Consensus methods to identify a set of potential performance indicators for systems of emergency and urgent care

Patricia Coleman, Jon Nicholl
Medical Care Research Unit, University of Sheffield, Sheffield, UK

***Objectives*: To identify a comprehensive set of indicators to enable Primary Care Trust (PCT) commissioners in England and other NHS decision-makers to monitor the performance of systems of emergency and urgent care for which they are responsible.**

***Methods*: Using a combination of Delphi RAND methods in three successive rounds of consultation and nominal group review, we canvassed expert opinion on 70 potential indicators as good measures of system performance. The two Delphi panels consisted of senior clinicians and researchers, and urgent care leads and commissioners in PCTs and Strategic Health Authorities (SHAs). The indicators were formatted into a questionnaire according to whether they were outcome, process, structure, or equity-based measures. Participants scored each indicator on a Likert scale of 1–9 and had the opportunity to consider their scores informed by the group scores and feedback. The questionnaire was refined after each round. To ensure that the indicators rated most highly by the Delphi panels covered all dimensions of performance, the results of the Delphi were reviewed by a nominal group consisting of two researchers and three clinicians from the local health services research network (LHSR).**

***Results*: Overall, the process yielded 16 candidate indicators. It also produced a core set of serious, emergency and urgent care-sensitive conditions (defined as conditions whose exacerbations should be managed by a well-performing system without admission to an inpatient bed), for use with the indicators.**

***Conclusions*: System-wide measures to monitor performance across multiple services should encourage providers to work for patient benefit in an integrated way. They will also assist commissioners to monitor and improve emergency and urgent care for their local populations. The indicators are now being calculated using routinely available data, and tested for their responsiveness to capture change over time.**

Journal of Health Services Research & Policy Vol 15 Suppl 2, 2010: 12–18

Background

In October 2001, the government in England introduced a ten-year strategy to reform the delivery of emergency health services and improve the quality of the patients' experience of care.[1] The strategy recognizes that patients may not attend or consult a single provider but are treated by systems of care. This is particularly important in emergency and urgent care where the numbers of services, access points and routes into care together with the heterogeneity of the case-mix raise questions for Primary Care Trust (PCT) commissioners and other National Health Service (NHS) decision-makers about how best to monitor and manage performance of the whole system.

Existing performance measures focus on separate services. For example, the four hour target in emergency departments (ED),[2] ambulance response time standards,[3] and General Practitioner (GP) appointments in 48 hours.[4] Protocols for the management of care at the interface between primary and secondary care tend to be disease-specific, for example, for cancer[5] or menorrhagia.[6] Each service or care package may be effective, but patients seeking care urgently (which for them may be a rare and worrying event) may not find the experience either efficient or satisfactory. Emergency and urgent care in the UK is provided by a range of services including 999 ambulance, ED, GP both in and out of hours, pharmacy, NHS Direct (NHSD),

Patricia Coleman BA, Research Fellow; **Jon Nicholl DSc**, Professor and Director, Medical Care Research Unit, University of Sheffield, Regent Street, Sheffield S1 4DA, UK.

Correspondence to: p.coleman@sheffield.ac.uk

  DOI: 10.1258/jhsrp.2009.009096

walk in centres (WIC), minor injury units (MIU), dental services, community mental health teams, and social care.[7] 'Progress', both into care, and towards recovery, is known to be a key concept in how patients judge the quality of their experience.[8] Existing service-based measures do not capture what for patients may be a protracted or smooth pathway to care. Adopting a 'systems' approach to emergency and urgent care would encourage providers to work for patient benefit in an integrated way across organizational boundaries. 'System wide' measures would also assist commissioners to monitor and improve the performance of the care they commission for their populations. We are undertaking a two-phase study to develop suitable system-wide measures and report phase one of this work.

Aims and objectives

Our aim was to draw on the expertise of different panels to develop a set of system-wide indicators which are judged to be related to quality of care or provision of emergency and urgent care so that changes over time can be interpreted simply as 'improvement' or 'deterioration'.

