A confidential study of deaths after emergency medical admission: issues relating to quality of care

E Seward, E Greig, S Preston, RA Harris, Z Borrill, TD Wardle, R Burnham, P Driscoll, BDW Harrison, DC Lowe, MG Pearson

ABSTRACT – In this retrospective pilot study we examine the feasibility of establishing a confidential enquiry into why some patients die after emergency admission to hospital. After excluding those who died in the first hour or who were admitted for palliative care, pairs of physicians were able to collect quantitative and qualitative data on 200 consecutive deaths. Both physicians reported shortfalls of care in 14 patients and one of the pair in 25 patients whose deaths would not have been the expected outcome. In 25, the shortfalls of care may have contributed to their deaths. Major problems were delays in seeing doctors, inaccurate diagnoses, delays in investigations and initiation of treatment. They occurred mostly in those admitted at night. It is possible that establishing the correct diagnosis and starting appropriate treatment may have been delayed in 64% of the 200 patients. The headline figures appear worse than some previous external assessment studies but this study did concentrate on those in whom problems were more likely. Nevertheless, the frequency is too high to be overlooked. In this feasibility study we have demonstrated that it is practicable for local staff to collect and assess data in hospitals and that the types of problems identified are relevant to anyone planning how to organise emergency care. A larger definitive study should be performed.

KEY WORDS: confidential inquiry, deaths, emergency medical admissions, medical errors, physician assessments, reproducibility

Emergency medical admissions have risen steadily. Approximately 90% of patients in acute hospital medical wards present as emergencies.[1,2] In most patients the acute episode can be reversed, but many are admitted in the terminal phases of their illness for palliative care. There is public concern, supported by evidence,[3] that mistakes are being made, and there are many unanswered questions.[4]

Confidential enquiries are well established to examine deaths that occur within either 30 days of a surgical procedure,[5] or within one year of delivery,[6] or after major trauma.[7] These enquiries have significantly improved practice and organisation in these surgical specialties. In the medical specialties, confidential enquiries have examined deaths associated with asthma.[8–12] These too have influenced practice.

Small studies have shown that there may be potentially avoidable factors or adverse events in 3–5% of medical emergency admissions but there has been no systematic attempt to audit why some patients admitted as medical emergencies die. Studies of patient records from a risk management perspective identified mistakes at all stages of patient care.[13,14] Any problems are likely to be compounded by the increasing number and complexity of medical admissions, coupled with reduced junior doctor hours and concerns over the role of the general physician.[15]

This pilot study was designed to assess the feasibility of establishing a confidential enquiry into deaths following admission as a medical emergency.

Method

Patient population

The sample comprised 200 consecutive deaths that occurred within seven days of admission as acute medical emergencies to three general hospitals after 1 January 2000. All had similar on-take teams, with at least one consultant-led post-take ward round every day. Patients were excluded if they died within an hour of arrival, or more than seven days after admission, or if the prime reason for admission was palliative care. One consultant and five specialist registrars (SpRs) (all with MRCP and qualified more than five years) reviewed, retrospectively, the patients' records. These included medical and nursing notes, temperature and drug charts, and laboratory results. A more detailed description of the proforma used and the steps taken to standardise the data collection will appear elsewhere but can be supplied on request.

Quantitative data about patient demographics, timing of medical contacts, investigations performed and management delivered were collected by one doctor using a standardised question proforma devised by the steering committee for the project.

E Seward MRCP, Specialist Registrar, Newham General Hospital

E Greig MRCP, Specialist Registrar*

S Preston MRCP, Specialist Registrar*

*Barts and the Royal London Hospital

RA Harris MRCP, Consultant Physician†

Z Borrill MRCP, Specialist Registrar, Respiratory and General Medicine, Salford Royal Hospital, Manchester

TD Wardle MD FRCP, Consultant Physician†

†Countess of Chester Hospital

R Burnham MA MD FRCP, Consultant Physician and Gastroenterologist, Director, Medical Workforce Unit, RCP

P Driscoll MD FRCSEd FFAEM, Consultant in Emergency Medicine, Hope Hospital

BDW Harrison MA FRCP FCCP, Consultant Physician, Norfolk and Norwich University Hospital

DC Lowe, Medical Statistician‡

MG Pearson FRCP, Consultant Physician, University Hospital Aintree, Liverpool and Director‡

‡Clinical Effectiveness and Evaluation Unit, RCP

Clin Med 2003;**3**:425–34

Two indicators of casemix were used:

1 A performance score[16] not routinely recorded in most hospital notes can be estimated from other available information and gives a simple measure of patients' usual capabilities;[17] it ranges from a score of 1 (normal activity without restriction) to 5 (completely disabled in a bed or chair).
2 The APACHE II score (Acute Physiology and Chronic Health Evaluation (system)) is an established index of severity in the intensive care setting.[18] Our pre-pilot data showed that the PaO_2 was often either unavailable or uninterpretable because the level of inspired oxygen was unknown. We therefore calculated a modified APACHE II of 14 items. This score ranged from 0 to 71, with higher scores indicating more severe illness.

