

## CHAPTER FOURTEEN

### The Monitoring of Mortality Rates among the Patients of General Practitioners

#### Introduction

- 14.1 At the time when Shipman was in practice, there was no system in place for monitoring mortality rates among the patients of general practitioners (GPs). There is not – and never has been – any requirement that primary care organisations (PCOs) should monitor mortality rates at GP or GP practice level. Nor is any such monitoring carried out by any national body, such as the Department of Health (DoH).
- 14.2 When the scale of Shipman's crimes became evident, many people were surprised that no system of monitoring GP patient mortality rates was in operation. It was suggested that, had such a system existed, it might have alerted the authorities, years before Shipman's eventual detection, to the fact that there was an unexplained excess of deaths among his patients.
- 14.3 In this Chapter, I shall consider the feasibility of setting up a system for the routine monitoring of mortality rates among the patients of GPs. I shall discuss the potential benefits of introducing a system for monitoring mortality rates, and shall consider also the problems that might be encountered by anyone seeking to set up such a system.

#### The Analysis of Mortality Data before Shipman's Arrest

##### In England and Wales Generally

- 14.4 The Inquiry team wanted to know what steps, if any, had been taken by PCOs in the past to monitor mortality rates among the patients of GPs in their areas. Accordingly, in August 2002, the Inquiry distributed a questionnaire to all strategic health authorities (SHAs) in England and health authorities (HAs) in Wales, requesting information about any monitoring of mortality rates which had been undertaken by the PCOs in their areas in the past, or was being undertaken currently. The responses to this questionnaire revealed that, before Shipman's arrest in 1998, there had been very few attempts by PCOs to collect and analyse GP patient mortality data of any kind. Those few attempts that had been made had been directed, not to detecting abnormal mortality rates, but to examining deaths from specific diseases or conditions or, in one case, to providing data to GPs about the deaths of their patients in order to enable them to carry out significant event reviews of those deaths. The Inquiry was told that, in some PCO areas, attempts to analyse mortality data had met with little success, largely because of problems with data linkage.

##### In the West Pennine Area

- 14.5 Neither the West Pennine Health Authority (WPHA) nor its predecessors, the Tameside Family Practitioner Committee and the Tameside Family Health Services Authority (FHSA), had undertaken any monitoring of the mortality rates among the patients of GPs on their list prior to Shipman's arrest. As I have explained, that was entirely typical of the vast

majority of PCOs in England and Wales at that time. There can be no criticism of the WPHA for the fact that it did not undertake any monitoring of this kind.

### **The Reasons Why Monitoring Was Not Undertaken**

- 14.6 There are a number of reasons why PCOs did not monitor GP patient mortality rates. First, as I have said, there was no requirement for them to do so. Nor were they advised that monitoring of this kind might be of value. It seems likely that most PCOs never even considered the possibility of doing so. It is doubtful that it would have occurred to anyone before Shipman's activities came to light that a GP might be guilty of criminality or neglect on such a scale that it would result in an observable excess mortality rate for patients of his/her practice. As I have said, those PCOs which attempted to collect and analyse mortality data did so for purposes other than monitoring the numbers of deaths.
- 14.7 The Office for National Statistics (ONS) holds mortality data collected from death certificates. That data includes details such as cause and place of death, together (except in the case of deaths certified by a coroner) with the name of the doctor who certified the cause of death. The ONS mortality data does not include the name of the GP or GP practice with whom the deceased person was registered.
- 14.8 Individual PCOs hold lists of patients registered with GPs and GP practices in their areas. In order to collect data about the deaths of patients registered with an individual GP or GP practice, it is necessary, under the present system, to link the data held by the ONS with data stored in the PCO systems.
- 14.9 I have mentioned that there were problems with data linkage. These problems arose from the absence of any national system linking the fact or details of a person's death to the GP practice with which that person had been registered. It was possible for a PCO to make the link in an individual case if it wished to do so, by matching ONS mortality data with its own patient list data. However, the task of matching the data from these different sources was laborious and time-consuming. Also, its success depended on the accuracy of the information contained in the PCO system. The Inquiry was told by Dr Peter Goldblatt, Chief Medical Statistician, ONS, that, even now, the accuracy of the data on PCO systems varies widely from area to area. There is a problem in some areas with inflation of lists, i.e. the inclusion on GP patient lists of the names of people who are no longer patients of the GP practice on whose list their name appears. The Inquiry was told that, in some urban areas with a mobile population, there is list inflation by as much as 15% to 20%. Also, the data on a PCO system includes only information relating to patients living within that area. It does not include information about patients who live outside the area but are registered with a GP in the area. Thus, the information on a PCO's system about the patient lists of GPs practising near to the area boundaries of the PCO tends to be incomplete.
- 14.10 As I have said, the problems with data linkage caused difficulties in the past for some PCOs which attempted to collect or analyse mortality data for GPs in their areas. It no doubt had the effect of deterring others from making the attempt at all. Until 1998, PCOs confined themselves in general to conducting analyses which examined deaths from specific causes among resident populations at electoral ward level. The data for such analyses was readily available and no complex linkage was required. The WPHA carried

out this type of analysis with the aim of identifying any socio-economic, environmental or other factors that might be affecting death rates.

## The Analysis of Mortality Data after Shipman's Arrest

### By the West Pennine Health Authority

- 14.11 After Shipman's arrest and after it became clear that he might have killed a large number of his patients, the WPHA conducted various analyses of the mortality rates among his patients. The object of these analyses was, first, to discover the likely extent of his criminality and, second, to ascertain whether it would have been evident that Shipman's mortality rates were excessive if those rates had been monitored during the period of his criminality. The task of linking the data used in these analyses was difficult and time-consuming.
- 14.12 The analyses showed that, in several years, Shipman had a statistically significant excess number of deaths. However, when the analyses were focussed on the deaths of patients which had occurred at the patient's home or on the deaths among his elderly female patients, there was a statistically significant excess of deaths among these groups of patients in most years from 1992 to 1998.
- 14.13 In April 1999, the WPHA submitted to the DoH a report on the policy implications for the NHS arising from what was then known about Shipman's activities. In that report, the WPHA recommended that routine monitoring of GP patient mortality rates should be introduced and that the value of such monitoring should be tested. The report acknowledged that there might be difficulties in interpreting apparent excesses in the mortality rates of the patients of individual GPs. However, it suggested that such excesses might give an indication that more in-depth investigation was warranted in the case of a particular GP. The monitoring of mortality rates might, it was suggested, form part of **'the overall performance assessment framework'**.

### By Other Primary Care Organisations

- 14.14 Shipman's arrest and conviction caused a number of other PCOs to consider conducting some form of analysis of mortality rates among patients of the GPs in their areas. Some PCOs carried out a 'one-off' exercise. Others were more ambitious and attempted to develop systems of monitoring mortality rates on an ongoing basis. A few developed systems for gathering and analysing mortality data which was then fed back to GPs, as part of a collection of performance indicators. Following receipt of the responses to its original questionnaire, the Inquiry sought and obtained further information from some PCOs which had undertaken analysis of mortality statistics over the last few years. Their accounts of their experiences were very illuminating and I shall refer to some of them later in this Chapter.

### By Professor Richard Baker

- 14.15 Following Shipman's convictions for murder in January 2000, the Chief Medical Officer, Professor (now Sir) Liam Donaldson, commissioned a clinical audit of Shipman's practice.

The audit was carried out by Professor Richard Baker, Director, Clinical Governance Research and Development Unit, University of Leicester. Professor Baker first analysed the pattern of deaths in respect of which Shipman had issued a Medical Certificate of Cause of Death (MCCD). As I explained in my First Report, a GP will generally be in a position to issue a MCCD only when a patient has died in the community (i.e. at home or in a care home) and where the death does not have to be reported to a coroner. If a patient dies in hospital, the MCCD will usually be issued by a hospital doctor. The process of issuing a MCCD is often termed 'certifying the death'.

- 14.16 Professor Baker carried out two analyses. His first analysis involved comparing the number and various features of the MCCDs issued annually by Shipman with those of the MCCDs issued annually by two control groups of other GPs who had been in practice at the same time and in the same localities (i.e. Todmorden and Hyde) as Shipman, and whose patients had similar socio-economic characteristics.
- 14.17 Professor Baker's second analysis compared the annual number of deaths which had occurred among persons who had been patients of Shipman from 1987 onwards (not just those patients whose deaths Shipman had certified) with the annual number of deaths that could have been expected to have occurred among that patient population. Shipman's patients were identified by means of the patient list data held by the WPHA. The ONS performed the necessary linkage of that data with information about deaths. The number of expected deaths was calculated by reference to figures which were provided by the ONS and which related to deaths among the population of the local (Tameside) district and of a group of districts sharing population and socio-economic characteristics similar to those of Tameside, and also by reference to deaths among the general population of England and Wales.
- 14.18 Professor Baker's analyses indicated that, while practising in Todmorden and Hyde, Shipman had issued an 'excess' total number of 236 MCCDs relating to deaths occurring at the patient's home or on his practice premises. This figure was very similar to the number of deaths that I found had been caused by Shipman during his time in Todmorden and Hyde.
- 14.19 At the conclusion of the report on the results of his clinical audit, Professor Baker made a number of recommendations. Among these was a recommendation that systems for the monitoring of GPs should be reviewed and extended to include routine monitoring of mortality rates among their patients. He pointed out that Shipman's case had demonstrated that, in the absence of the monitoring of mortality rates, it was possible for a GP with a sustained excess rate of mortality among his/her patients to go undetected for many years. However, Professor Baker acknowledged that there were a number of difficulties with the monitoring of mortality rates. The first was their inherent variability. The second was the fact that, under normal circumstances, many of a GP's patients will die at a time when they are not under his/her management. The most common example of this is when a patient dies in hospital. Professor Baker suggested that his clinical audit had highlighted some factors that might improve the ability of monitoring systems to detect abnormal mortality rates or criminal activity. He also suggested that the monitoring of the number of deaths certified by GPs – rather than the mortality rates among their registered

patients – would be more indicative of the GP’s clinical activities. He also referred to the potential value of analysing the excess cumulative mortality rate of GPs so that all excess deaths of a GP throughout his/her career would be assessed, not just those occurring in individual years.

- 14.20 Professor Baker went on to point out that the collection of data resulting from the monitoring of mortality rates would not of itself be sufficient. There would have to be a regular review of the findings. He said that one argument against the monitoring of mortality rates was that it would require too great an investment of time and other resources and that GPs were already over-stretched. It might, he said, also be contended that such monitoring was unnecessary, since cases such as Shipman’s were exceptionally rare. It might also be said that individuals determined on murder would adopt strategies to avoid detection. Nevertheless, he believed that steps should be taken to introduce a system of monitoring mortality rates.

### **The Response of the Department of Health**

- 14.21 On 1<sup>st</sup> February 2000, immediately following Shipman’s conviction, the then Secretary of State for Health, the Rt Hon Alan Milburn MP, made a statement to the House of Commons, in which he announced that the DoH was working with the ONS **‘to find new and better ways of monitoring deaths of GPs’ patients’**. Since then, the DoH has been considering the feasibility and practicability of collecting and using mortality data to monitor deaths of patients at GP or GP practice level. I shall refer to the outcome of the work undertaken by the DoH later in this Chapter.

### **The Inquiry’s Approach**

- 14.22 It was evident to me from an early stage that the Inquiry had to examine the possibility of setting up a system for the routine monitoring of mortality rates among the patients of GPs. I wanted to discover whether, using existing data sources, it would be possible to provide complete and accurate information on which such monitoring could be based. I also wanted to find out whether such monitoring would be capable of identifying abnormalities in mortality rates and, if so, at what level (i.e. GP, GP practice or primary care trust (PCT) level) it would be capable of doing so. If routine monitoring was likely to be capable of detecting abnormalities at any level, I wanted to know who would be best placed to organise such monitoring and by what method it could best be done. In order to assist me in answering these questions, it was clear that the Inquiry needed the services of an expert.