Methods

We reviewed the scientific and policy literature and identified a long list of 70 potential indicators and candidate conditions.[9] We then used Delphi consensus methods based on the RAND design[10] to elicit expert opinion on the suitability of each indicator separately to be a 'good' measure of performance in three consecutive rounds of consultation and feedback. During this process the lists of indicators and candidate conditions were revised and refined. To ensure that as a set the indicators rated highly by the Delphi panels were also comprehensive and inclusive, an overview of the results was provided by a nominal review group.

Questionnaire development

The 70 indicators were grouped according to whether they were measures of (i) health outcomes (ii) processes (iii) structures, or (iv) equity (for example, having the potential to identify variations in outcomes or processes between weekdays and weekends[11]). The research team did not have any prior expectations of the results. All 70 indicators were considered to be capable of providing a reasonable and different perspective on the performance of systems of emergency and urgent care. However, it was felt that to cover all perspectives, the final set should contain at least one measure based on outcomes, process, structural and equity. All the items were formatted into a questionnaire that could be completed and returned electronically. The questionnaire also listed 27 serious emergency conditions (including external causes, injuries, poisoning and violence), and 16 urgent conditions, as candidate conditions to be used with the indicators. 'Urgent care-sensitive conditions' were defined as *'conditions whose exacerbations could be managed by a well-performing EUC [emergency and urgent care system] out of hospital or in emergency departments (EDs) without admission to an inpatient bed.'*

Recruitment of Delphi panels

To include a diverse range of opinion from stakeholders involved in emergency and urgent care, we recruited two panels of experts. Panel one was recruited from a list of 33 senior clinicians and researchers nominated by our local health services research (LHSR) network as having a special clinical and or research interest in some aspect of emergency and urgent care.[12] The specialty profile of the peer-nominated group included emergency nursing, acute medicine, MIU/WIC nursing, consultants in anaesthesia/intensive care, emergency medicine, psychiatry, public health, obstetrics and gynaecology, paediatrics and pre-hospital care. General practice, NHS Direct, pharmacy and two patient representatives were also included. Panel two was recruited from 43 PCTs that had responded to a previous survey of emergency and urgent networks.[9] These PCTs were assumed to have a particular interest in emergency and urgent care and be potential users of performance indicators.

Our target sample for both panels was between 13 and 20, which is the size of panel recommended in the RAND manual.[10] However, given the range of expertise in emergency and urgent care and the importance when recruiting Delphi panels to include a wide range of views we aimed to recruit as many participants from our 'pools' as were willing to assist.

Recruitment to each panel followed the same process. Invitations and information about the study were sent to potential participants either by an email or letter 'alert' several days before the questionnaire. The 'alert' indicated that the questionnaire would be mailed unless the potential panellist declined to take part by a particular date. The consultation was sequential. After a small pilot study to test the acceptability and feasibility of the methods, the questionnaire was sent to panel one. Panel one participated in rounds one and two of the Delphi; panel two participated in round three.

Scoring the indicators

In all rounds, each panellist was asked to score his or her agreement with a statement that *'this measure is likely to be a good indicator of the performance of the emergency and urgent care system'*, on a Likert scale of 1–9 (1 being 'disagree strongly' and 9 'agree strongly'). The emergency and urgent conditions were scored on the same scale. There was also space for members to add any comments

about the usefulness or otherwise of each of the items included and to propose other candidate conditions. A Delphi approach is inductive. As the literature warns against imposing preconceptions that may inhibit the contribution of other perspectives,[12] we did not specify which services were constituents of the EUCS. Conceptually, these will vary between different areas depending on how pathways into emergency and urgent care are configured locally. Similarly, the criteria by which participants should interpret 'good' were not specified. Our aim was to draw on the expertise of panellists who might have different ways of judging what would be a 'good' indicator. Participants were asked not to limit their views about the potential usefulness of an indicator by perceived difficulties in collecting or processing the data required to calculate them.

Analysis

After each round, the average score of the panel for each measure was calculated. A revised questionnaire incorporating the panellist's individual score, the panel average score, and comments sorted by indicator, was circulated to the panellists to enable them to consider their own score in the light of the group score and feedback from the previous round if they so wished.

