal charts could be used for large practices and quarterly charts for smaller practices.

In keeping with Shewhart's original work, we set control limits at 3σ.[8] Our findings confirm that this level — which has been empirically found to be appropriate in many other settings — was appropriate for these data. Like any statistical technique guidance from Shewhart control charts is subject to misclassification errors.[22] It is tempting to see misclassification errors as analogous to false positives and false negatives in a diagnostic test. However this is incorrect. Diagnostic test characteristics are determined by the extent of agreement between test outcome and a known true

Table 1. Possible explanations for special cause practices.

Practice Code	Primary explanation
1	Nursing home effect
2	Nursing home effect
3	Nursing home effect
4	Nursing home effect
5	Nursing home effect
6	Nursing home effect
7	Very high levels of deprivation and excess morbidity not adequately captured by the case-mix adjustment method
8	High levels of deprivation and excess morbidity not adequately captured by the case-mix adjustment method
9	High levels of deprivation and excess morbidity not adequately captured by the case-mix adjustment method. Practice population becoming younger
10	Practice situated in a violent, deprived area and excess morbidity is not captured by case-mix adjustment method
11	Practice situated in a violent area. Deprivation and excess morbidity not captured by case-mix adjustment method
12	Low levels of deprivation not adequately captured by the case-mix adjustment method. Practice population becoming older
13	Low levels of deprivation not adequately captured by the case-mix adjustment method
14	Very low levels of deprivation and comorbidity not adequately captured by the case-mix adjustment method

An extended version of this table is available as Supplementary information.

state (determined by a gold standard test). They are therefore hypothesis testing. In Shewhart's quality improvement methodology we do not know the true state (nor is there a gold standard test). We are simply determining when it is likely to be useful to generate hypotheses about the reasons for variation. Shewhart control charts are not the only the only graphical tool that could be used for this purpose, techniques such as CUSUMS,[23,24] and SPRT[25] have advocates. We found the strict use of the language associated with Shewhart's theory of variation (common and special cause) to be helpful. It avoids the implication that high mortality is bad performance, that low mortality is good performance or that common cause variation is the 'norm' and therefore acceptable. Educating participating GPs and stakeholders in these concepts was an integral part of the pilot project. Like others[26] we found both support for the principle of monitoring mortality rates and anxiety around public disclosure of such data.

Implications of the study

Several further issues arise from our pilot study. If a practice does not consent to monitoring how should we proceed? Mortality monitoring could be made a contractual requirement. However, as many

GPs found mortality monitoring valuable, providing GPs' concerns are addressed, voluntary participation may be sufficient. How far should we proceed in seeking an assignable cause for special cause variation? To whose satisfaction should this be done? Inclusion of additional stakeholders, such as the Coroners service (currently under review[27]), might make this process more robust and independent. Our experience indicates that such a process benefits from involving stakeholders with in-depth knowledge of the data, analytical methods and the local context.[28] Funding for our pilot project was approximately £25 000, however, as this figure excludes opportunity costs such as GP time, data collection, cleaning and analysis (carried out routinely by the CSA) and the site visits; the true cost of monitoring is certainly higher. Finally, if investigation of a special cause does necessitate a focus on individual carers, how should this proceed?

A recurrent closing question in our workshops was: what is the aim? This pilot was undertaken to monitor mortality rates in response to the Baker report.[1] But our ultimate aim is to use these data to improve quality of care. Mortality rates may not be the most appropriate measure. Several of our practices and others[2] have suggested that quality improvement might be better served by analysis of cause of death. This is possible with the CSA's linked GP mortality database. It remains a challenge to devise a monitoring system that can lead to improvement in quality of care and yet maintain the support and confidence of stakeholders. We believe that we have taken important steps towards meeting this challenge.

REFERENCES
1. Baker R. *Harold Shipman's clinical practice 1974–1998: a clinical audit commissioned by the Chief Medical Officer.* Norwich: Stationery Office. 2001.

2. Baker R, Jones D, Goldblatt P. Monitoring mortality rates in general practice after Shipman. *BMJ* 2003; **326:** 274–276.

3. Bhopal R. Fallout from the Shipman case. Death registers in general practice would be a means of preventing malpractice and murder [letter]. *BMJ* 2000; **320:** 1272.