### **The Commissioning of Work from Dr Paul Aylin and His Team**

- 14.23 The Inquiry team commissioned Dr Paul Aylin, Clinical Senior Lecturer in Epidemiology and Public Health, Imperial College of Science, Technology and Medicine, to carry out the necessary work and to report. Dr Aylin was a consultant to the Public Inquiry into children’s heart surgery at the Bristol Royal Infirmary (the Bristol Royal Infirmary Inquiry) and had previously been employed by the ONS as a medical statistician. He assembled a team, all from Imperial College, consisting of Dr Nicky Best (Senior Lecturer in Statistics), Dr Alex

Bottle (Researcher) and Dr Clare Marshall (Lecturer in Statistics). Dr Aylin and his team were asked to determine:

- if there was any benefit in monitoring mortality data at local level
- at what level (i.e. GP, GP practice or PCO) such monitoring should be conducted
- how the monitoring should be done
- what method of analysis would be appropriate
- at what point there should be concern about the patient mortality rate of the unit (i.e. GP, GP practice or PCO) being monitored
- who should do the monitoring.

14.24 Dr Aylin and his team prepared a report and, in July 2003, gave a presentation of their work to the Inquiry. An account of their work has subsequently been published in *The Lancet*<sup>1</sup> and is at Appendix E to this Report.

### **The Inquiry's Seminar**

14.25 The work of Dr Aylin and his team, together with wider issues relating to the monitoring of GP patient mortality rates, was discussed at a two-day seminar held by the Inquiry in October 2003. Participating in that seminar were a number of experts in the field. As well as Dr Aylin and his team, there were Dr John Fox (Director of Statistics, DoH), Dr David Spiegelhalter (Medical Research Council Biostatistics Unit), Dr Mohammed A Mohammed (Senior Research Fellow, Department of Public Health and Epidemiology, University of Birmingham), Dr Christopher Roberts (Senior Lecturer in Medical Statistics, School of Epidemiology and Health Sciences, University of Manchester) and Dr Goldblatt. In addition, the seminar was attended by Professor Baker, Dr Maureen Baker (representing the Royal College of General Practitioners (RCGP)), Dr John Grenville (representing the British Medical Association (BMA)) and Professor Gwyn Bevan (then Acting Head of Information, Commission for Health Improvement (CHI) and a member of the transition team for the new Commission for Healthcare Audit and Inspection (now known as the Healthcare Commission)). In addition, the Inquiry invited to the seminar representatives from a number of PCOs who had experience of carrying out work involving the analysis of mortality rates among the patients of GPs or GP practices. A full list of seminar participants is at Appendix F to this Report.

14.26 Participants in the seminar submitted written material in advance and expanded on that material during the course of discussions at the seminar. The discussions were led by Leading Counsel to the Inquiry. The seminar produced a very interesting discussion about the benefits that might accrue from the monitoring of mortality data, and about the problems that might be associated both with the monitoring process and with the investigation that must necessarily follow the detection of any apparent abnormality. The discussion concluded, as will become clear, with an encouraging degree of consensus on the way forward.

---

<sup>1</sup> Aylin P, Best N, Bottle A, Marshall C (2003) 'Following Shipman: a pilot system for monitoring mortality rates in primary care', *The Lancet*, Vol 362: pp 485–491.

- 14.27 I am most grateful to Dr Aylin and his colleagues for the work that they have done for the Inquiry. It is innovative and, as I had hoped, it has made a real contribution to the debate about the feasibility and the value of setting up a system for the routine monitoring of mortality rates among the patients of GPs. The work of Dr Aylin and his team appears to have been well received by their peers and was the subject of much praise at the seminar.

## **The Work of Dr Paul Aylin and His Team**

- 14.28 The report prepared by Dr Aylin and his team is on the Inquiry's website. Their work is well described in the published account to which I have already referred. I shall refer here to some of the main points of the system which they have devised.

### **Data Linkage**

- 14.29 As I have already explained, in the past, the absence of any national system of linking mortality data held by the ONS to the GP or GP practice with whom the deceased person had been registered made it difficult to collate the data necessary to carry out an analysis of mortality rates among the patients of GPs. Even if a link was made, the resultant data was likely to be incomplete and, probably, inaccurate. Dr Aylin and his team carried out a pilot exercise aimed at linking national death registration records to the GP patient list data for five former HAs, including the WPHA. A special computer program was used to link ONS mortality data for the years from 1993 to 1999 with anonymised patient list data. Deaths were linked to the GP with whom the deceased person was registered, using patients' NHS numbers, date of birth, gender and postcode. The NHS Information Authority provided Dr Aylin and his team with patient list sizes for every GP and GP practice in the five HA areas during the years from 1993 to 2000. The data divided the patient population into three age bands (0–64, 65–74 and 75 and over). These age bands were broad, but constituted the best information available for the period.
- 14.30 The overall degree of success in linking death registration records to patient lists varied between each of the five HA areas. There was a low success rate (no more than 60%) for the WPHA up to 1999. This apparently resulted from loss of data at the time when two former FHSAs merged to create the WPHA. In general, the linkage improved markedly after 1997, when individual patient NHS numbers first became available. Dr Aylin told the Inquiry that this improvement suggested that data linkage in any future monitoring exercise would be more successful. The overall success rate in linking records was 92%. In 2000, it was 98%–99%.
- 14.31 Since 2000, the DoH has been working on the provision of a single database containing information on every deceased NHS patient, including the identity of the GP or GP practice with whom the patient was registered and the cause of death. The database is to be held by the NHS Information Authority. It has already been tested by several PCTs and it seems likely, as Dr Aylin suggested, to provide better linkage and more accurate data than was available to Dr Aylin and his team when they carried out their analysis.
- 14.32 When the pilot linkage exercise was carried out by Dr Aylin and his team, the patient list data for GPs whose practices were near the boundary of a HA area appeared to be

incomplete. This was probably because some of the patients of those GPs lived in the areas of neighbouring HAs. The data relating to such a patient would be contained in the database for the HA area in which s/he lived. In an attempt to overcome the problem of incomplete data, Dr Aylin and his team excluded from their analysis those GP practices with patient list sizes of fewer than 1000 patients for any of the years from 1993 to 1999. They did so on the assumption that those practices' patient lists must be incomplete. However, exclusion of some practices is not a solution which could be adopted for a national system of monitoring mortality rates. Dr Aylin emphasised that, in order to avoid the problem in the future, it would be necessary to have a national system of data linkage.

- 14.33 Dr Aylin and his team chose to look at annual death rates, rather than death rates over a shorter period. This was primarily because of the small number of deaths involved. Another factor was the delay between the occurrence and registration of a death and the compilation of mortality statistics nationally. In the past, there has been a considerable time lag (amounting to several months) before the relevant mortality data was available and could have been linked with patient list data. This would have meant that monitoring would have had to take place several months in arrears. Dr Goldblatt said that it should be possible in the future to obtain information about 97% – possibly more – of deaths within about two weeks of the death occurring. This would mean that analysis of mortality statistics could take place much earlier than would previously have been possible. Dr Mohammed suggested undertaking analyses on a six-monthly or quarterly basis. Dr Aylin said that this would be possible under his system. However, he would not advise more frequent analysis in view of the small numbers of deaths involved.

### **The System of Prospective Monitoring**

- 14.34 Dr Aylin and his team developed a system that could be used to monitor mortality rates prospectively. The system was intended to be a screening tool for the purpose of identifying units which appeared to have mortality rates higher than would be expected. Dr Aylin and his team described the system as a 'surveillance' system. Dr Spiegelhalter observed at the seminar that surveillance would act as an 'automatic whistleblowing procedure' which would alert people to the fact that something might be going wrong. Once such a signal had sounded, a decision would then have to be taken as to what, if any, action was needed in response to it, in order to detect whether there was a problem with poor clinical practice affecting death rates or even with potential criminal activity. He distinguished the 'surveillance' of mortality data from the 'monitoring' of mortality data. The latter, he suggested, would be undertaken as part of the monitoring of a pool of information which could be examined and used for the purpose of quality improvement. Clearly, both 'surveillance' and 'monitoring' fulfil useful functions. Rather than attempting to distinguish between 'surveillance' and 'monitoring' throughout this Chapter, I shall use the word 'monitoring' to describe both.
- 14.35 The distinctive feature of prospective monitoring is that it involves monitoring continuously over time, rather than taking a single snapshot look and comparing performance at one time point. Data is accumulated over time and the analysis is repeated at every time point. This has particular benefits when viewing the small numbers of deaths among patients of



GPs. The aim is to detect unusual variation in the underlying mortality rate of any unit (e.g. of an individual GP) as soon as possible after it has occurred.

- 14.36 It should be noted that the analysis carried out by Dr Aylin and his team related to the deaths of all patients registered with the relevant GPs, not just to those deaths which the GPs had certified. The number of deaths for each GP was compared with the number of deaths which could be expected among the total patient population for that GP. It was not possible to perform an analysis by reference to the number of deaths certified by each GP. This is because there was no comparison patient population available by which to calculate the number of deaths which each GP could be expected to certify.

### **Statistical Process Control Charts**

- 14.37 The method developed by Dr Aylin was derived from statistical process control (SPC) which was originally developed for use in the quality control of industrial processes. It is based on the recognition that the outputs of even the most stable and perfectly tuned production process will inevitably show some variation (known as common cause variation). This common cause variation means that, even under ideal conditions, a group of similar doctors will never match one another's performance from one period to the next.
- 14.38 In developing a system of monitoring, it is important to be able to distinguish 'common cause' variation from variation that has occurred as a result of a 'special cause'. A special cause is a factor extrinsic to the normal functioning of the process being monitored, which gives rise to variation over and above the common cause variation that can inevitably be expected in any stable process. Common causes will be responsible for most of the variation observed in a monitoring system and will produce 'background noise' in the monitoring of a stable process. Ideally, the monitoring system should be designed to 'signal' the fact that variation due to some special cause has or might have occurred and to indicate that steps should be taken to identify the special cause.
- 14.39 SPC charts were designed to detect and signal the point when an industrial process goes from being stable, in control and exhibiting only common cause variation to being unstable, out of control and exhibiting special cause variation. The degree of common cause variation may be acceptable. If it is not, action on the underlying process will be necessary to eliminate or reduce it. If the process becomes unstable – because there is some special cause variation – an investigation will be necessary in order to determine what special cause is responsible for the variation. That special cause must then be addressed as necessary. Special cause variation is not necessarily a bad thing; whether it is good or bad is a matter for judgement depending on the circumstances of the particular case.
- 14.40 SPC charts are created with 'control limits', representing the upper and lower limits of common cause variation. Suitable limits, like the limits for any diagnostic test, must balance the disadvantage of erroneously signalling that a special cause is present when it is not (i.e. a false positive) and the disadvantage that would result from failing to detect the presence of a special cause (i.e. a false negative). Control limits with high specificity (i.e. capable of discriminating to a high level) are the most appropriate.