Delphi techniques incline scores towards the middle and the median for most of the indicators was 5–6. The change in the median scores of panel one between rounds one and two was ± 0.5 with no change in the scores for those items that scored ≥ 7. After three rounds, although there were small differences in the higher end median scores between the two panels, the indicators ranked ≥ 7 by panel two in round three were exactly the same as those rated ≥ 7 by panel one in rounds one and two. At this stage, we felt that a consensus had been reached and did not mail a further round to panel two. The candidate measures were ranked by their median scores and classified as potentially good (≥ 7), moderate (5–6) or poor (<5). We combined the scores of the two panels in order to assist the subsequent review by the nominal group. The maximum possible score therefore for each measure was 18. Categories were assigned as follows.

14–18 = good
10–13 = moderate
<10 = poor

We calculated a measure of disagreement expressed by the mean absolute deviation of the individual scores from the group median (MAD-M).

Nominal review group

The Delphi panels had scored each candidate indicator in isolation. However, individual indicators that scored highly may not meet the requirements for a set of indicators to be comprehensive over different dimensions of performance and inclusive of different patient groups. The brief of the nominal review group was to consider the results of the Delphi as a whole. It was tasked with identifying a set of indicators that conformed to the attributes of a good performance indicator,[13–16] and were comprehensive and inclusive. This was felt to be necessary to allow the indicators to be used flexibly across boundaries of general and specialized systems of care (for example, ED, paediatrics, and mental health); and be relevant to all sub-groups of the population (for example, young men, children, women, elderly). The nominal review group discussions were held face-to-face.

Results

For the Delphi survey, we recruited 30 members to panel one and 19 to panel two. The nominal review group consisted of two members of the research team and three clinical researchers with expertise in urgent and emergency care from the LHSR network.

The Delphi consultation revealed a consensus amongst researchers, clinicians and NHS commissioners ($n = 49$) that 14 of the 70 candidate indicators were likely to be 'good' measures of the performance of emergency and urgent care systems. The range of disagreement was 0.9–1.7, see Table 1. The range of agreement and disagreement for the serious, emergency, and the urgent care-sensitive conditions rated by the panels as 'good' is shown in Table 2.

Outcome, process, and equity-based indicators were all rated highly by the two Delphi panels with process measures being rated most strongly. The original list of 70 indicators contained two structural measures only. Both measures achieved median scores of 5 and were classified as 'moderate'. In providing an overview of the Delphi results, the nominal group agreed that given the development of the indicators was still at its pre-testing stage it was important to retain at least one indicator in each of the four categories (outcomes, process, structural and equity) to provide measures of performance from each of these perspectives. It was also felt that a structural measure of location of services in relation to the population may be useful to capture the impact of opening or closing services. The original wording of one of the two structural indicators was '*. . .Proportion of population living within 10 km of 24 hour emergency or urgent ambulatory care facilities*'. The range of disagreement in the scoring of this particular indicator was wider than for the other structural indicator (1.7 vs 0.97) and it had prompted one panellist to comment:

> might have perverse effects in services with low usage being provided at times when not needed.

Table 1 Delphi consensus of what is likely to be a 'good' performance indicator for the emergency and urgent care system (EUCS) ($n = 14$)