We arbitrarily divided patients into severe (modified APACHE ≥ 20) or less severe (<20) groups for analysis. Questions were piloted on 10 sets of notes in each hospital to exclude ambiguities and ensure that data were collectable.

The second part of the proforma required a series of qualitative judgements about the appropriateness or otherwise of the initial assessment, of the diagnosis reached, the investigation and management plan, and whether the overall care was satisfactory. Two doctors reviewed each patient's record, and reached an independent opinion. The six doctors met before the study to discuss the standards to be applied, and agreed that:

1 The general standard should be 'one they would be happy to share with, and defend to an outside body'.
2 Clinical decisions would be assessed against what it was reasonable to expect a doctor to achieve in the light of knowledge that could or should have been available prospectively in each individual case.

The most important question asked the assessor's opinion as to whether the overall management of the patient was satisfactory. They were asked to judge whether death was:

- expected at the time of admission as a natural course of the illness
- unexpected, but with acceptable management and within the disease process
- unexpected, with some evidence of management problems.

Finally, there was space for free-text comments on patient management.

Statistical note

Data were collected and recorded as for local audit procedures. The records were anonymised to ensure that no individual patient or caregiver was identifiable. Databases were sent to the Clinical Effectiveness and Evaluation Unit (CEEu) at the RCP for collation and analysis.

The results are largely descriptive, using percentages, medians and interquartile ranges (IQR). Standard statistical testing was added to show the stronger correlations between quantitative data and the overall qualitative assessment. Agreement between doctors was assessed using the kappa statistic. Values of 0.20 and below are said to reflect poor agreement, while values of 0.21 to 0.40 indicate fair, 0.41 to 0.60 moderate, 0.61 to 0.80 good, and 0.81–1.00 very good agreement.

Ethical approval

This was requested at, but not required by, each hospital. The medical director and audit committee from each hospital supported the project. Each unit was internally responsible for local data management. Only fully anonymised data were collated centrally.

Results

Quantitative data (Table 1)

The age, sex, performance score and APACHE II score, and presence of co-morbidity were similarly distributed in the three hospitals; the remainder of the analyses are therefore of the amalgamated cohort.

Demographic and admission data Most patients (91%) were aged over 65 years (median 79, IQR 73–87 years), and 62% were females. Most (70%) were admitted via A&E, half of them between 5 pm and 9 am, with slightly more on Fridays and Saturdays.

Estimation of severity/casemix The diagnostic mix is shown in Table 2. The performance score was calculated from 85% of records. Nearly half were graded as 4 or 5. The modified APACHE II score was calculated for 59% of patients, with 29% (34/118) ≥ 20. This percentage was similar for admissions between midnight and 9 am (33%, 8/24), from 9 am to 5 pm (25%, 16/63) and later in the day (33%, 10/30). Most patients (88%) had co-existent medical conditions, with three or more present in 37%.

Initial medical contact The first contact was with a senior house officer (SHO) for 86% of patients and a pre-registration house officer (PRHO) in 10%. The time to being seen by a doctor was less than an hour in 16% (21/155), and less than three hours in 63% (97/155). Ten patients (5%) waited more than six hours to be seen, including 4/34 (13%) of those admitted between midnight and 9 am. The most acutely ill (APACHE II ≥ 20) were recorded as being seen more quickly: 24% (7/29) within one hour, and 83% (24/29) within three hours.

Patient review by a senior doctor Consultant physicians reviewed 81% (161) of cases, and SpRs saw 22 of the 39 patients not seen by a consultant, mostly (16/17 with times recorded) within 24 hours. Of those reviewed by a senior doctor, 62% were seen within 12 hours and a further 29% within 24 hours. For the 17 (9%) cases with no recorded review by a senior professional, all died within three days and 12 within 24 hours – in 10 cases

Table 1. Documented quantitative data. The table gives the number (percentage) documented for each subgroup. Denominator is 200 unless stated in first column.