- 14.41 The system of monitoring advocated by Dr Aylin and his team is similar to that described above, in that it aims to identify and signal when a GP's mortality rates appear to be subject to (special cause) variation over and above that variation which would be common to all GPs with similar features. The application of SPC charts involves defining when mortality rates may be considered acceptable or 'in control' and when they are deemed to have become unacceptable or 'out of control'.
- 14.42 Plotted on the SPC chart is the value of the chart statistic, which represents the difference between the observed outcomes (i.e. the annual mortality rates) for the relevant unit (in this case a GP or GP practice) and the outcomes which would be expected if the GP or GP practice was performing acceptably. The chart statistic is plotted against time. A pre-defined alarm threshold is set. All the time the chart statistic stays below that threshold, the mortality rates of the GP are considered to be in control and showing common cause variation only. However, if the chart statistic crosses the alarm threshold, it signals that there may be some factor affecting the GP or GP practice which has caused the mortality rate to become unstable or out of control and to exhibit special cause variation. The signal should then be followed up and investigated in order to find an explanation for it.
- 14.43 The earliest type of SPC chart is known as the Shewhart control chart (named after Walter Shewhart, who first devised it). By convention, the control limits on a Shewhart chart are set in such a way that 99.8% of the common cause variation would fall within these limits. A sequence of observed outcomes over time is plotted on the chart and when, at any time point, the observed outcome goes above the upper control limit or below the lower control limit, the chart signals that there is special cause variation. The chart statistic on a Shewhart chart is independent at each time point. It does not depend on or reflect historical data. There is no accumulation of previous values of the chart statistic. There is, therefore, no accumulation of evidence of any unusual variations that may have occurred over a number of time points.

### **Cumulative Sum Charts**

- 14.44 By contrast, the type of SPC chart used by Dr Aylin and his team (known as a cumulative sum (CUSUM) chart) adds in 'new' observations to the previous value of the chart statistic. Thus, each observation is based on both historical and new observations. The reason why Dr Aylin and his colleagues used this approach was that they believed that what was needed was a system that would detect substandard or aberrant practice over a period of time. The effects of such substandard or aberrant practice may not be evident at a single time point. They may become obvious only when evidence of performance over an extended period is accumulated and considered. I have already mentioned that, in the report of his clinical audit, Professor Baker had recognised the potential benefits of monitoring cumulative mortality rates and had recommended that the possibility of using a cumulative method of monitoring should be investigated.

### **Variation**

- 14.45 One of the problems with transferring the technique of monitoring used in an industrial process control setting to a healthcare setting is that the process going on within a GP

practice is very much more complex than an industrial process. As a consequence, the extent of intrinsic variation (common cause variation) is typically quite large. It therefore becomes very important to identify variation which is intrinsic and to distinguish it from the special cause variation which the monitoring is designed to detect.

- 14.46 When comparing the mortality rates for the patients of GPs from year to year, there are likely to be considerable differences between individual GPs and even between the mortality rates for patients of the same GP from year to year. These differences can be caused by several types of variation.

### ***The Type of Variation that the Monitoring System Is Designed to Detect***

- 14.47 When setting up a monitoring system, it is necessary to identify in advance the particular type of variation that the monitoring is intended to detect. In the case of the monitoring of mortality rates, monitoring is intended to detect unusual variations in the mortality rate or special cause variation. Such special cause variation would occur, for example, in the presence of criminal behaviour such as Shipman's or where poor quality of care by a GP was giving rise to an increased mortality rate among his/her patients. Special cause variation of this kind will exist in the case of only a small number of GPs. When monitoring mortality rates, the object is to ensure that the GPs who are subject to special cause variation are the ones who signal. It is also important to ensure, insofar as it is possible to do so, that their signals are not obscured by 'background noise' from other types of common cause variation.
- 14.48 In order to eliminate (or, at least, to reduce) the effect of variations other than those resulting from systematic difference between GPs, it is necessary to recognise the types of variation that are inevitable and to use statistical methods to quantify and adjust for those types of variation. Dr Aylin and his team attempted to quantify and to adjust for a number of different types of common cause variation that, if not adjusted for, would potentially have had an effect on mortality rates. Dr Best described those types of variation to the Inquiry.

### ***Chance Variations***

- 14.49 The first type of common cause variation that Dr Best described was the type that occurs entirely randomly, as a matter of chance. The number of deaths among the patients of a GP will vary from year to year, purely as a matter of chance. There are statistical methods, which Dr Aylin and his team employed, of adjusting for the effects of chance.

### ***Case Mix***

- 14.50 The second type of common cause variation is known as 'case mix'. Case mix factors are factors associated with a GP's patient population which are known to have an effect on mortality rates. The most obvious such factors are age and gender. If no adjustment were made for age, GPs with elderly patient populations would be likely to signal as having higher than expected mortality rates whereas, when the age of their patient population was taken into account, their mortality rates might be within a normal distribution. In

addition, a GP with significantly more elderly male patients than elderly female patients would be expected to have a higher mortality rate than one with equal numbers of elderly male and female patients. Another case mix factor known to have an effect on mortality is socio-economic deprivation. If no adjustment is made for that factor, GPs with a deprived population would be more likely to signal.

- 14.51 Dr Best emphasised that care must be taken when making adjustments for known case mix factors. Adjustments should be made only for factors that are beyond the control of the GP. Otherwise, adjustment might have the effect of masking important variations caused by systematic factors within the GP's practice, which are the very type of variations that the monitoring system is intended to detect. In this context, other participants in the seminar mentioned the dilemma as to whether to adjust for socio-economic deprivation. Mrs Catherine Scott, Specialist in Public Health, Croydon PCT, said that patients in deprived areas may not currently receive the health care they need. This can happen because of a shortage of GPs in deprived areas or for other reasons which might be associated with variations in GP performance or poor clinical practice. These are matters which, in the interests of public health, the PCT might wish to have highlighted, rather than to risk their being masked. Mrs Scott said that, in 2003, her PCT had prepared two sets of mortality data, one adjusted for socio-economic deprivation and one not. The intention was to see how useful the figures proved to be.
- 14.52 As mentioned earlier in this Chapter, because of the lack of data held by the NHS about patients registered with particular GPs and GP practices during the years from 1993 to 1999, Dr Aylin and his team were only able to adjust for age within three broad age bands (0–64, 65–74 and 75 and over). They used an indirect standardisation method to estimate the expected number of deaths for each GP and GP practice. They had no information on gender or on socio-economic deprivation. Thus, their adjustment was not as sensitive as they would have wished.

### ***Variations Due to External Factors***

- 14.53 Another type of common cause variation which must be taken into account is that caused by external factors unrelated to case mix. An example of this would be the effect of a known event such as a 'flu epidemic, or a spell of particularly harsh weather, either of which could cause a sudden increase in mortality rates.

### ***Variations Due to Unaccounted-for Factors***

- 14.54 A fourth type of common cause variation is attributable to factors not otherwise accounted for. Even after allowing for the effects of chance, for known case mix and for external factors, there may be other factors which could be expected to affect a GP's mortality rate but which are not known to the people carrying out the monitoring. The individual unknown factors may be small, but their combined effect may lead to significant variations. Indeed, work carried out by Dr Aylin and his team suggested that the variation in mortality rates at GP level caused by unaccounted-for factors was about twice that which would be expected as a result of chance variations alone. Dr Best explained that unaccounted-for factors are important but are often ignored by those conducting monitoring of this kind.

She described the method which Dr Aylin and his team used to adjust for this type of variation. It is unnecessary to describe it in detail. However, it is perhaps worth noting that Dr Best considered that there was a need for further work to be done in this area before a monitoring system was implemented.

### **The Incidence of False Alarms**

- 14.55 When SPC charts are used in an industrial setting, they are used to monitor only a single process at a time. They have also been used for public health and surgical monitoring but, even then, such monitoring has involved only one or two processes at a time. This is very different from a national system for monitoring the mortality rates among the patients of all GPs or GP practices simultaneously. The fact that multiple units (rather than just one unit) are being monitored over time will increase the chances of getting a signal. It will also increase the chance of getting false alarms. It is obviously desirable to reduce the number of false alarms since they will divert attention and resources away from the true alarms that really require investigation. Over time, a large number of false alarms with no positive results would bring the whole monitoring system into disrepute.
- 14.56 It is important to define exactly what is meant by a 'false alarm'. A statistical false alarm will occur when a GP's patient mortality rate triggers an alarm but the rate is in fact in control and is not exhibiting special cause variation. By contrast, a statistical true alarm occurs where the signal is valid, i.e. the mortality rate *is* higher than would be expected after appropriate adjustments have been made for common cause variation. Many statistical true alarms will be medical false alarms, in that there will be an acceptable clinical explanation for the fact that the patient mortality rate of a GP or GP practice is statistically out of control. There will, therefore, be no reason for concern. A case such as that of Shipman would, however, be both a statistical and a medical true alarm since the mortality rate was indeed out of control and this was due to a special cause variation, namely his criminality.
- 14.57 The alarm threshold should be set in such a way that the number of statistical false alarms is minimised while, at the same time, the system remains sensitive to true alarms from unstable units which are really exhibiting special cause variation. Dr Aylin and his team ran computer simulations to calculate the false discovery rate (i.e. the proportion of signals occurring before a given time point that will be statistical false alarms) and the successful discovery rate (i.e. the proportion of unstable units exhibiting special cause variation that will be successfully detected before the same time point). They carried out the simulations using different alarm thresholds. They suggested that the false and successful discovery rates could form part of a cost benefit calculation to assist in setting an appropriate alarm threshold for a monitoring system. They suggested also that factors to be taken into account in making such a cost benefit calculation might include the importance attached to detecting GPs whose mortality rates showed special cause variation, relative to the importance attached to avoiding false alarms. The resource implications associated with investigating the cause of alarms would also be a relevant factor.
- 14.58 Dr Aylin and his team explained that it would be possible to 'tune' a CUSUM chart by identifying two (or possibly more) different thresholds. For example, a lower threshold

would pick up a lower level excess of mortality rates and would constitute an 'early warning', signalling a GP whose mortality rate was at the extreme end of the normal distribution of mortality rates for GPs. This might merit a low-level investigation. A higher threshold might also be set, with a view to sounding an alarm if the mortality rates of a GP became divergent from the normal distribution. A more detailed investigation might then be undertaken. The charts could thus be 'tuned' in such a way as to meet the goal of the monitoring exercise. For the purpose of illustrating the application of their SPC charts, Dr Aylin and his team used two different thresholds.

### **The Results of the Analyses Carried Out by Dr Aylin and His Team**

- 14.59 Dr Aylin and his team used their CUSUM charts to examine mortality data for 1009 GPs (including Shipman) who practised in the five HA areas during the years 1993 to 1999. The data used was, of course, past data, although Dr Aylin and his colleagues looked at what the charts would have revealed had they been viewed contemporaneously. Data was available only as far back as 1993, so no account could be taken of any excess deaths among Shipman's patients prior to that time.
- 14.60 The CUSUM chart for Shipman shows that his chart statistic was above zero for 1993 indicating that he had had a greater number of deaths than would have been expected. It fell in 1994, which suggested that he had fewer deaths than expected in that year, given the age distribution of his patients. In fact, the Inquiry found that he killed 11 patients that year, but he might have had few other patient deaths. Thereafter, the chart statistic rose steadily as the excess deaths accumulated. At the lower of the two thresholds set by Dr Aylin and his team, Shipman would have triggered the alarm first during 1996. At the higher of the two thresholds, he would have triggered it first during 1997. He would have continued to signal up to 2000, despite the fact that he ceased practice in September 1998. This is because his GP practice code was taken over by a locum and Shipman's historical data continued to exert an effect. However, after the time when the locum started, a fall in the chart statistic was evident.
- 14.61 Dr Marshall told the Inquiry that the appearance of the chart for the period after the arrival of the locum in September 1998 highlighted two issues. First, the fact that Shipman's GP practice code had been passed from one doctor to another would not normally be known to the person or body carrying out the monitoring. This could lead to misleading results if monitoring continued. It would clearly be desirable that where, as in this instance, a single-handed practitioner is replaced by another doctor, there should be a change of GP practice code, so that monitoring could be restarted.
- 14.62 The second issue was what to do when a GP signalled. In this instance, the practice had continued to signal, despite the fact that the special cause variation (i.e. Shipman's criminality) had been removed and the mortality rate had returned to being in control. One option was that, once an acceptable explanation for a signal had been found, the chart should be re-set at zero. That would have the effect of 'wiping the slate clean' and the GP would then have a fresh start. The alternative would be to re-set the chart to some value greater than zero. The effect of that would be to put the GP 'on probation'. The latter approach would make a chart more sensitive to sustained excess mortality.