	Agreement	Disagreement
Outcomes		
Mortality rates for serious, emergency conditions for which a well-performing EUC system could improve chances of survival	15	1.2
Case fatality ratios for serious emergency conditions for which a well-performing system could improve chances of survival	15	1.3
Process		
Admission		
Hospital emergency admission rates for urgent conditions, the exacerbations of which could be managed out of hospital or in emergency departments (EDs) without admission to a hospital bed	16	1.4
Arrivals at EDs referred by any EUCS services and discharged without treatment or investigation(s) that needed hospital facilities	15.5	1.2
Arrivals at EDs by emergency ambulance and discharged without treatment or investigation(s) that needed hospital facilities	15	1.4
Adherence to evidence-based good practice guidelines for serious emergency and urgent conditions	15	1.3
Service users		
Multiple transfers between EUCS services	14	1.4
Timings		
Time from first contact with a EUCS service to clinical assessment. For example:		
a) call to NHS Direct (NHSD) to nurse contact	15	1.5
b) call to ambulance service (AS) to paramedic contact (ie. time on scene)	15.5	1.3
c) call to GP in-hours or out-of-hours (OOH) to clinical assessment by primary care team	15.5	1.3
For patients with indicator conditions who are admitted, time from first contact with a EUCS service to time of admission. For example:		
a) call to NHSD to admission	15	1.6
b) call to AS to admission	15	1.4
c) call to GP in-hours or OOH to admission	15	1.4
d) call to mental health team to admission	14.5	1.4
Time from first contact to definitive care for indicator conditions e.g. for patients having thrombolysis – call to needle time; for patients having percutaneous coronary intervention (PCI) call to Cath lab; patients with serious head injury, undergoing neurosurgery – call to theatre; for mental health crisis – call to contact with mental health crisis team	17.5	0.9
Equity		
Variations in relative case fatality ratios (i.e. deaths as a proportion of contacts) for serious emergency conditions between contacts made		
a) in hours vs. OOH	15	1.2
b) weekdays and weekends	15	1.2
Variations in times from first call to any EUCS service to first clinical assessment, for example with NHSD, AS, GP, mental health team between:		
a) in hours vs. OOH	15	1.5
b) weekdays and weekends	14	1.3
c) area of residence		
Variations in times from first contact with any EUCs service for example, NHSD, AS, GP, mental health team, to admission between:		
a) in-hours vs. OOH	15	1.7
b) weekdays vs. weekends	14	1.5
Variations in times from first contact to definitive care for example, patients having thrombolysis – call to needle time; for patients having PCI – call to cath lab; patients with serious head injury – call to theatre, between:		
a) in-hours vs.OOH	16.5	0.9
b) weekdays vs. weekends	16	0.9
c) area of residence	14.5	1.4

The review group felt this point was well-made and that including *'24 hour emergency or urgent ambulatory care facilities'* in the original wording may have been too specific. The group decided to keep a structural measure but to adjust the wording to: *'proportion of population living within 10 km of emergency or urgent ambulatory care facilities open for more than 12 hours per day 7 days per week'*, see item 12, Table 3.

Based on the feedback from the Delphi panels the review group also decided to retain a process indicator based on '*emergency readmissions following discharge*', see item 6, Table 3, but adjust the period for readmission from 28 days to seven days. The justification for this decision was that, although the median score for this indicator was 5, the panel feedback suggested that the reason it had not been rated more highly (Box 1) may have been due to the period of 28 days used by the NHS quality regulator, the Healthcare Commission.[17]

The review group also made recommendations about the need for measures using serious emergency conditions to be reported separately for older (>75 years) and younger patients, and on the inclusion of a number conditions that had been proposed by the panels during the three rounds of consultation but which, because they were not included on the original list, had not been scored systematically by the panels.

Table 2 Delphi consensus of serious emergency and urgent conditions

Serious emergency conditions	Agreement	Disagreement
Myocardial infarction	18	0.6
Stroke/CVA	17	1.5
Meningitis	17	1.1
Cardiac arrest	17	1.1
Anaphylaxis	17	1.0
Major haemorrhage (e.g. ruptured aneurysm)	16.5	1.6
Fractured neck of femur	16	1.1
Asthma	16	1.2
Falls	16	1.2
Septic shock	15.5	1.5
Asphyxiation	15.5	1.5
RTA injuries	15	1.4
Acute heart failure	15	1.3
Isolated extradural haematoma	14.5	1.6
Childbirth	14	1.4
Exacerbation of COPD	14	1.2
Urgent conditions	**Agreement**	**Disagreement**
COPD	15.5	1.2
Hypoglycaemia	15.5	1.2
Urinary tract infection	15.5	1.6
Acute mental health crisis	15.5	1.1
Non-specific chest pain	15	1.2
Asthma	15	0.9
Elderly falls	15	1.1
Minor head injuries	14.5	1.3
Angina	14	1.1
Epileptic fit	14	1.1

Table 3 Set of performance indicators for the EUCS ($n = 16$)

Outcome based indicators

Mortality rates for serious, emergency, conditions for which a well-performing EUCS could improve chances of survival.