Patient and admission data							
Age	<65: 18 (9)	65-74: 37 (18)	75-84: 76 (38)	85+: 69 (35)			
Sex	Males: 77 (39)	Females: 123 (62)					
Source of admission	GP: 55 (28)	A&E: 139 (70)	Other: 6 (3)				
Day of admission	Mon: 23 (12)	Tues: 28 (14)	Wed: 23 (12)	Thurs: 23 (12)	Fri: 35 (18)	Sat: 38 (19)	Sun: 26 (13)
Time of day on admission	00:00-08:59: 37 (19)	09:00-16:59: 98 (49)	17:00-23:59: 61 (31)	NK: 4 (2)			
Number of admissions on day of admission	<30: 44 (22)	30-39: 90 (45)	40+: 59 (30)	NK: 7 (4)			
Modified APACHE score	<20: 84 (42)	20+: 34 (17)	NK: 82 (41)				
Performance status (PS)	PS1: 17 (9)	PS2: 33 (17)	PS3: 38 (19)	PS4: 63 (32)	PS5: 19 (9)	NK: 30 (15)	
Comorbidity	None: 25 (13)	One: 49 (25)	Two: 53 (27)	Three: 37 (19)	≥ Four: 36 (18)		
Process of care							
Who (gen med) first saw patient	HO: 20 (10)	SHO: 171 (86)	SpR: 6 (3)	CONS: 1 (1)	NK: 2 (1)		
Hours (gen med) to first seeing patient (198)	<1: 25 (13)	1-5: 120 (61)	6–11: 9 (5)	12+: 1 (1)	NK: 43 (22)		
Hours to review by SpR (50/70)	<1: 3 (2)	1-5: 32 (25)	6–11: 9 (7)	12+: 6 (5)	NK: 20 (15)		
Hours to review by consultant (134/161)	1-5: 20 (12)	6-11: 29 (18)	12–17: 37 (23)	18–23: 32 (20)	24+: 16 (10)	YES, NK: 27 (17)	
Review by another team	Anaesth: 29 (15)	Gen surg: 12 (6)	Other: 4 (2)	NO: 152 (76)	NK: 3 (2)		
First major investigations	CT scan: 25 (13)	Nuclear medicine scan: 1 (1)	Ultrasound: 6 (3)	Endoscopy: 5 (3)	Other: 7 (4)	None: 156 (78)	
Hours to getting first major investigation (44)	<1: 5 (11)	1-5: 15 (34)	6–11: 5 (11)	12-23: 4 (9)	24+: 8 (18)	NK: 7 (16)	
Delay in getting old notes	YES: 9 (5)	NO: 191 (95)					
Minutes before treatment specific to diagnosis started	No treatment: 59 (30)	<60: 48 (24)	60–239: 53 (27)	240+: 40 (20)			
Completeness of relevant charts	Present & complete: 138: (70)		Present & incomplete: 40 (20)		Not present: 22 (11)		
Adverse events recorded during admission	YES: 9 (5)	NO: 191 (95)					
Patient seen by medical team member (not weekends)	YES: 194 (97)	NO: 6 (3)					

NK: not known, NA: not applicable.
APACHE = Acute Physiology and Chronic Health Evaluation (system); CT = computed tomography; CONS = consultant; HO = house officer; SHO = senior house officer; SpR = specialist registrar.

both reviewers agreed this was the expected outcome. All but six of those surviving longer than 24 hours were seen by a member of the medical team every day (except on weekends and Bank Holidays).

Investigations Major investigations (eg computed tomography (CT) scan) were requested in 22% of cases, with a wait of at least 12 hours for 32% (12/37) and of at least 24 hours for 21% (8/37). Investigation rates were similar irrespective of the APACHE score (<20 = 18%, 15/84; ≥ 20 = 24%, 8/34).

Treatment A treatment specific to diagnosis (eg anticoagulation, antibiotics) was documented in 70% of cases. A delay of at least four hours was noted for 39% (12/31) of admissions between midnight and 9 am and for 25% (27/109) of those at other times. The delay was shorter for patients with APACHE scores of 20 or more (median 20 minutes; 4% (1/23) waiting for at least four hours) than for those with lower APACHE scores (median 120 minutes; 31% (20/64) waiting for at least four hours). Delay in treatment was also related to delay in first being seen. Of the 21 patients seen within an hour of admission, only one (5%) had to wait more than four hours for treatment compared with 35% (34/98) of those seen at least one hour after admission.

Qualitative data

Specific assessments Both doctors agreed that an initial diagnosis was recorded in 96% (190) of cases and that it was likely to have been correct in 84% of cases overall; in 69% (11/16) of patients seen first by a PRHO, in 86% (129/150) by an SHO, and in 7 of 7 by an SpR or consultant. Both assessors agreed that there were acceptable records of the history in 63% (125), and examination in 47% (94) and that appropriate investigations were ordered in 79% (158). The initial treatment plan in the notes was judged acceptable for 72% (143), implying concern that initial management could have been better in 28%, and when a plan was set out the assessors considered it appropriate

Table 2. Death certificated causes of death for the 200 cases and the distribution of the 39 cases about whose care concern was expressed. Problems were noted for: 25/146 (17%) of the five common diagnoses versus 14/54 (26%) of the less common diagnoses, suggesting that less familiar conditions are more often associated with management problems.

	Patients (*n*)		No. with problems (groups D and E)	
Pneumonias	43		7	
COPD	30		6	
Stroke	27	73%	5	17%
Cardiac (but not MI)	24		4	
Myocardial Infarction	22		3	
Cerebral bleeds	8		1	
Gastrointestinal bleeds	7		2	
Cancers	6		2	
Renal failure	6	19%	1	29%
Septicaemia	4		1	
Pulmonary embolism	4		2	
Aortic aneurysms	3		2	
Other respiratory failure	3		0	
Old age	3		0	
Perforated viscus	2		0	
Hepatic failure	1		0	
VSD	1	8%	0	19%
Multi-organ failure	1		0	
Intestinal infarct	1		1	
Overdose drugs/alcohol	1		0	
Others	3		2	
	200	100%	39	100%

COPD = chronic obstructive pulmonary disease; MI = myocardial infarction; VSD = ventricular septal defect.

and prompt in only 57% (79/139). Thirty-six per cent (71) of all patients were judged to have had both a correct diagnosis and received prompt appropriate treatment. The results of investigations performed were recorded in 61% (121) of the case notes. In only 59% (71/121) of these did both doctors agree that appropriate action had followed. For 7% (14) both reviewers were concerned that more senior specialist help had not been sought.