- 14.63 For illustrative purposes, Dr Aylin and his team assumed that 5% of GPs would show special cause variation. This of course does not mean that they were assuming that 5% of GPs have excess mortality rates because of criminality or poor clinical practice. They may have a statistical excess of deaths for quite innocent reasons, for example because many of their patients live in nursing homes or because they care for patients being treated in a local hospice. Again for illustrative purposes, Dr Aylin and his team set two different alarm thresholds for detecting GPs with special cause variation. Shipman was one of 33 GPs out of the 1009 monitored who would have signalled during the period of monitoring between 1993 and 1999 if the lower threshold was applied. At the lower threshold, Dr Aylin and his team calculated from their computer simulations that the false discovery rate would be 5.2%. That would mean that 5.2% of the 33 signals were likely to be statistical false alarms. They also calculated that the successful discovery rate was 96.6%. That meant that the system had a 96.6% chance of correctly identifying any GP whose mortality rate was subject to special cause variation.
- 14.64 If the alarm threshold were set at a higher level, then only 12 GPs (including Shipman) would have signalled. Dr Aylin and his team calculated from their computer simulations that the higher threshold had a false discovery rate of less than 0.01%. It could therefore be expected that only a tiny proportion of signals would be statistical false alarms. There is, therefore, a trade-off in setting different alarm thresholds. A low alarm threshold would have the effect of identifying a larger number of GPs who appeared to show special cause variation and would need to be investigated; some of those signals would turn out to be statistical false alarms. A higher threshold would identify a smaller number of GPs and would have a very low statistical false alarm rate. However, it might miss some GPs who would merit investigation because they had an excess of deaths.

### **The Lessons to Be Learned from the Work**

- 14.65 Dr Aylin concluded by saying that the work carried out by himself and his team demonstrated that CUSUM charts could be used to monitor patient mortality rates at GP level and that they would have been capable of detecting Shipman if they had been in use at the relevant time. However, detection of excess mortality rates among the patients of GPs (unless very large indeed) would be impossible from an analysis at PCT level or higher. He recognised that, in reality, patients of group practices may receive treatment from more than one GP, so that monitoring at GP practice level might be more appropriate than at the level of an individual GP. However, he acknowledged that it would be more difficult to detect an unusual mortality pattern attributable to an individual GP if monitoring was being done at GP practice level, rather than at individual GP level. I shall return to this topic later in this Chapter.
- 14.66 Dr Aylin emphasised that CUSUM charts can highlight unusual patterns of mortality but cannot themselves shed any light on the reasons for them. That must be an issue for investigation. He also observed that, if monitoring is to be undertaken, there must be some infrastructure to undertake the necessary investigations. There would be no point in running the CUSUM charts and taking no notice of the signals that they produced. He suggested that those responsible for investigating the signals would need to have the

necessary information to enable them to investigate effectively. He thought that PCTs would be best suited to this task. Others do not agree with this, as I shall explain.

- 14.67 Dr Aylin and members of his team told the Inquiry that, when they looked at the mortality data of the 1009 GPs, there was an enormous degree of variation between them. That variation was greater than would be expected by chance alone and was only partially explained by case mix factors such as age, gender and socio-economic deprivation. Dr Aylin and his team attributed the high degree of variation in part to the poor quality of some of the historical data available to them. I have already mentioned that the data available in the future is likely to be of very much higher quality. Another factor was the inadequacy of the available information about case mix factors, which meant that a proper adjustment for case mix factors could not be made. Dr Aylin said that there is now more detailed data available. Data by electoral ward is available and can be linked with census data to give better adjustments for age, gender and socio-economic factors. Another very important element which was likely to have contributed to the variation in mortality rates was the unaccounted-for factors to which I have already referred. Dr Aylin said that a key message was that the variation caused by these factors must be adequately taken into account in any monitoring system that is implemented. It seems to me that, if a prototype system were devised and put into operation, much could be learned about variation from the experience gathered in the first year or two.
- 14.68 Dr Aylin said that it would be possible to set the CUSUM charts to detect either very high levels of mortality within a short time of their occurrence or more consistently low levels of mortality over a longer period of time. Dr Best observed that the crucial element needed to give the necessary flexibility to detect different types of outcome was good quality data, not only data relating to actual deaths, but good quality 'denominator data', i.e. information relating to the patient population being used to calculate the number of expected deaths. The denominator data is used to calculate the chart statistic for the CUSUM control chart. She said that this denominator data was likely to improve in the future, particularly with the introduction of electronic patient records. Dr Aylin also emphasised that, in statistical analysis, the denominator data was as important as the data relating to observed deaths. Dr Spiegelhalter agreed that the same statistical process could be used to detect both sudden changes in mortality rates and more gradual changes. It was just a matter of tuning the system to be 'interested' in different types of variation.

## The Possible Framework for a Monitoring System

### The Need for Analysis to Be Performed by a Central Unit

- 14.69 There was agreement at the seminar that the collection and linkage of data should be performed centrally. Dr Aylin said that the quality of the data contained on the current NHS system varied from area to area and there were still difficulties with patients who lived in one area but were registered with a GP practice in another area. The data would be more accurate if it were collected countrywide. Dr Goldblatt, of the ONS, agreed that there was a 'huge variability' in the quality of the patient list data available for different parts of the country. If the quality of mortality data was to be comparable in all areas, there must be a national system for linking data and for feeding it back to those responsible for clinical



governance locally, i.e. the PCTs. As I have said, the DoH plans that the database should be held by the NHS Information Authority.

- 14.70 Dr Aylin believed that analysis of the data could be undertaken efficiently and relatively cheaply by a small national unit. He suggested that many PCTs would not have the expertise necessary to carry out this task. The results of the analysis at the central unit could be fed back to PCTs, which would in turn pass on the results to individual GP practices. Dr Spiegelhalter agreed and said that central analysis of mortality data and dissemination of the results would be valuable in terms of efficiency and would also enable guidance to be given on the interpretation of the results of the analysis. Dr Fox told the Inquiry that the DoH had no plans to undertake any analysis of mortality data itself. At the time of the seminar, it was envisaged that this would be done by the new Healthcare Commission.

### **The Importance of Local Knowledge**

- 14.71 Dr Aylin and Professor Baker both emphasised that PCTs must examine the results of the central analysis themselves and monitor them. It would not be sufficient for them merely to pass the results back to GP practices. Of course, the PCTs would know that a high mortality rate does not necessarily indicate that a doctor is failing in some way. Still less is it indicative of criminal activity. In most cases, the high rate will be explained by some characteristic of the doctor's practice or of his/her patient population. An obvious example would be a doctor who cared for patients in a local hospice or in one or more local care homes. The death rates among such patients might well account for a higher death rate among patients of the GP than would be expected in a practice where patients did not have these characteristics. This type of information specific to a GP practice is not likely to be available outside the immediate locality and may become known to a PCT only after discussions with the GP or GP practice concerned.
- 14.72 Professor Baker said that he expected that the first two or three years after the introduction of a monitoring system would be spent accumulating a lot of additional knowledge about practices and their patient populations. This would be a very important phase. Once the knowledge had been acquired, monitoring would move into a rather different phase of long-term observation. Professor Baker said that he hoped the point would be reached where all practices – not just those which had been flagged up as being statistically abnormal in some way – would be able to explain their mortality rate by reference to their patient populations.

### **The Part to Be Played by the Healthcare Commission**

- 14.73 At the Inquiry's seminar, Professor Bevan and Dr Spiegelhalter spoke about the part which it was envisaged the Healthcare Commission might play in any future system of monitoring the mortality rates among GP patients. As I have explained, Professor Bevan was a member of the transition team which was working on the development of the future arrangements for the Healthcare Commission. Dr Spiegelhalter was an expert adviser to the Performance Assessment Subcommittee of the CHI; the CHI formerly carried out many of the functions which are now to be performed by the Healthcare Commission. Professor

Bevan said that it was expected that monitoring and surveillance of mortality rates, as well as of other indicators, would be carried out at a central analytical unit run by the Healthcare Commission. The results of the monitoring would be shared with PCTs and GP practices. Professor Bevan described how it was expected that the 'beating heart' of the Healthcare Commission would be a system of local offices. It was expected that these local offices would cover the areas of the SHAs. Professor Bevan said that discussions at the CHI had suggested that one of the key roles of the local Healthcare Commission offices would be to apply their local knowledge to the questions thrown up by the central analysis of data.

- 14.74 If the central analytical unit identified a statistical abnormality (whether in relation to mortality rates or any other indicator), it would inform the local Healthcare Commission office and the relevant PCT. The local office would examine any data (including clinical governance data) it held about the relevant GP or GP practice. The PCT would do the same. The findings would be shared. In the rare cases where no plausible explanation was found for the abnormality, the Healthcare Commission would conduct a targeted investigation. He suggested that its response would probably be tailored according to its assessment of the capacity of the individual PCT to carry out its own investigations. Dr Spiegelhalter thought that the usual pattern was likely to be that the routine monitoring of various indicators would result in a number of problem areas affecting a GP or GP practice arising simultaneously, giving rise to the need for investigation. However, it was possible that a single indicator (such as mortality rates) would exhibit striking variation and would of itself require investigation.
- 14.75 Dr Fox told the Inquiry that the description given by Professor Bevan and Dr Spiegelhalter was in accordance with the DoH's understanding of the likely arrangements for the Healthcare Commission.
- 14.76 Professor Bevan stressed that the primary function of the Healthcare Commission would be quality improvement. However if, in the course of its monitoring, the Commission were to find that something was seriously wrong, plainly it would have to act. He thought that the combination of monitoring and surveillance (see paragraph 14.34) would be 'quite a tricky balance to strike'. Dr Spiegelhalter hoped that the balance could be struck. He envisaged the DoH rapidly making mortality data available. The Healthcare Commission would then monitor the data, using a highly organised surveillance procedure. The results of that procedure would then be presented in a simple and informative way to PCTs and GP practices.
- 14.77 This was the position at the time of the Inquiry's seminars, in January 2004. In July 2004, the Inquiry wrote to the Healthcare Commission, asking for updated information about the role it was to play in the monitoring of GP patient mortality rates. In October 2004, the Healthcare Commission responded by saying the monitoring of the mortality rates was to **'be available'** through the DoH's National Programme for Information Technology, while routine assessment of individual GP rates would be the responsibility of the PCTs. It seems, therefore, that it is not at present intended that the Healthcare Commission should house the central analytical unit envisaged by Professor Bevan and Dr Spiegelhalter at the time of the seminars. It is not clear to me whether it is envisaged that the DoH would undertake the central analysis of mortality data or whether its function would, as Dr Fox

suggested be confined to the provision of linked mortality data. Nor is it clear whether the Healthcare Commission is to have the system of local offices described by Professor Bevan. In a recent letter to the Inquiry, the Healthcare Commission referred to its intention to **'manage relationships at local level'** and to **'engage with local stakeholders'**. It also referred to an intention to test a variety of ways of **'achieving the objectives for local presence'**. There was no mention of establishing local offices. Thus, there is uncertainty about which body would carry out any necessary investigation consequent upon the monitoring of mortality rates.

- 14.78 It is disappointing that the arrangements for the monitoring and analysis of mortality data that were envisaged at the time of the Inquiry seminar have not progressed. Indeed there is now uncertainty as to whether this function is to be carried out at all and, if so, by whom. It is also unclear who might be responsible for the investigation of the cause and significance of any data that lies outside the norm.