Case fatality ratios for serious, emergency conditions for which a well-performing EUCS could improve chances of survival.

Process based indicators

Hospital emergency admission rates for acute exacerbations of urgent conditions that could be managed out of hospital or in other settings without admission to in inpatient bed.

Arrivals at EDs referred by any EUCS services and discharged without treatment or investigations(s) that needed hospital facilities.

Arrivals at EDs referred by emergency ambulance and discharged without treatment or investigations(s) that needed hospital facilities.

Emergency readmissions within 7 days for serious, emergency or urgent conditions as a proportion of all live discharges.

Adherence to evidence-based good practice guidelines for serious, emergency and urgent conditions.

Multiple transfers between EUCS services.

Time from first contact with a EUCS service to clinical assessment.

For patients with serious, emergency conditions who are admitted to a hospital bed, time from first contact with a EUCS service to time of admission.

Time from first contact with EUCS service to definitive care.

Structural indicator

Proportion of population living within 10 kilometres of emergency or urgent care facilities open for more than 12 hours per day and 7 days per week.

Equity indicators

Relative case fatality rates between all contacts with EUCS services for serious emergency or urgent conditions, that die within 7 days, by defined comparators.

Variations in times from first call to any EUCs service for serious emergency or urgent conditions, and clinical assessment, by defined comparators.

Variations in times from first call for any EUCs service for serious emergency or urgent conditions, and admission, by defined comparators.

Variations in times from first call to definitive care for serious emergency, by defined comparators.

The adjustments made by the review group increased the number of indicators to be tested in the next phase of the study from 14 to 16, see Table 3. The final list of serious, emergency and urgent conditions consisting of all the conditions scoring ≥ 14 by the Delphi panels together with additional conditions proposed by the panellists and endorsed by the review group is shown in Table 4.

Discussion

Principal findings

Using a combination of Delphi and nominal group review, we have identified a set of indicators which a consensus of different experts considers to have the potential to be reliable and meaningful measures of the performance of EUCS. We also identified a core group of serious, emergency conditions, and urgent conditions, to use with the indicators.

Comparability with other findings

Delphi techniques have been used extensively in health services research to identify performance indicators for the quality of patient care in EDs,[18] cardiovascular disease[19] difficult asthma[20] and current treatment strategies for rheumatoid arthritis.[21] This is the first study that has used these methods to identify a set of indicators for the performance of EUCS. 'Ambulatory-care-sensitive' conditions for which hospital admissions may be prevented have been identified previously in the UK using consensus methods.[22] We were not looking at the ability of ambulatory care to prevent an acute episode of a disease in our study. Instead, our focus was to identify a set of measures with the potential to monitor the ability of systems of care to manage serious emergency and exacerbations of urgent conditions outside a hospital inpatient setting, in a timely and effective way.

Box 1 Excerpts from nominal group feedback on Delphi panels

'Much more important both politically and realistically – the true measure of a system is how well managed they are the first time hopefully avoiding likelihood of admission again.'

'Good indicator of failure to sort initial problem out.'

'Even better if 7 days, as many 28 days are because of disease rather than failed discharge.'

'Agree that re-admission in <7 days better measure, would have scored this 6.'

Table 4 Core set of serious, emergency, and urgent conditions to be tested for use with the performance indicators