Overall assessment Five composite groups were derived from data in Table 3 as follows:

A Both doctors felt the death was expected as a natural course of the illness (66 patients).

B One doctor felt the death was expected whilst the other felt it was unexpected, but both agreed that the management was acceptable (52 patients).

C Both doctors felt the death was unexpected but that management was acceptable (41 patients).

D One of the two assessors felt the death was unexpected with some evidence of management problems (25 patients).

E Both doctors felt the death was unexpected with some evidence of management problems (14 patients) (Table 4).

If we accept the more critical view of the two assessing doctors, then problems were suspected in 39 of 198 cases. The range 14 (7%) to 39 (20%) measures the uncertainty between doctors, and provides a current working range for the prevalence of unexpected death with evidence of management problems. There was a moderate amount of agreement between the two observers (see Table 3).

Key Points

This pilot study shows it is possible for physicians to assess quality of care retrospectively

A combined quantitative/qualitative approach detected potential 'faults' in 7% of deaths (both observers) and in a further 12% (one observer)

Problems included errors of diagnosis, delays in treatment, and were most common in those admitted at night

The frequency of medical errors implies a need for further studies to enable routine identification of problems so action can be taken to reduce the problem

The groups A–E were compared in regard to demographic and process of care data (Table 4). Those who died unexpectedly (C+D+E) had more often been admitted between midnight and 9 am than other patients (A+B), had lower APACHE scores, and had to wait longer for their first major investigation. Where there was suspicion of unacceptable management (D+E), patients experienced significantly longer delays before being seen by a doctor and for treatment specific to their diagnosis. Across all five groups there were no obvious associations regarding age, sex, the presence of comorbidities, source of admission, and the total number of admissions on the day of admission. Overall there were few differences between patient groups D and E, that is between whether only one or both doctors rated the management as unacceptable. One or both doctors felt that a wrong diagnosis had been made in 53%

Table 3. Agreement between doctors when making a subjective conclusion about whether death was expected at the time of admission. Kappa coefficient of agreement between doctors was 0.36 for the whole table excluding the missing data, 0.43 for whether the death was expected or not, and 0.47 for whether the death was unexpected with some evidence of management problems or not.

	Death expected at time of admission as a natural course of the illness	Death unexpected but acceptable management and within the disease process	Death unexpected with some evidence of management problems	Missing	Total
Death expected at time of admission as a natural course of the illness	66	17	1	–	84
Death unexpected but acceptable management and within the disease process	35	41	6	1	83
Death unexpected with some evidence of management problems	4	13	14	–	31
Missing	1	–	1	–	2
Total	106	71	22	1	200

Table 4. Associations between documented quantitative data and whether death was unexpected at the time of admission.