## Recent Work on Analysing and Investigating Mortality Rates

### The Pilot Project Undertaken by the Northern Ireland Eastern Health and Social Services Board

- 14.79 The Inquiry requested and received information about one pilot project that was being undertaken in Northern Ireland to develop a system for monitoring general practice mortality rates. The pilot project, which was being undertaken by a team from the Eastern Health and Social Services Board (the Eastern HSS Board), was conducted in association with the Department of Public Health and Epidemiology, University of Birmingham. At the Inquiry's seminar, Dr Kathryn Booth, Chair of the Northern Ireland GP Practice Mortality Regional Group, and Dr Mohammed, who assisted with the pilot project, gave further details of the project.
- 14.80 The Northern Ireland Central Service Agency established a database and provided each of the four Northern Ireland HSS Boards with mortality data for the GP practices in their area. Each HSS Board carried out its own pilot project. As I have said, the Inquiry heard about that conducted by the Eastern HSS Board.
- 14.81 The Eastern HSS Board team used Shewhart control charts to identify GP practices with adjusted mortality rates which were in excess of the upper control limit or below the lower control limit of the charts and were therefore exhibiting special cause variation. A longitudinal control chart was used to compare the mortality rate of four GP practices with their past mortality rates and with Northern Ireland average mortality rates over a period of eight years. Dr Mohammed explained that this latter technique avoided problems with case mix factors and showed changes in the pattern of mortality within an individual practice over time.
- 14.82 Dr Mohammed explained that, before the Eastern HSS Board pilot project started, there was great concern within the medical profession about the ultimate aim of the work. There was also concern about the possible publication of mortality data and about the way in which that data would be interpreted by patients and by the public. Dr Booth observed that, in particular, there was a fear that the data would be presented to the public in the

form of rankings. She said that another fear was that, if a practice was shown to be an outlier, members of the practice might not be allowed to contribute to any investigation into why its mortality rates fell 'outside the tramlines' and might have their names 'blackened' among their peers or with the public without having an opportunity to explain the apparent abnormality. The Eastern HSS Board team addressed the concerns of the profession by holding a series of seminars for GPs, patients' representatives and representatives from coroners' offices at the beginning of its pilot project. Members of the team sought to reassure those concerned that the object of the pilot project was primarily to improve the quality of health care. They had managed to secure support from GPs' representatives. Dr Booth said that some fears were still expressed in the course of practice visits conducted by those working on the pilot project, although both she and Dr Mohammed felt that progress had been made in persuading the profession of the value of the project.

- 14.83 If the mortality rate of a GP practice appeared to be exhibiting special cause variation, the Eastern HSS Board team used a pyramid model of investigation, which had been adapted from industry, in order to discover why this had occurred. The first stage of the investigation involved checking the existing data with the Central Service Agency and with the practice itself. A visit was arranged and, in advance of that visit, the practice was provided with all the data on which the team's analysis had been based. The practice was asked to carry out a check to validate the data.
- 14.84 The second stage of the investigation process was to check patient case mix. Members of the Eastern HSS Board team prepared information derived from the basic data about patient deaths and adjusted it for case mix factors such as age, gender and socio-economic deprivation. They would explain to the practice the adjustments they had made. Members of the practice, using their own local knowledge, would then have an opportunity to challenge those adjustments if they considered that they were inappropriate in some way. The team would invite members of the practice concerned to identify any characteristics of the practice, or of its patient population, that they considered might have had an effect on its mortality rates.
- 14.85 During a practice visit, members of the Eastern HSS Board team made use of a piece of software called a 'data mining tool', contained on a laptop computer. This enabled them to access a wealth of information, comparing the data for the practice year by year and also comparing the practice statistics with the Northern Ireland average. All this information was sent to the practice in advance of the visit. Before the visit, the team would analyse the data in an attempt to generate a hypothesis to explain the high (or low) mortality rates of the practice. The team would not disclose its hypothesis to the practice, since one of the purposes of the visit was to see whether the practice itself could explain the reasons for its special cause variation.
- 14.86 In each case where an investigation was undertaken, the team was satisfied by the end of the second stage of the investigation that the special cause variation was related to data and/or case mix factors. Dr Mohammed said that the team had not foreseen the impact of residence in nursing homes on mortality rates. This had become evident only in discussions with the practices concerned. He regarded local knowledge as the key to the investigation process. He gave two examples of this. In one case, a practice had had a

very low mortality rate for at least three of the years analysed; that mortality rate suddenly jumped up and then levelled off again. A preliminary investigation revealed that, during the first period of the analysis, the practice had been serving a university population, so that its patients were in general very young. The university then closed and the practice began to attract older patients, which explained the higher mortality rate. A second practice had a very high mortality rate which was decreasing over a period. Investigation of this revealed that the mortality rate was dropping because the practice had at one point had a very high proportion of patients who were resident in nursing homes but had subsequently stopped taking such patients. The proportion of patients who were resident in nursing homes had, therefore, dropped – hence the decrease in mortality rates.

- 14.87 In none of the cases which it investigated was it necessary for the Eastern HSS Board team to proceed beyond the second stage of the pyramid model of investigation. Had it proceeded to the third stage, this would have involved checking practice resources. It would have been necessary to discover from the practice and the PCT whether the practice had any specific problem with resources (e.g. poor access to cardiac surgery) which might have affected its mortality rate. If this third stage had failed to reveal an explanation for the excess, the team would have proceeded to the fourth stage of the pyramid model. This would have consisted of checking the process of care, i.e. looking at the processes of care adopted by the practice and by other agencies when treating particular groups of patients in the practice. It might, for example, have involved considering protocols in use in the practice for treating various patient groups. The fifth stage would have been reached if none of the preceding stages had provided a plausible explanation for the excess in mortality rates. It would have involved a formal investigation of the GP concerned.
- 14.88 Dr Mohammed said that the Eastern HSS Board's investigations were designed to generate knowledge in a systematic way. Once data was found to explain the special cause variation, the investigation was halted. As I have said, the explanations were generally dependent on local knowledge. Dr Mohammed said that, if a continuous monitoring system were implemented, those operating it would not necessarily wish to investigate every occasion when a unit crossed the control limit. If resources for investigation were limited, it might be appropriate to prioritise by investigating only those units whose chart statistics were most distant from the control limits.

### **The West Midlands Investigation**

- 14.89 As I have said, the work undertaken by Dr Aylin and his team revealed that, had their monitoring system been in operation at the relevant time, 12 GPs (including Shipman) would have signalled at the higher of the thresholds applied. Having been informed of that fact, the Inquiry decided that the DoH and the PCTs on whose lists the 11 other GPs were now included should be notified of the fact that they had signalled. That would enable the PCTs to undertake any investigations they considered appropriate in order to ascertain the reason for the apparent excess in mortality rates.
- 14.90 Two of the investigations arising from this notification were of particular interest to the Inquiry. The first was carried out in the West Midlands by a team from Shropshire County

and Telford and Wrekin PCTs, with technical assistance from Dr Mohammed (the West Midlands team). The investigation was reported subsequently in the *British Medical Journal*<sup>2</sup>.

- 14.91 The West Midlands team used the pyramid model of investigation which I have already described. Its first step was to check the data. The second was to check the patient case mix of the two GPs being investigated. Before starting its investigation, the West Midlands team informed the two GPs of what it was doing. The team then obtained the raw data used by Dr Aylin and his colleagues, together with the results of their analyses. The team also obtained additional information about where each death had occurred. By this means, they were able to see how many deaths had occurred in nursing homes. The team then formulated a preliminary working hypothesis that the excess of deaths might be caused by the fact that, at the relevant time, both GPs had a significant number of patients who were resident in nursing homes. Patients admitted to nursing homes are known to have a high mortality rate.
- 14.92 The West Midlands team held discussions with the two GPs and with the local primary care medical adviser, who had responsibility for GP performance. Members of the team examined administrative data from the practices of the two GPs. They compared the excess mortality rates for each GP with the number of deaths among their patients that had occurred in nursing homes. They carried out various analyses of the data. On the basis of their investigations, they concluded that the high mortality rates of the two GPs were attributable to a 'nursing home effect'. They judged that it was not necessary to proceed to the third stage of the pyramid model.

### **The West Sussex Investigation**

- 14.93 The second investigation was conducted by Adur, Arun and Worthing PCT into five GPs, one from its own area and four from other PCTs in the West Sussex area. The investigating team provided a copy of its report to the Inquiry and Ms Julie Billett, Associate Public Health Specialist, spoke about it at the Inquiry's seminar.
- 14.94 The investigation was undertaken without the knowledge of the GPs in question. Ms Billett said that this was because of the background to the investigation, the possibility of negative publicity and the concerns that might have been generated locally had the fact of the investigation become known. The first part of the investigation consisted of a statistical analysis, using local data to make more refined adjustments for age and gender than had been possible in Dr Aylin's analysis. The PCT's statistical analysis, which covered only the year in which each of the GPs first signalled, confirmed that the five GPs did indeed have statistically significant excess mortality rates when compared with the age and gender specific mortality rates for West Sussex in the same years. The second part of the investigation consisted of a review of the medical records for the patients whose deaths had been certified by each of the five GPs in the year that s/he signalled. In the case of two of the GPs, problems were experienced in identifying from the patient medical records

<sup>2</sup> Mohammed MA, Rathbone A, Myers P, Patel D, Onions H, Stevens A (2004) 'An investigation into general practitioners associated with high patient mortality flagged up through the Shipman Inquiry: retrospective analysis of routine data'. *British Medical Journal*, Vol 328: pp 1474–1477.

evidence of the medical condition which the GPs had certified as being the cause of the patient's death. Those problems were attributed to serious inadequacies of the patient medical records. Indeed, such were the inadequacies of the records that the reviewers suggested that the practices might have failed to include the most recent computerised records when returning the records to the PCT after the patients' deaths. An alternative explanation was that a parallel set of notes might have been held in the nursing or residential homes in which some of the patients had resided. As a result of the gaps in the medical records, there were some concerns about the GPs' management of chronic disease. It was not established in the course of the investigation whether any other records did in fact exist. Subsequently, one of the GPs explained in a letter that his practice had been paperless for some years and that he had **'felt it unnecessary to produce a full paper discharge summary from the computer for a deceased patient'**.

- 14.95 At the conclusion of the investigation, the PCT decided that the likeliest explanation for the GPs' higher than expected mortality rates was that a significant proportion of their patient populations were nursing or residential home residents, with correspondingly high rates of mortality. At that stage, the fact that they had been the subject of an investigation was revealed to the GPs. It was left to the Directors of Public Health and the Clinical Governance Leads for the relevant PCTs to take up with the doctors the issues about record keeping and chronic disease management which had been identified during the course of the investigation.

## Future Investigations

- 14.96 Adur, Arun and Worthing PCT had to devise the whole process of investigation for itself. Ms Billett said that one of the difficulties was knowing where the investigation should stop. She told the Inquiry that it would be of great assistance if there was some clear guidance available to PCTs on a structure and a process to be followed when conducting investigations. There was general agreement at the seminar that such guidance was required. Professor Baker supported the development of a structured protocol. He suggested that the potential outcomes of an investigation might be, first, that there was a satisfactory clinical explanation for the apparently abnormal findings. That outcome would require no further action. The second possible outcome might be that there was a problem with the performance of a GP or GP practice; that might result in the provision of support or remediation or in the use of disciplinary procedures. The third possible outcome might be the identification of possible criminal behaviour. Professor Baker suggested that there must be clarity about the point which an investigation had reached and why it had reached that point. He suggested that the police should have some input into the formulation of a protocol. They should advise at what point they might wish to become involved in an investigation. Dr Grenville, representing the BMA, suggested that the medical defence organisations should also be consulted in connection with the formulation of an investigation protocol.
- 14.97 Dr Fox, for the DoH, suggested that there were two different types of investigation that might result from monitoring mortality rates. The first would occur in the context of local routine prospective monitoring when a PCT or a GP practice was looking at data in a 'learning, improving mode'. This type of investigation would be very different from that

which would occur when a PCT had been alerted to the fact that there might be a 'problem' with a practice. Dr Fox suggested that different guidelines were needed to deal with those two types of investigation. He emphasised the importance, when a formal investigation was to be undertaken, of identifying the question which it was intended that the investigation should answer. Without that, it would be impossible to know when the question had been answered and when the investigation should be halted. Dr Fox anticipated that formal investigations would be undertaken fairly rarely. He was concerned that individuals would be involved in such investigations infrequently and would not be in a position to gain any experience of how to handle them. He suggested that it would be important to develop a bank of knowledge about how such investigations should be carried out.