4. Frankel S, Sterne J, Smith GD. Mortality variations as a measure of general practice performance: implications of the Shipman case. *BMJ* 2000; **320:** 489.

5. Pinder DC. Monitoring the death rates of general practitioners' patients in a single Health Authority. *J Public Health Med* 2002; **24(3):** 230–231.

6. Landrum MB, Bronskill SE, Normand S-LT. Analytical methods for constructing cross-sectional profiles of health care providers. *Health Service Outcomes Research Methodology* 2000; **1:** 23–47.

7. Deming WE. *Out of the crisis.* Cambridge, MA: Massachusetts Institute of Technology 1986.

8. Shewhart WA. *Economic control of quality of manufactured product.* New York, NY: D Van Nostrand Company, 1931.

9. Wheeler D. *Advanced topics in statistical process control. The power of Shewhart's charts.* Knoxville, TN: SPC press Inc, 1995.

10. Mohammed MA, Cheng KK, Rouse A, Marshall T. Bristol Shipman, and clinical governance: Shewhart's forgotten lessons. *Lancet* 2001; **357:** 463–467.

11. Marshall T, Mohammed MA, Lim HT. Understanding variation for clinical governance: an illustration using the diagnosis and treatment of sore throat. *Br J Gen Pract* 2002; **52:** 277–282.

12. Harrison WN, Wall MK, Marshall T. Analysis of inadequate cervical smears using Shewhart control charts. *BMC Public Health* 2004; **4:** 25.

13. Marshall T, Mohammed MA, Rouse A. A randomised controlled trial of league tables and control charts as aids to health service decision-making. *Int J Qual Health Care* 2004; **16(4):** 309–315.

14. The Shipman Inquiry. Reports. http://www.the-shipman-inquiry.org.uk/reports.asp (accessed 8 Aug 2005).

15. The Census Office of Northern Ireland. *Census Small Area Statistics Northern Ireland 1991.* http://www.data-archive.ac.uk/findingData/snDescription.asp?sn=5048 (accessed 3 Aug 2005).

16. Jarman B. Identification of underprivileged areas. *BMJ* 1983; **286:** 1705–1709.

17. Ishikawa K. *Guide to quality control.* Tokyo: Asian Productivity Organisation, 1991.

18. Social Disadvantage Research Centre, University of Oxford. *Measures of deprivation in Northern Ireland.* 2001. http://www.nisra.gov.uk/whatsnew/dep/Measures%20of%20Deprivation%20for%20Northern%20Ireland%20(28th%20June).pdf (accessed 3 Aug 2005).

19. Woodall WH. Controversies and contradictions in statistical process control. *Journal of Quality Technology* 2000; **32(4):** 341–350.

20. Iezzoni LI. The risks of risk adjustment. *JAMA* 1997; **278:** 1600–1607.

21. Lilford R, Mohammed MA, Spiegelhalter D, Thomson R. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *Lancet* 2004; **363:** 1147–1154.

22. Deming WE. *The new economics.* 2nd edn. Cambridge, MA: Massachusetts Institute of Technology,1994.

23. Aylin P, Best N, Bottle A, Marshall C. Following Shipman: a pilot system for monitoring mortality rates in primary care. *Lancet* 2003; **362:** 485–491

24. Steiner SH, Cook RJ, Farewell VT, Treasure T. Monitoring surgical performance using risk-adjusted cumulative sum charts. *Biostatistics* 2000; **1(4):** 441–452.

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

26. Marshall MN, Hiscock J, Sibbald B. Attitudes to the public release of comparative information on the quality of general practice care: qualitative study. *BMJ* 2002; **325:** 1278.

27. Northern Ireland Court Service. *The Coroners Service of Northern Ireland. Proposal for administrative redesign 2004.* http://www.courtsni. gov.uk/en-GB/Publications/ Public_Consultation/ p_pc_TheCoroners ServiceofNorthernIreland ProposalsforAdministrativeRedesign.htm (accessed 3 Aug 2005).

28. Mohammed MA, Rathbone A, Myers P, *et al.* An investigation into general practitioners associated with high patient mortality flagged up through the Shipman inquiry: retrospective analysis of routine data. *BMJ* 2004; **328:** 1474–1477.