		Qualitative assessment of death					
		A (n = 66)	B (n = 52)	C (n = 41)	D (n = 25)	E (n = 14)	
198 cases unless stated in this column.	Statistic	Both doctors felt the death was expected in the admission as a natural course of the illness	One doctor felt the death was expected whilst the other felt it was unexpected with acceptable management	Both raters felt the death was unexpected with acceptable management	One or doctor felt the death was unexpected with some evidence of management problems	Both doctors felt the death was unexpected with some evidence of management problems	p value KW: Kruskal-Wallis CHI: Chi-squared, 4 df
Patient and admission data							
Age	Median (IQR)	82 (73-88)	79 (75-86)	79 (75-89)	82 (75-90)	74 (63-81)	KW: p = 0.24
Sex	Males (%)	23 (35%)	21 (40%)	16 (39%)	9 (36%)	7 (50%)	CHI: p = 0.86
Source of admission	A&E (%)	47 (71%)	36 (69%)	28 (68%)	16 (64%)	10 (71%)	CHI: p = 0.97
Day of admission (194)	Weekend (%)	19 (29%)	18/49 (37%)	16 (39%)	8/24 (33%)	1 (7%)	CHI: p = 0.22
Time of day on admission (194)	00:00-08:59 (%)	5 (8%)	7/49 (14%)	10 (24%)	9/24 (38%)	5 (36%)	CHI: p = 0.004
Admissions on day of admission (191)	Median (IQR)	38 (31-42)	35 (29-40), n = 47	33 (30-42), n = 39	37 (32-44)	38 (31-43)	KW: p = 0.59
Modified APACHE score	Yes (%)	35 (53%)	39 (75%)	23 (56%)	13 (52%)	7 (50%)	CHI: p = 0.11
Modified APACHE score (117)	Median (IQR)	20 (14-24), n = 35	17 (13-21), n = 39	14 (8-16), n = 23	11 (10-17), n = 13	12 (10-16), n = 7	KW: p<0.001
Performance status	Yes (%)	56 (85%)	39 (75%)	40 (98%)	22 (88%)	12 (86%)	CHI: p = 0.05
Performance status (169)	PS 4&5 (%)	30/56 (54%)	24/39 (62%)	17/40 (43%)	10/22 (45%)	1/12 (8%)	KW: p = 0.03
Comorbidity (n: range 0-7)	Median (IQR)	2 (1-3)	2 (1-4)	2 (1-3)	2 (1-3)	2 (2-4)	KW: p = 0.30
Process of care							
Who (gen med) first saw patient (196)	HO (%)	3/65 (5%)	5 (10%)	5 (12%)	4/24 (17%)	3 (21%)	CHI: p = 0.25
Time (gen med) to first seeing patient (153/198)	6+hours (%)	2/48 (4%)	2/41 (5%)	1/33 (3%)	3/20 (15%)	2/11 (18%)	KW*: p = 0.01
Review by registrar	Yes (%)	21 (32%)	20 (39%)	17 (42%)	7 (28%)	5 (36%)	CHI: p = 0.77
Time to review by registrar (50/70)	6+hours (%)	3/15 (20%)	5/15 (33%)	3/14 (21%)	3/5 (60%)	1/1 (100%)	KW**: p = 0.11
Review by consultant	Yes (%)	48 (73%)	44 (85%)	36 (88%)	19 (76%)	12 (86%)	CHI: p = 0.29
Time to review by consultant (132/159)	24+hours (%)	2/37 (5%)	5/39 (13%)	3/32 (9%)	4/15 (27%)	2/9 (22%)	KW**: p = 0.16
Review by another team	Yes (%)	16 (24%)	10 (19%)	7 (17%)	6 (24%)	6 (43%)	CHI: p = 0.53
First major investigations	Yes (%)	19 (29%)	10 (19%)	7 (17%)	3 (12%)	5 (36%)	CHI: p = 0.24
Hours to getting first major investigation (37/44)	Median (IQR)	2 (1–6), n = 17	2 (3–5), n = 7	24 (12–36), n = 7	6 & 122, n = 2	35 (5–120), n = 4	KW: p = 0.003
Delay in getting old notes	Yes (%)	2 (3%)	1 (2%)	2 (5%)	4 (16%)	0 (0%)	–
Minutes before treatment specific to diagnosis begun (140)	Median (IQR)	75 (25–158), n = 42	69 (16–240), n = 37	90 (45–195), n = 29	120 (62–405), n = 20	360 (38–510), n = 12	KW: p = 0.04
Completeness of relevant charts	Present & complete (%)	41 (62%)	42 (81%)	26 (63%)	19 (76%)	8 (57%)	CHI: p = 0.14
Adverse events recorded during admission	Yes (%)	2 (3%)	2 (4%)	3 (7%)	1 (4%)	1 (7%)	–
Patient seen by medical team member (not weekends)	No (%)	3 (5%)	2 (4%)	0 (0%)	1 (4%)	0 (0%)	–

NK: not known; NA: not applicable; IQR = interquartile range.
Weekend: from 00:00 Saturday to 23:59 Sunday.
* Ordered categories used in Kruskal-Wallis test: <1 hour, 1-5 hours, 6-11 hours, 12+ hours. ** Ordered categories used in Kruskal-Wallis test: <1 hour, 1-5 hours, 6-11 hours, 12-17 hours, 18-23 hours, 24+ hours.

(16/30) of groups D+E, but in only 8% (11/43) of other patients.

On further record review, of the 39 cases in groups D and E (Table 4) it was felt that management problems had contributed to the death of 11/14 patients in group E and 14/25 in group D (Table 5). In 11 cases diagnoses were incorrect, either because symptoms and signs were not properly investigated or because results were misinterpreted. Lack of treatment, often by junior staff, was the problem in six cases and delayed treatment in a further five. Six patients had been inadequately reviewed by senior staff.

Free text

We identified problems of varying severity in some aspects of care in 172 patients. Many were not thought to contribute directly to death, but could have had an effect on morbidity. Issues identified included:

- incomplete examination (46 patients)
- failure to request bedside glucose estimation in patients with acute neurological symptoms/signs (10 patients)
- failure to use appropriate oxygen concentration in shocked patients (9 patients)
- inappropriate fluid replacement in the shocked patient (9 patients)
- incorrect management of hyperkalaemia (4 patients)
- inappropriate management of renal failure (4 patients)
- misdiagnosis of subarachnoid haemorrhage (1 patient).

Discussion

This pilot study had three objectives:

1 to examine the feasibility of a confidential study of patient deaths after acute admission
2 to define the scale of any management problems
3 to explore the technical, medical and political obstacles to be resolved if a more extensive study is to be planned for the future.

Management problems of varying degrees of severity were present in 172/200 patients. In 14 (7%) of cases, both assessors felt that death was unexpected and there was some evidence of management problems which probably contributed to death in 11 patients (Table 5). In a further 25 (12.5%) one or other assessor expressed concern over the patient's acute care. These headline

Table 5. Summarised comments of the assessors for the 25 patients in whom management problems were possible contributors to their deaths.