- 14.98 Doubts were expressed about whether PCTs, which are small organisations with many other responsibilities, would have the necessary resources and expertise to conduct the investigations which would become necessary as a result of GPs signalling within a monitoring system. Ms Billett felt that the task of investigation should be undertaken by PCTs, which could bring to the task their understanding of the local factors that could influence the mortality rates of a GP practice or of an individual GP. However, she doubted whether all PCTs would have the capacity to undertake the task. The exercise had been useful for her PCT, in that it had identified problems with two GPs and had highlighted the importance of preserving the quality and integrity of the PCT's historical data systems. However, it had occupied a considerable amount of time and resources.
- 14.99 Professor Baker believed that an investigation of this complexity was likely to be beyond the ability of most PCTs. He accepted the value of local knowledge but felt that there was a danger of becoming focussed solely on a local explanation, rather than adopting a wider and more objective view. He suggested that local input should be blended with accumulated regional or national experience, so as to ensure that expertise and objectivity were brought to investigations. Dr Grenville referred to the high turnover of staff in PCTs and suggested that it was likely to be difficult for individual PCTs to build up the pool of expertise that would be required to undertake such investigations.
- 14.100 The alternative to a local investigation would be investigation by an external body. The most likely candidate appeared to be the Healthcare Commission. Ms Billett said that an investigation by an external body was likely to be perceived by GPs as more threatening. However, how such an investigation was received would depend on the way in which it was conducted. Both she and Mrs Scott agreed that their PCTs would want to have an input into any such investigation which affected a GP in their area.
- 14.101 In my view, it is important that real expertise and objectivity should be brought to such investigations. I recognise the need for PCTs to be involved, so that use may be made of their local knowledge. However, I do not think it would be appropriate, either in principle or in practice, for PCTs to have primary responsibility for a formal investigation into a doctor who had signalled within a monitoring system. As Professor Baker said, not only would they not have the necessary skills or sufficient opportunity to develop them, they might also lack objectivity. There would be a danger that local knowledge and reputation might be given too great a prominence and might be seen to provide an innocent



explanation for figures which should give rise to real concern. I express no view as to who should be responsible for the conduct of such investigations. I do not know how frequently they will be required. If it transpires that they are rare, as Dr Fox anticipates, a small national team should suffice, possibly organised by the Healthcare Commission. If, on the other hand, it is found that such investigations are required quite frequently, it may well be necessary to provide a facility on a regional basis. If so, those responsible in the regions must work to common methods and protocols. The existence of national or regional expertise would also provide a source of advice for PCTs who came across the possibility of a problem when collecting data in the 'learning, improving mode' mentioned by Dr Fox.

### **After an Investigation**

14.102 Dr Spiegelhalter raised the issue of what would happen after an investigation had discovered an acceptable clinical explanation for excess mortality. If monitoring continued and the high mortality rate persisted, there would at some point be another signal. A decision would then have to be taken as to whether to investigate again (which could well be a waste of resources and upsetting for the doctor concerned) or to assume that the explanation previously identified satisfactorily accounted for the continued excess (which might be a false assumption). He said that decisions would have to be taken about the approach to be adopted in such circumstances.

14.103 I agree that careful consideration would have to be given to such a case. I would have thought that the right approach would be for a second (and possibly more experienced) investigator to review the first investigation, in the light of the fact that the high mortality rate appeared to have persisted. If s/he were satisfied with the methodology and conclusion of the first investigation and considered that the explanation then accepted also accounted for the continuance of the high mortality rate, it should not be necessary to embark upon a second investigation. However, a second investigation would have to take place if any concern remained after the review. In any event, such a review should not be undertaken locally, for the reasons I have already given.

### **Potential Problems with Monitoring**

14.104 At the seminar, there was discussion about some of the problems that might be associated with monitoring mortality rates.

### **The Practicability of Monitoring at Individual General Practitioner Level**

14.105 During the whole of the period covered by the analysis carried out by Dr Aylin and his team, Shipman (and, subsequently, the locum who replaced him) was practising single-handed. The analysis showed that he would have signalled during that period. However, the question then arose whether, if he had been a member of a larger GP practice (as he was until 1992), his excess deaths would have caused the practice to signal.

14.106 Dr Aylin told the Inquiry that, if excess deaths caused by one doctor were to be subsumed within a large GP practice, they would inevitably become much more difficult to detect. He

said that it would be possible to tune the CUSUM chart to pick up differences in mortality rates within a practice caused by a single doctor. However, if the alarm threshold were to be set at a level that would pick up such differences, this would result in a large number of statistical false alarms. Dr Aylin said that, with better data and with better adjustment for case mix and for unaccounted-for factors than had been possible in the statistical analysis carried out by himself and his team, it might be possible to pick up, in the mortality data for a practice of six doctors, an excess of 25 deaths caused annually by a single GP. This would be a large excess; Shipman had an excess of 25 deaths during only three years of his practice.

- 14.107 Dr Aylin pointed out that, if avoidable deaths were being caused by practice systems which were affecting more than one doctor, an excess mortality rate might show up more readily. He suggested that, from the point of view of quality of care delivered to patients, a GP practice was a reasonable unit of analysis. However, in terms of detecting excess mortality attributable to the actions of a single GP within the practice, it is plain that it is a blunt instrument. While I agree with Dr Aylin that monitoring of mortality rates at GP practice level may have some value, it seems to me that, in order to be sufficiently sensitive to detect an excess of deaths caused by the activities of a single GP, monitoring must be at individual GP level. Monitoring at individual GP – rather than GP practice – level would also enable a GP to be tracked throughout his/her career. This is important because of the increasing tendency of GPs to move between practices.
- 14.108 If monitoring at individual GP level is to be undertaken, two conditions must be met. First, there is a need to link the actual care of a patient to a named GP, so that, when the patient dies, the death can sensibly be counted as being ‘attributable’ to that GP for monitoring purposes. A merely administrative allocation of patients to a particular GP would be valueless. Second, it must be possible to calculate a total patient population for each GP. This could be done by making a link between named patients and that GP, or by allocating a proportion of the practice patient population to the GP by reference to information such as the number of sessions worked by the GP.
- 14.109 Recent changes to the arrangements for general practice have made it difficult to link a patient to an individual GP. Under the new General Medical Services Contract (and under the personal medical services contracts which have recently been introduced), patients are now to be registered with a GP practice and not with an individual GP. Thus it will not be possible, on the basis of the available data, to link a deceased patient with an individual GP for monitoring purposes. Even before the recent changes, the link between a patient and the doctor with whom s/he was registered might have been misleading, since patients might have been registered with one doctor, but their care might have been shared by other doctors within the practice. This type of shared care has become more common in recent years.
- 14.110 Professor Baker told the Inquiry that he could envisage the possibility of requiring some allocation of patients by a GP practice so that there would be a named GP for each individual patient. However, he felt that this might be ‘asking rather a lot’ of some practices and patients. He believed that it should be possible to estimate with some confidence the proportion of patients within a practice that an individual GP was caring for, by reference

to the patterns of work of members of GPs within the practice. In the longer term, he suggested that improved electronic information systems (including electronic patient records) might enable the number of patients being cared for by an individual GP to be measured directly. He pointed out that, even now, almost all consultations between a patient and a healthcare professional within a practice are recorded electronically. If anonymised data were made available, it would be possible to determine the proportion of consultations between an individual patient and each GP within the practice and to allocate patients to named GPs on the basis of those proportions. It would also be possible to obtain information about the activities of locums who had been employed within the practice. Professor Baker pointed out that this system would also enable precise adjustments to be made to reflect the age and gender of patients allocated to a particular GP. Dr Aylin endorsed Professor Baker's idea of focussing on the number of consultations between a patient and an individual GP in order to determine that GP's patient population.

- 14.111 Some potential drawbacks to Professor Baker's approach were mentioned at the seminar. Dr Goldblatt believed that, if information about patient consultations had to be collected before mortality data could be analysed, this would introduce a delay into the monitoring system. Dr Fox was concerned that the provision of information about patient consultations would impose an undue burden on GP practices at a time when the DoH is attempting to reduce that burden.
- 14.112 Professor Bevan observed that, if the information about patient consultations was being collected for the purpose of monitoring mortality rates with the sole intention of 'spotting another Shipman', it would be difficult to justify. However, the information might have a broader purpose to serve in the context of monitoring and of improvements in the quality of care. He said that there had been concern within the CHI about the 'massive gap' in the statistical information available about general practice. This was in contrast to the amount of information available in relation to hospital care. He pointed out that 90% of patient contact takes place at general practice level. He said that there might be the potential for setting up databases to collect the relevant data as part of a wider move to improve the quality of care in general practice.
- 14.113 Professor Baker mentioned that there were in existence at least two schemes whereby groups of practices operating similar computer systems had undertaken to collect and share data relating to a range of issues. He said that those schemes enabled searches to be conducted of databases containing information about contacts of up to a million patients with doctors and nurses. Dr Baker told the Inquiry about a number of projects being undertaken by the RCGP. She said that work was being done on a method by which data could be automatically extracted from a GP practice's computer system without any action by the practice being necessary.
- 14.114 If there is to be effective monitoring of mortality rates at individual GP level, some way must be found to make a real linkage between an individual patient registered with a practice and a named GP within that practice. That must be based upon the care actually provided and not on any administrative factor. I agree with Professor Baker that it should be possible to do this by reference to actual patient contact. There will, of course, be many young, healthy patients who rarely consult their doctors and who may have to be notionally

allocated to one member of the practice. However, allocation would have to be reviewed periodically so that, if a patient began to require regular care, s/he could be reallocated if necessary to the doctor providing it. I should have thought that it would be possible to do that. I should also have thought that it would be helpful to a GP's understanding of his/her own practice to have access to data about the patients whom s/he treats and their characteristics. From what Professor Bevan told the Inquiry, it appears that the data might have a wider use than just in the monitoring of mortality rates. It seems to me also that, with the advances in technology about which the Inquiry has heard, it should be possible to collect the necessary data without placing an undue burden on practices. I hope that such measures will be explored because I am firmly of the view that analysis of the collective mortality statistics of a group practice will be a very blunt instrument indeed in detecting either poor clinical practice or criminality.

- 14.115 Over recent years, there has been an increase in the number of locums providing care. There has also been an increase in the use of out of hours services. This will continue with the changes to the arrangements to the provision of out of hours services due in January 2005. At the seminar, Dr Aylin was asked whether he had considered how the mortality rates of locum GPs and out of hours services might be monitored. He explained that the difficulties lay, not in the statistical issues, but in collecting the necessary data linking a death with a particular locum or out of hours agency, and in determining the total number of patients cared for by that locum or agency. He suggested that it might be possible to monitor the mortality rates of patients being cared for by an out of hours agency. The patient population being cared for by such an agency could be determined by reference to the GP practices which used the agency. If a monitoring system were to be introduced, it would be necessary to explore the different possibilities for collecting this data.