Serious emergency conditions	ICD-10 codes
Stroke/CVA	I61; I63; I64; I629
Meningitis	G01; G02; G03; A39; A321
Anaphylaxis	T780; T782; T805; T886
Myocardial Infarction	I21; I22; I23
Fractured neck of femur	S72
Asthma	J45; J46
Cardiac arrest	I46
Pregnancy and birth related	O
Serious head injuries Glasgow coma score <9	S02; S03; S04; S05; S06; S07; S08; S09
Self harm	X6; X7; X8
Ruptured aortic aneurysm	I710; I711; I713; I715; I718
Falls <75 years	W0; W1
Road traffic accidents NEC	V0; V1; V2; V3; V4; V5; V6; V7; V802; V803; V804; V805; V821; V892; V830; V832; V833; V840; V841; V842; V843; V850; V851; V852; V853; V860; V861; V862; V863; V870; V871; V872; V873; V874; V875; V876; V877; V878
Septic shock	A40; A41
Asphyxiation	T71
Acute heart failure	I50
'Urgent conditions' *defined as the exacerbations of which could be managed by a well-performing EUC system without admission to an inpatient bed*	
COPD	J40; J41; J42; J43; J44
Acute mental health crisis	F
Non-specific chest pain	R072; R073; R074
Falls >74 years	W0; W1
Non-specific abdominal pain	R10
Deep vein thrombosis	I80; I81; I82
Cellulitis	L03
Pyrexial child <6 years	R50
Blocked tubes, catheters and feeding tubes	T830
Hypoglycaemia	E10; E11; E12; E13; E14; E15; E161; E162
Urinary tract infection	N390
Angina	I20
Epileptic fit	G40; G41
Minor head injuries	S00

Strengths and weakness of methods

The two Delphi panels were recruited purposively to achieve a broad range of expertise and perspectives in key aspects of emergency and urgent care, as well as different geographical areas in England and Wales. The electronic mailing of the survey was simple, efficient and cost-effective. There was no attrition between the rounds of consultation. The method allowed diverse views to be collected and fed back anonymously to the panels allowing each member to consider his or her initial score informed by the aggregate scores in their own time and space without the potential bias associated with group dynamics in face-to-face meetings.[23] In reaching a consensus, we used the median as the measure of agreement, and the mean absolute deviation (MAD-M) as the measure of dispersion. The range of the MAD-M was small indicating that the values in the set were clustered closely around the median increasing our confidence in the consensus about which items were likely to be good indicators of EUCS. A potential limitation of our study is that scoring each indicator in isolation did not yield a set of indicators that covered performance from all perspectives. The task of adjusting the final composition to ensure that the set of indicators to be tested in the next phase of the work conformed to pre-defined perspectives of outcome, process, structure and equity was provided by a nominal review group.

As a set, the indicators are generic 'context-free' and conform to the attributes associated with good performance indicators. The serious emergency and urgent conditions are 'adapters' to be used with the indicators to monitor the performance within and across general EUCS, and also specialized systems such as paediatrics or mental health. The list of conditions is not intended to be exclusive – other local and national priorities can be added. However, having a core list of conditions provides for the possibility of examining changes over time within systems routinely, and also for comparison of trends between systems.

The Delphi literature discourages administrators giving prescriptive feedback to the panels as this is associated with limiting the potential of other perspectives within the group.[12] The 'discussions' are facilitated by the comments of the panellists themselves so each participant can see where concerns are clustering and judge for themselves whether they are persuaded by the points made. We did not define 'the emergency and urgent care system'. Conceptually, the composition of the system will vary between different areas depending on how emergency and urgent care locally is configured and commissioned, for example, by PCTs or urgent care networks involving more than one PCT. We did ask the panels not to consider the availability of or feasibility of collecting the data required to calculate the indicators. This was because opinion of the potential of data systems to support decisions around the health care provision relies on different expertise than that for which the two panels were recruited. We also hope that what data are collected in the future may change in response to the need for accurate, reliable performance measures for systems of care.

The next steps

In the next phase of the development of a set of indicators to monitor the performance of systems of emergency and urgent care, we are addressing issues around the data required to calculate the indicators and testing their responsiveness to change over time in collaboration with four PCT populations in England.

Acknowledgements

The authors wish to thank all members of the panels for taking part in this study.

This work was undertaken by the Medical Care Research Unit which is supported by the Department of Health. The views expressed here are those of the authors and not necessarily those of the Department.

Conflict of interest: None.