Age	Eleven patients in whom *both* assessors felt that management issues may have contributed to the death
58, 63	Profound hypoxia, neither oxygen nor ventilatory support
72, 79	Oliguric, acute renal failure, no USS, insuffient volume replacement
72, 75	Aortic aneurysm missed, received either heparin for non-existent DVT or streptokinase for non-existent MI in hypotensive patient with recorded arm BP variation
80	Community-acquired pneumonia missed, raised white count, no antibiotics given
72	New onset atrial fibrillation, no rate control, no anticoagulation. Stroke after 5 days
50	Severe acidosis and hyperkalaemia, only PRHO review for 48 hours
61	Almost no records of initial assessment, investigations not performed and subsequent management delayed
82	Acute MI missed despite ECG and enzyme evidence
	Fourteen patients in whom *one* assessor felt that management issues may have contributed to death
66	Anaemic, no investigation. Certified as MI – no evidence
64, 68	Misdiagnosed septic shock. Insufficient volume replacement. No antibiotics given
76	Oliguric, acute renal failure, no USS, insuffient volume replacement
77	Stated for resuscitation, previously well – not resuscitated
78, 84	Septic, pleural effusion not aspirated, no antibiotics given
90	Community-acquired pneumonia missed, raised white count, no antibiotics given
95	Acute MI missed despite ECG and enzyme evidence
85	Heart failure missed initially, then not treated for 48 hours
96	Penicillin allergy, given penicillin. Certified as PE with no evidence
69	Acute confusion, only investigation FBC, U&E, not even BM
89	Subarachnoid haemorrhage. Possible herald bleed ignored. Patient pyrexial, GCS 15, presumed to be viral meningitis. No LP. Head CT 4 days later when GCS 8
91	Head laceration, lateralising signs. No Head CT. Death atfter 72 hours

CT = computed tomography; DVT = deep vein thrombosis; ECG = electrocardiogram; FBC = full blood count;
GCS = Glasgow Coma Scale; LP = lumbar puncture; MI = myocardial infarction; PE = pulmonary embolus; U&E = urea and electrolytes; USS = ultrasound scanning.

figures should be interpreted with caution as they are higher than in previous studies,[3,19,20] partly because the denominator population excluded many of those expected to die.

Retrospective data collection

In a retrospective snapshot it is easy to be wise after the event. The logistics of prospective data collection are close to impossible, and the retrospective approach mirrors the way in which others judge medical care when there have been critical incidents or complaints. The assessments depend on the accuracy, detail, legibility and quality of the case notes, and we adopted the 'legal' approach that 'if not documented, then presumed not to have been done'.

There are two approaches to collecting data retrospectively. Either all possible items of information about every aspect of each patient are collected on a standardised proforma, from which a neutral observer can produce analyses of what was done well or badly; or a more limited range of data is collected, relying on the medical skill of the assessor to sort out what is relevant to the particular case – the approach we adopted. While the former is more objective, the volume of data to be collected when including all causes of admission was beyond the project's resources. Instead we employed experienced physicians from within the hospital and used their experience of what was best care to get a 'feel' for what was important in each case while making allowances for information that was – or should have been – available. This was augmented with a limited range of more objective variables.

Consultants are expected to perform audit assisted by SpRs who collect appropriate data as part of their training recommended by General Medical Council and Royal Colleges. This keeps patient-identifiable information confidential to the medical team and avoids the need to share identifiable data with outsiders. These practical advantages of using internal medical assessors to review decisions made by colleagues and friends raise concerns about bias. We cannot tell if the assessors behaved as 'hawks' or 'doves', but analysis of the free-text comments does not suggest any tendency to hold back on criticism. However, if they have been too 'lenient' on their colleagues the data are even more worrying and if they overestimated errors, even by a factor of 2, the proportion of serious 'faults' would still be cause for concern.

The analysis of the duplicate case assessment suggests that both collectors were broadly in agreement about the scale of the problems and that there was a reasonable kappa score for the judgemental questions about overall management. The medico-legal process will often produce two strongly held polarised views of the same case for the courts, demonstrating that medicine remains an inexact science with few absolute standards. We have reported two standards of assessment in this paper. When both assessors agreed, it is likely that there really was a management problem. We also report those cases of unexpected death where one but not the other assessor felt that care was poor. If one doctor is able to criticise a process then either a patient or relative could have cause for concern, but equally in this situation the other doctor would be able to argue against that. Confounding factors such as doctor fatigue and non-availability of support services have not been measured.

Magnitude of the problems found

Many of those who died had severe medical problems. There was at least one co-existing medical condition in 88%, and half were physically impaired such that they had difficulty with washing and dressing (performance score of 4 or 5). The good news is that the sickest patients were seen the quickest. However, an appropriate management plan was judged to have been delivered promptly in 57% of all patients, and only 36% had both correct initial diagnosis and appropriate treatment when viewed retrospectively. Although these figures are discouragingly low, it is not known what the best achievable figure would be, but it is perhaps worth noting that the more senior the doctor seeing the patient initially the more likely the diagnosis was to accord with the final diagnosis.