### **Where to Set the Threshold**

- 14.116 Another problem which was identified at the seminar was knowing where to set the alarm threshold, so as not to cause a large number of false alarms.
- 14.117 Dr Aylin said that he and his team had tried not to be 'too explicit' about where the threshold should be set. He said that one strength of the CUSUM methods was that the false and successful discovery rates could be calculated so that those setting the alarm threshold would know how the chart was likely to perform. That meant that the thresholds were not chosen arbitrarily. Dr Mohammed pointed out that, under Dr Aylin's proposed method, the calculation of the false and successful discovery rates depended on an assumption that the mortality rates of 5% of the GPs and GP practices being examined are out of control. He observed that, in reality, we do not know how many units are out of control. He suggested that the calculations were useful for the purposes of illustration but could not resolve the fundamental problem that the proportion of GPs with an excess mortality rate is not known. Dr Goldblatt made a similar point. Dr Best and Dr Aylin agreed that statistical calculations should not dictate where the thresholds were set. They should merely inform the selection of thresholds by others. It would be necessary to take into account practical considerations relating to resources and to the types and number of investigations which it was thought appropriate to carry out. Dr Aylin observed that the

number of true and false medical alarms signalling under the monitoring system would inform, on an ongoing basis, the decision as to where to set the alarm thresholds.

### **When Would Shipman Have Signalled?**

- 14.118 Some concern had been expressed before the seminar about the fact that, on the basis of the statistical analysis conducted by Dr Aylin and his team, it appeared that Shipman would have signalled only in 1996. The time lag in collecting the relevant data would have meant that the signal could not have been detected until late 1997. This would not have been long before March 1998 when the late Dr Linda Reynolds reported her suspicions about Shipman. Her report gave rise to the first – and unsuccessful – police investigation. Doubts were expressed about the value of a monitoring system which would have taken so long to detect Shipman.
- 14.119 When asked about this, Dr Aylin pointed out that the period covered by his analysis had started only in 1993. Had the analysis been in operation for the whole of the period when Shipman was in practice, he could have been expected to signal much earlier. Also, the poor quality of the data and the limited adjustment for case mix factors made it less easy to detect an unusual pattern of deaths. It seems to me that, with improved data quality, finer adjustment for case mix factors and the use of the cumulative method, a serial killer or a doctor who habitually neglected his/her patients would signal within a reasonable time. However, it must be recognised that a doctor who killed a patient only occasionally, or who hastened the deaths of patients by only a short period, would not be detected by such a monitoring system.
- 14.120 At the presentation of their work to the Inquiry by Dr Aylin and his team in July 2003, Dr Aylin suggested that, statistically, every GP would signal at one time or another under his system. This suggestion caused a considerable amount of concern. At the seminar, he explained that his comment had been made in the context of monitoring for an infinite time period in the future. He had calculated that, using the thresholds employed in his analysis, 1.5% of GPs might signal purely by chance over a 50 year period. It would, therefore, take several thousand years for every GP being monitored to signal as a result of chance alone.

### **The Effect on Mortality Rates of Residence in a Care Home**

- 14.121 The outcomes of the investigations in Northern Ireland, the West Midlands and West Sussex that I have described all highlighted the relationship between the mortality rates among patients of a GP or GP practice and the number of those patients who are resident in nursing homes. This effect may extend to patients living in residential homes also. The fact that a patient is resident in a nursing home suggests that s/he is in a poorer state of health and has a higher level of dependency than a patient in a residential home. The Inquiry was told, however, that this is not always the case; sometimes, practical considerations govern the admission of a patient to one type of home or another. For the purposes of this discussion, I shall refer to nursing and residential homes collectively as 'care homes' and assume that residence in either has an effect on mortality rates.

- 14.122 If the pattern of the three investigations is followed, it seems likely that many statistical true alarms will be accounted for by the fact that a higher than average proportion of the patient population of the GP under investigation is resident in care homes. Investigation of these alarms will, therefore, inevitably result in the unnecessary expenditure of resources.
- 14.123 One way of avoiding this would be to make a case mix adjustment in order to account for the number of patients of a GP who are resident in care homes. There are at least three problems with that approach. One problem is that adjustment might mask the existence of a high mortality rate of which the authorities should be aware. It is possible that a doctor might be responsible – through lack of care or criminal activity – for the deaths of patients in a care home. If an adjustment were to be made for care home residents, an excess mortality rate from that cause might not be detected. The second problem is that admission to a care home can, to an extent at least, be influenced by a GP. As I have said, there are dangers in adjusting for a case mix factor that is within the control of the subject of the monitoring. The third problem relates to the collection of the necessary data to enable an accurate adjustment to be made. Dr Farhang Tahzib, Director of Public Health, Adur, Arun and Worthing PCT, suggested that the fact that a patient had been admitted to a care home should be recorded on the NHS central data system. He suggested that this might be done by linking information from care homes and from the National Care Standards Commission (now part of the Healthcare Commission) with the NHS system. Such information would enable the number of patients registered with every GP who were resident in care homes to be ascertained. Ms Billett suggested that the information should also be sent to GP practices, which may be unaware that one of their patients had been admitted to a care home.
- 14.124 In theory, this seemed to be a good idea. However, Dr Goldblatt doubted whether it would be possible to collect the data in the way suggested by Dr Tahzib. He said that the designation of care homes as nursing or as residential homes could change quite frequently and suddenly, so that it would be difficult to keep data up to date. Also, patients were admitted to care homes for different periods of time. Some were admitted for short periods of respite care. Others were admitted only a short time before their death. Some patients moved between homes in different PCT areas. Dr Grenville confirmed these points and added that some hospital wards were very much akin to nursing homes, although they would not be categorised as such.
- 14.125 Mrs Scott, of Croydon PCT, said that, when preparing comparative data on GP patient mortality rates, her PCT had excluded deaths which had occurred in care homes. It was felt to be inappropriate to compare practices which had large patient populations living in care homes with practices which had no or few care home residents. The PCT had good information about the number of care home beds and about which GPs cared for patients in the different homes. The PCT was able to identify deaths which occurred in care homes. It was considering setting up a system for analysing deaths in care homes. Mrs Scott said that the main difficulty was the frequent change in the designation of care homes – even of individual beds within care homes. She said that it would require a lot of local knowledge and careful monitoring to deal with these changes.
- 14.126 Dr Aylin told the Inquiry that it would never be possible to adjust for all aspects of case mix. He said that the fact that it would not be possible to adjust for every single eventuality did

not prevent an analysis being performed on the basis of the available data. He suggested that there might be a case for performing analyses with and without adjustment for residents in care homes.

14.127 I should have thought that it would be feasible to collect data about persons admitted to care homes. It may be possible to identify those admitted on a wholly temporary basis, e.g. for the purpose of convalescence or respite care. I should have thought also that it would be helpful to GP practices and to PCTs to have information about the number of patients resident in care homes and also in other institutions such as hospices. Ways of analysing and comparing the mortality rates of residents in care homes and other types of institution could no doubt be devised. A first step would be to investigate with the Healthcare Commission precisely what information it holds and how up to date it is.

### **Other Possible Statistical Approaches**

14.128 At the seminar, there was some discussion about different statistical methods that might be used in the analysis of mortality rates and that might have advantages over the use of the CUSUM charts advocated by Dr Aylin and his team. I shall deal with these very briefly.

14.129 The first discussion related to the relative merits of Shewhart control charts and CUSUM charts. Dr Aylin's main reason for favouring CUSUM charts was, as I have said, their ability to accumulate historical data. He also pointed to their efficiency in determining gradual, sustained rises in mortality and large and sudden deviations in mortality. Dr Mohammed, who favoured the use of Shewhart charts, pointed out that the longitudinal control charts showing the mortality rates of a single practice revealed evidence of the historical performance of that practice. He referred to the transparency of Shewhart control charts and the ease with which they could be understood.

14.130 Dr Roberts' view was that the fact that CUSUM charts used information from the past meant that they were better predictors of what would happen in the future. Dr Spiegelhalter also favoured the use of CUSUM charts for use in a central monitoring system. He agreed with Dr Aylin about the importance of accumulating historical data. He also favoured the more formal technique for setting alarm thresholds advocated by Dr Aylin and his team, rather than what he described as the 'rule of thumb' used to set the control limits for Shewhart charts. However, Dr Spiegelhalter favoured a slightly different approach to the use of historical data.

14.131 When plotting their annual chart statistic, Dr Aylin and his colleagues simply added up the outcomes from previous years and kept adding year by year. The effect of this was to give equal weight to historical and recent observations. An alternative approach would be to adjust previous outcomes so as to give more weight to more recent data and less weight to data in previous years. Dr Aylin and his team did not attempt to do this, since it was not clear by how much it would be appropriate to downweight information from the past.

14.132 Dr Spiegelhalter said that his preference was to discount historical data gradually. The effect of that would be to accord most weight to the most recent information and to give less weight as data became older. He advocated a gradual discounting, so as not to precipitate a sudden change by the removal of historical data.

- 14.133 Dr Spiegelhalter emphasised that a monitoring system need not be confined to using one statistical method. Different methods could be used for different purposes and to present the data to different groups of people. He suggested that the system devised by Dr Aylin and his team would be the most suitable for use as a quality assurance tool. It could be used centrally for monitoring large numbers of units and could provide warnings which would then be passed down to those who were to investigate them. In other words, the system could be used for surveillance purposes.
- 14.134 Dr Spiegelhalter suggested that, when the data was supplied to PCTs and practices, it should be presented clearly and simply. This might be best done by means of Shewhart charts. Dr Spiegelhalter said that it would be possible to move between the two statistical methods and thus to have multiple views of the data. No extra resources would be required. Dr Aylin agreed that CUSUM charts were not appropriate for feeding back direct to GPs. Other methods of presenting the information would have to be adopted.
- 14.135 Another issue arose in relation to the treatment on a CUSUM chart of a GP who, for a period, has a low mortality rate. The CUSUM chart is set at zero, so that if the mortality rate of a GP is less than that which would be expected, the chart statistic is plotted at zero. It is not allowed to go below zero and to build up credit by registering negative values. This is in distinction to the sequential probability ratio method for which Dr Spiegelhalter had originally expressed a preference. Under that method, the chart statistic builds up 'credit'. The danger is that it might build up too much credit so that, if the mortality rate of a practice increased, it would take longer for the practice to signal, because the deterioration would to some extent be compensated for by the previous good performance. Dr Spiegelhalter said that, when using the sequential probability ratio method, he had found that the building up of credit was a real problem and led to a lack of sensitivity. While the sequential probability ratio method had some advantages in his view, he now felt that the CUSUM chart was preferable. He also favoured the approach advocated by Dr Aylin and his team of using computer simulations and actual past data to work out the false and true discovery rates and of then leaving the decision as to where the threshold should be set to those responsible for running the monitoring system.
- 14.136 Dr Spiegelhalter suggested that there should be a pilot scheme to design the system, to set the thresholds and to fine-tune the way it works. He suggested that the pilot scheme should use data that resembled as closely as possible the data that would eventually be used. Once the system was implemented, it would be necessary to monitor its use closely and to adjust it as necessary.
- 14.137 Dr Spiegelhalter said that he would be very concerned at the suggestion, which had been put forward by some, that the monitoring of mortality rates should be left to visual comparisons of crude mortality data by PCTs. He spoke of his experience in the Bristol Royal Infirmary Inquiry and said that the hospital involved had been 'awash with data', with plenty of opportunity for visual comparison, but nothing had been done about the information shown by the data. A formal system was, in his view, required. Professor Baker expressed the view that a visual comparison would not be sufficiently sensitive or reliable. It was likely, he thought, to miss some information.