References

1 Department of Health. *Reforming Emergency Care. First steps to a new approach*. London: Department of Health; 2001. See: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyand Guidance/DH_4008702
2 Locker T, Mason S, Wardrope J, Walters S. Targets and moving goal posts: changes in waiting times in a UK emergency department. *Emergency Medicine Journal* 2005; **22**: 710–14
3 Price L. Treating the clock and not the patient: ambulance response times and risk. *Quality and Safety in Health Care* 2006;**15**:127–30
4 Department of Health. *The NHS Plan: a plan for investment, a plan for reform*. Norwich: HSMO; 2000
5 Nielsen JD, Palshof T, Mainz J, Jensen AB, Olesen F. Randomised controlled trial of a shared care programme for newly referred cancer patients: bridging the gap between general practice and hospital. *Quality and Safety in Health Care* 2003;**12**:263–72
6 Julian S, Naftalin NJ, Clark M, Szczepura A, Rashid A, Baker R, *et al*. An integrated care pathway for menorrhagia across the primary-secondary interface: patients' experience, clinical outcomes, and service utilisation. *Quality and Safety in Health Care* 2007; **16**:110–15
7 O'Cathain A, Coleman P, Nicholl J. Characteristics of the emergency and urgent care system important to patients: a qualitative study. *Journal of Health Services Research and Policy* 2008;**13**:19–25
8 Preston C, Cheater F, Baker R, Hearnshaw H. Left in limbo: patients' views on care across the primary/secondary interface. *Quality in Health Care* 1999;**8**:16–21
9 Nicholl J, Coleman P, Knowles E, O'Cathain A, Turner J. *Medical Care Research Unit Programme 2006–2010. Emergency and Urgent Care systems. Final interim report to Department of Health*. Sheffield: University of Sheffield; 2009. See: http://www.shef.ac.uk/content/1/c6/05/91/04/final%20report.pdf
10 Fitch K, Bernstein SJ, Aguilar MS, *et al*. *The RAND/UCLA appropriateness method user's manual*. Santa Marica: RAND Corporation; 2001
11 Bell CM, Redelmeier DA. Mortality among patients admitted to hospitals on weekends as compared with weekdays. *N Engl J Med* 2001;**345**:663–8
12 Hsu CC, Sandford BA. The Delphi technique: making sense of consensus. *Practical Assessment, Research and Evaluation* 2007;**12**:1–8
13 Audit Commission. *On target: the practice of performance indicators*. London: The Audit Commission; 2000
14 Pencheon D. *The Good Indicators Guide: Understanding how to use and choose indicators*. Coventry: NHS Institute for Innovation and Improvement; 2008
15 Pringle M, Wilson T, Grol R. Measuring "goodness" in individuals and healthcare systems. *British Medical Journal* 2002;**325**:704–7
16 Bird SM, Cox Sir D, Farwell VT, Goldstein H, Holt T, Smith PC. Performance indicators: good, bad and ugly. *Journal of the Royal Statistical Society* 2005;**168**:1–27
17 Healthcare Commission. Performance indicators for the performance ratings 2004/2005. See: http://ratings.healthcarecommission.org.uk/indicators_2005/
18 Beattie E, Mackway-Jones K. A Delphi study to identify performance indicators for emergency medicine. *Emergency Medicine Journal* 2004;**21**:47–50
19 Normand SL, McNeil BJ, Peterson LE, Palmer RH. Eliciting expert opinion using the Delphi technique: identifying performance indicators for cardiovascular disease. *Int J Qual Health Care* 1998;**10**:247–60
20 Prys-Picard CO, Campbell SM, Ayres JG, Miles JF, Niven RM. Defining and investigating difficult asthma: Developing quality indicators. *Respiratory Medicine* 2006;**100**:1254–61
21 Wolfs JF, Peul WC, Boers M, van Tulder MW, Brand R, van Houwelingen HT, *et al*. Rationale and design of The Delphi Trial-I(RCT)2: international randomized clinical trial of rheumatoid craniocervical treatment, an intervention-prognostic trial comparing 'early' surgery with conservative treatment [ISRCTN65076841]. *BMC Musculoskeletal Disorders* 2006;**7**:14
22 Sanderson C, Dixon J. Conditions for which onset or hospital admission is potentially preventable by timely and effective ambulatory care. *Journal of Health Services Research and Policy* 2000;**5**:222–30
23 Mitroff II, Turoff M. Philosophical and methodological foundations of Delphi. In: Linstone HA, Turoff M, Eds. *The Delphi Method: Techniques and Applications*. Reading, MA: Addison-Wesley Publishing Co; 1975:17–35