Only 12 out of a possible 200 post-mortems were performed. This was too few to be a useful gold standard and so we had to rely on the clinical expertise of our medical assessors to evaluate all the data available at the time of death to judge the appropriateness of the diagnosis, the likelihood that the patient might have survived and whether the necessary care for that survival had been delivered. Both assessors agreed that medical care had fallen short of reasonable standards in 14 (7%) cases who would have been expected to survive the admission, and one or other assessor expressed concern in a further 25 cases. However, the absolute figures quoted must be qualified in that the study excluded some patients who were expected to die (see below). The numbers of exclusions were not recorded but their inclusion would have reduced the headline figure, although it would probably not be lower than that reported in other studies.[3,20] Most doctors would like to deny that such error rates applied to their practice but audit is an unforgiving tool. Experience from other large audits also shows standards of practice well below what most would consider reasonable.[17,21]

There are positive significant correlations between the subjective assessments of management and some of the objective measures. While there may be an element of self-fulfilment about such analyses, that many of them seem to make good clinical sense supports the validity of the overall assessment.

Choice of cases to be studied We included consecutive deaths within the hospitals to avoid selection bias, but did exclude three groups in whom problems were less likely. The death of patients admitted primarily for palliative care is not unexpected. Deaths in the first hour include many who had arrived in a pre-morbid state. Deaths at more than seven days after admission are less likely to be related to acute medicine but more to the reasons for staying in hospital that many days. The investigators paid most attention to areas where there was the greatest chance of finding remediable management problems.

The frequency of identified problems was independent of age and of diagnosis, but there was a trend for more problematic

management to be identified among the least common admission diagnoses. It was not possible to ascertain whether this was because of lack of familiarity with less common conditions or because the admission system is better suited to cope with common problems, such as rapid triage of myocardial infarction to the coronary care unit. Emergency medicine covers all diagnoses and ages, and if the complex organisational factors needed to cope with this diversity are to be understood, future studies will also need to be based on unrestricted consecutive admissions. Narrow studies of one diagnosis or specialty could make it harder to solve the wider organisational issues of the admissions system.

Overall quality of care delivered The 39 'problem' cases were more likely to wait longer to see a doctor and to have been admitted in the night. Diagnostic errors, delays in investigation and in treatment appear to be the more common problems. Incorrect diagnosis was related to clinical features that were misinterpreted or not properly investigated. Diagnostic accuracy was related to experience, but nearly one in 10 patients was apparently not reviewed by either SpR or consultant. The problems identified in Table 5 and the free-text comments support these observations.

Would an increase in senior staff have improved matters? The National Confidential Enquiry into Peri-operative Deaths studies[22] showed that the hazards of surgery were less when operations were performed by senior experienced surgeons during the day. There is some evidence in this study that staff of greater seniority would have made fewer diagnostic errors, with continuing benefits to patients. The inability to recognise – at an early stage – those who are either severely ill or deteriorating may have been an important factor. With the current numbers of consultants it would not be feasible to have 24-hour consultant physician cover for all medical admissions to hospital. The pressure to increase the commitment to medical emergencies has to be balanced against the continuously expanding demands on consultants.

Alternative solutions have been advocated,[4,23] including simple protocol-driven triage schemes that could be developed to ensure proactive early identification of patients whose condition is deteriorating.[24–26] One short-term benefit of such a scheme might be earlier senior review of the sickest patients.

The future

This pilot study set out to investigate whether it was possible to perform an inquiry into the deaths of patients after medical admission and whether the effort justified the process. The answer to both questions would seem to be affirmative. It is possible to collect data that identify clinical management problems that deserve to be investigated. While some problems are either unavoidable or inconsequential, it is probable that many could – or should – be prevented.

Methodologically there are still issues to be resolved. These include:

- improving the questions (especially the subjective ones) to achieve more reliable data collection
- defining quality of care more objectively
- managing the whole process with medical assessors within the hospital and making this process adequately robust and impartial
- incorporating this as a part of routine audit in hospitals and ensuring that the audit loop is closed and action follows.

This study identified problems of diagnosis, wrong treatment and delayed treatment as significant issues, but did not collect data on the organisation of care. There is huge interest in how best to organise acute care under the challenges of reduced junior hours and rising numbers of referrals. Studies like this should be more widely available to assess the effectiveness of the various patterns of care in different hospitals. The serial national audits of stroke care have shown that hospital staff are rarely aware of the problems identified, but once alerted are keen to improve and prove that improvement is possible.[27] If this study could be repeated across multiple sites to yield some benchmark figures, it would add more objective data against which changes in organisation could be planned and assessed.

This study concentrated only on deaths, but the issues raised are likely to be relevant for all medical admissions. The costs of this study were small since all staff gave freely of their time. A formal cost analysis has not been done, but the techniques of case review are not expensive. With more refined questions, the process could be repeated widely, provided it was seen as a routine part of the work plan of SpRs and consultants. Any costs must be balanced against the potential savings from avoidance of error that may reduce length of hospital stays, improve outcome, and potentially reduce the cost of litigation.