## The Concerns of the Profession

- 14.138 A number of participants in the seminar suggested that many GPs find the idea that the mortality rates of their patients might be monitored very threatening. Dr Grenville, representing the BMA, made the point that, if 'entirely blameless' GPs were to be over-investigated, this would have an adverse effect on morale and, ultimately, on patient care. I have already mentioned the concerns of members of the medical profession about the work that has been undertaken in Northern Ireland.
- 14.139 Dr Booth mentioned that there had been particular concern in Northern Ireland that mortality data would be presented to the public in the form of rankings or league tables. There was general agreement at the seminar that the publication of rankings of GPs by mortality rates would be inappropriate and unhelpful. Dr Aylin pointed out that simple ranking can be dangerous. It inevitably means that one practice is at the bottom of the league table and is labelled 'worst'. It would be necessary to have some understanding of the distribution of mortality between GPs, in order to detect whether a practice was a true outlier, or whether it represented merely the tail of a normal distribution. Professor Baker and Dr Spiegelhalter agreed that rankings would be unhelpful. Dr Spiegelhalter said that there were other means by which divergent cases could be clearly spotlighted.
- 14.140 Dr Aylin and his team did, however, believe that rankings of mortality rates for GPs in a PCT area – appropriately adjusted for case mix factors – could have a use in communicating comparative information about mortality data to GPs, provided the rankings were presented in an appropriate way. Typically this would be done graphically by presenting individual GPs' mortality rates with a range of values (known as confidence intervals) within which their mortality rates were likely to vary.
- 14.141 Rankings of this kind are already produced in some areas. Mrs Scott told the Inquiry that her PCT provides comparative information about 60 or 70 performance indicators to GP practices in its area on an annual basis. For the last few years, mortality data has been one of the indicators. A practice's indicators are discussed at annual clinical governance visits.
- 14.142 I can well understand why GPs would feel that the publication of rankings of GPs' mortality rates might be misleading, unfair and damaging to doctors' morale and patient care. In my view, it should not happen. The data should be presented to GPs and GP practices as in a scientific study, with an explanation of the methodology and its limitations and the confidence intervals that apply. Then, assuming that GPs are able to understand and interpret the material in that form, it should be of real value to them for their clinical governance purposes at both an individual and a practice level.

## The Potential Advantages of a Monitoring System

- 14.143 Since the issue of monitoring mortality rates has gained prominence in the wake of Shipman's activities, it is perhaps not surprising that there is a perception that the only purpose of such monitoring is to 'catch another Shipman'. However, those who are in favour of setting up a system of monitoring argue that it would have benefits extending well beyond the detection – and possible deterrence – of a murderous doctor.

- 14.144 There was agreement among participants at the seminar that the process of monitoring mortality data could have two distinct and separate uses. The first of these was quality improvement. Professor Baker believed that, if mortality data were available and GP practices had the necessary expertise to understand the data and to learn from it, it should be possible to use the data for local quality improvement purposes. The second use of monitoring would be for quality assurance purposes. This would involve the use of the data to observe abnormalities which could indicate a performance problem, or even criminal activity.
- 14.145 Dr Mohammed could not give any specific examples of improvements in quality of care that had occurred as a result of the Eastern HSS Board project, although, of course, it was early days. He suggested, however, that monitoring would assist in understanding the factors that contribute to an excess in mortality. This might result in an improvement in the quality of care and might potentially also identify criminal behaviour. He would like a monitoring system to look at both high and low mortality rates. He believed that this might lead to identification of good – as well as poor – practice and might make the process less threatening to the medical profession. Dr Aylin confirmed that it would be possible, using CUSUM charts, to look at both high and low mortality rates.
- 14.146 Dr Grenville observed that the monitoring of mortality rates could serve a dual purpose. It could be a 'powerful tool for improving practice'. At the same time, it might also be a tool for detecting 'another Shipman' should one appear. Dr Grenville felt that quality improvement should be the primary aim, with the detection of criminality as a 'by-product'. He wanted to ensure that monitoring provided 'the most benefit for the most people'. He felt that it would do this by improving practice. He cited as an example the situation which might arise if monitoring showed that a practice had an excess of deaths due to cardiovascular disease. He said that that information might lead to consideration of changes which should be made to primary and secondary strategies for the prevention of cardiovascular disease within the practice and within the area. If, in the process of using mortality data to improve the lot of patients, it was also possible to detect poor practice or instances (which he believed would be very rare) of criminal practice, Dr Grenville said that he would be delighted. His concern about setting up a monitoring system was that, once it was in operation, it might encourage a false belief that it would inevitably catch 'the next Shipman'. He did not think that it would necessarily do so, since, if a doctor was minded in the future to act in the same way as Shipman, s/he was unlikely to do it in the same setting of general practice. However, with that proviso, Dr Grenville could see real benefits in monitoring mortality rates.
- 14.147 Professor Baker felt that there was real value in GP practices receiving their own mortality data and having available the necessary information to enable them to compare their own data with that of other practices. He suggested that if, for example, a GP practice found that it was having more cardiovascular deaths than could be expected from the figures for other practices, this could be 'a good trigger' for them to look at how they were treating patients with risk factors for cardiovascular disease. Dr Baker, representing the RCGP, agreed with Professor Baker about the benefits of practices having an understanding of their own mortality rates. She believed that a system of monitoring mortality rates might act as a deterrent to a potential criminal. However, she also expressed the view that such a

system would not necessarily be effective in detecting the activities of locums or doctors engaged in out of hours work and would not, therefore, necessarily detect criminality on the part of a doctor working in those fields.

- 14.148 Some reservations were expressed about the extent to which the behaviour of GPs had the potential to influence the mortality rates of their patients. Dr Goldblatt pointed out that many diseases and causes of mortality are closely linked to socio-economic and other factors (e.g. the treatment received by a patient in hospital or the care received in a nursing home or the compliance of the patient with treatment) which are outside the control of the patient's GP. Some conditions are not amenable to treatment at all. It seems to me likely also that some deaths may have contributory factors that lie many years in the past, before the deceased became a patient at the GP practice with which s/he was registered at the time of his/her death.
- 14.149 Dr Goldblatt made the point that each individual GP has only a few deaths per thousand patients annually and only a relatively small proportion of the conditions causing those deaths are amenable to treatment or to intervention by the GP. It is only that small proportion of annual deaths which can be related to GP performance. Dr Goldblatt said that he believed that the monitoring of mortality rates among GP patients was of value, but it had to be recognised that a high level of mortality within a practice might be caused, not by the quality of care within that practice, but by some extraneous factor. He believed that the monitoring of general practice had to be complemented by the monitoring of other institutions and organisations.
- 14.150 Dr Roberts expressed some doubt, based on his experience of clinical trials in primary care, about the extent to which a variation in GPs' behaviour (e.g. by improving their practice) was capable of influencing the outcome for patients. He acknowledged that it might be useful for a GP to look at the deaths in his/her practice and, having seen for example that the practice had a number of deaths from substance misuse, to consider changing its system of dealing with substance misuse. However, he pointed out that that was not 'a statistical observation', by which I think he meant that it was not necessary to conduct any analysis of mortality rates in order to make the observation. Nevertheless, I suppose that an examination of mortality data might cause a practice to focus on deaths which had occurred and on the causes of those deaths in a way that it might not otherwise do. That was, as I understand it, the point that was being made by Professor Baker and by Dr Baker.
- 14.151 Professor Baker said at the end of the seminar that he felt optimistic about the capability of a monitoring system to combine both the quality improvement and quality assurance roles. He firmly believed that the monitoring of information about mortality was likely to be important to local quality improvement activities. He observed that quality could be measured only if there were several different ways of measuring it. This is because quality of health care is such a complex and confusing concept. Mortality, although not a very sensitive measure of quality, is, he said, an important one. It should be part of any overall scheme to assess quality. Professor Baker believed, from what he had heard at the seminar, that it would be possible to have a system which could detect doctors with high and clinically unexplainable mortality rates and could also help doctors and primary

healthcare teams to consider the mortality patterns among their population and to plan their key policies. He went on to say:

**‘I do not think it is possible for us to tell the public, let alone relatives of Shipman’s victims, that we are unable to detect and will be unable to detect in the future a doctor who murders a large number of patients. I do not think that is sustainable.’**

## Conclusions

- 14.152 Several participants in the Inquiry’s seminars observed that if, in the future, there were another doctor with a criminal intent similar to that of Shipman, s/he was unlikely to pursue his/her criminal behaviour in the same way as Shipman did. That being the case, it was said there was no guarantee that a system of routine monitoring of mortality rates among GPs’ patients would ‘catch another Shipman’.
- 14.153 I recognise, of course, that a system of routine monitoring of mortality rates would not, on its own, provide any guarantee that patients would be protected against a homicidal doctor. I also agree with those who have emphasised the importance, if a system of routine monitoring is to be introduced, of ensuring that PCTs, the medical profession and the public are not lulled into a false sense of security, whereby they believe that the system of monitoring itself will afford adequate protection for patients. Routine monitoring of mortality statistics would be only one element of that protection. I have already made detailed recommendations for the reform of the systems of death certification and investigation, which I believe would provide significant protection. Indeed, it is my view that the introduction of those reforms, coupled with a system of monitoring of mortality statistics would together provide a real deterrent to misconduct, as well as a greatly improved opportunity to detect the activity of a doctor such as Shipman. For example, if the results of monitoring had been available when Dr Reynolds made her report to the Coroner, it would have been possible to ascertain immediately whether Shipman had an abnormally high mortality rate. If the information that I have suggested should be collected for death certification had also been available, it would have then been possible to pick up the abnormalities in the pattern of deaths among his patients even if they had not already come to light as a result of routine monitoring. In short, it seems to me that, if routine monitoring of mortality rates were shown to be workable, its introduction must be seriously considered.
- 14.154 Of course, the case for routine monitoring would be even stronger if it were shown that the monitoring of mortality rates would have benefits over and above the possibility of detecting aberrant or criminal behaviour. Quite apart from any other consideration, if GPs were able to see that the data derived from monitoring mortality rates was making a positive contribution to the quality of patient care within their own practices – and possibly more widely – the process of monitoring would appear less threatening than might otherwise be the case. It seems to me, as a lay person, that Professor Baker and others must be right when they say that it is important for GPs and GP practices to have access to, and to be able to interpret and understand, their mortality data. Of course, mortality data is not the only – or necessarily the strongest – indicator of the quality of patient care.

Nevertheless, it is an important indicator and must in my view form part of the data which should be collected in the course of clinical governance.

- 14.155 I recognise that there are practical difficulties in the way of setting up a satisfactory system of monitoring mortality statistics. The greatest of these seems to be the problem of providing a reliable method of linking an individual patient with an individual GP. My suspicion is that, unless that can be done, a monitoring system might be too blunt a tool to be worthwhile. However, I am optimistic that, with a will, it could be done. I agree with Professor Baker that it is simply not acceptable to say to the public – and particularly to the families of Shipman’s victims – that it cannot be done or that it is not worth doing. In my view, Dr Aylin and his team have shown that it can be done and I am satisfied, having heard the views of the seminar participants, that it will be worth doing.
- 14.156 In my view, the analysis of mortality data is important. The value of this process was acknowledged by all participants at the seminar, although the difficulties which lie ahead were recognised. There was also agreement that the work must be done centrally. In my view, the DoH must now make provision for it to be done. At the seminar there was also agreement that the investigation of the cause and significance of outlying data should be undertaken locally but that it required considerable expertise. Such investigations must be well organised, consistent and objective. Unless such cases can be properly examined, the system will not be useful and may cause concern and resentment. It was not envisaged that an individual PCT should undertake this function. PCTs could not be expected to develop the necessary expertise, as the need for such investigations will arise only rarely. If the Healthcare Commission does not, after all, intend to set up local or regional bases, I think the work of investigation may have to be assigned to the inter-PCT investigation teams, which, in Chapter 27, I recommend should be set up for the purpose of investigating complaints and concerns. That would probably not be the ideal solution but, if training could be provided and a protocol for such investigations could be developed, I think the arrangement would be adequate.