A larger-scale investigation into deaths in medical patients within one week of admission to hospital as an acute emergency is practicable and needed. Its main aim should be to refine the data-collecting techniques, make the assessments more objective, and develop ways of handling the data as part of the clinical governance procedures. Thereafter the procedures should become part of every hospital's routine audit of medical care quality.

References

1 NHS Trust Federation and the King's Fund. *Analysing changes in emergency medical admissions*. London: NHS Trust Federation, 1995.
2 Houghton A, Hopkins A. Acute medical admissions: results of a national audit. *J R Coll Physicians Lond* 1996;**6**:551–9.
3 Vincent C, Neale G, Woloshynowych M. Adverse event in British hospitals: preliminary retrospective record review. *BMJ* 2001;**322**:517–9.
4 Alberti KGGM. Medical errors: a common problem. *BMJ* 2001;**322**;501–2
5 Royal College of Surgeons. *Report to the National Confidential Enquiry into Perioperative Deaths (1996-1997)*. London: Royal College of Surgeons, 1998.
6 Department of Health. *Why mothers die. Report on confidential enquiries into the maternal deaths in the United Kingdom (1994-1996)*. London: The Stationery Office, 1998.
7 Anderson ID, Woodford M, de Dombal FT, Irving M. Retrospective study of 1000 deaths from injury in England and Wales. *BMJ* 1988; **296**:1305–9.

8 Wareham NJ, Harrison BDW, Jenkins PF, Nicholls J, Stableforth DE. A district confidential enquiry into deaths due to asthma. *Thorax* 1993;**48**:1117–20.

9 Bucknall CE, Slack R, Godley CM, Mackay TW, Wright SC. Scottish Confidential Enquiry into Asthma Deaths (SCIAD) 1994-6. *Thorax* 1999;**54**:978–84.

10 Burr ML, Davies BH, Hoare A, Jones A *et al.* A confidential enquiry into asthma deaths in Wales. *Thorax* 1999;**54**:985–9.

11 Mohan G, Harrison BDW, Badminton RM, Mildenhall S, Wareham NJ. A confidential enquiry into deaths caused by asthma in an English health region: implications for general practice. *Br J Gen Pract* 1996;**46**:529–32.

12 Harrison BDW, Slack R, Berrill WT, Burr ML *et al.* Results of a National Confidential Enquiry into Asthma Deaths. *Asthma J* 2000;**5**:180–86.

13 Neale G. Risk management in the care of medical emergencies after referral to hospital. *J R Coll Physicians Lond* 1998;**32**(2):125–9.

14 Neale G. Clinical analysis of 100 medicolegal cases. *BMJ* 1993;**307**:1483–7.

15 Wardle TD. Whither the acute medical patient? *J Accid Emerg Med* 1997;**14**:167–9.

16 Oken MM, Creech RH, Torney DC, Horton J, Davies TE. Toxicity and response criteria of the Eastern Co-operative Oncology Group. *Am J Clin Oncol* 1982;**5**:649–55.

17 Roberts CM, Ryland I, Lowe D, Kelly Y *et al.* Audit of acute admissions of COPD: standards of care and management in the hospital setting. *Eur Respir J* 2001;**17**:343–9.

18 Gunning K, Rowan K. Outcome data and scoring systems. *BMJ* 1999;**319**:241–4.

19 Wilson RM, Runciman WB, Gibberd RW, Harrison BT *et al.* The quality in Australian health care study. *Med J Aust* 1995;**163**:458–71.

20 Leape LL, Brennan TA, Laird NM, Hebert L *et al.* Incidence of adverse events and negligence in hospitalised patients: results of the Harvard medical practice study II. *N Engl J Med* 1991;**324**:377–84.

21 Rudd A, Irwin P Rutledge Z, Lowe D *et al.* The national sentinel audit of stroke: a tool for raising standards of care. *J R Coll Physicians Lond* 1999;**33**:460–5.

22 Campling EA, Devlin HB, Hoile RW, Lunn JN. *The report of the National Confidential Enquiry into Peri-operative Deaths*. London: National Enquiry into Perioperative Deaths, 1991/2.

23 Rhodes J, Harrison BDW, Black DA, Spiro S *et al.* General internal medicine and speciality medicine – time to rethink the relationship. *J R Coll Physicians Lond* 1999;**33**:341–7.

24 Baudouin S, Evans T. Improving outcomes for severely ill medical patients. *Clin Med* 2003;**2**:92–9.

25 Whitehead MA, Puthucheary Z, Rhodes A. The recognition of the sick patient. *Clin Med* 2003;**2**:95–8.

26 Puthucheary Z, Whitehead MA, Grounds M. Early management of the critically ill patient. *Clin Med* 2003;**2**:98–100

27 Rudd AG, Lowe D, Irwin P, Rutledge Z, Pearson M. National stroke audit – a tool for change? *Qual Health Care* 2001;**10**:141–51.