

about one quarter of ambulance crews in London are paramedics, but analysis of workload suggests that less than one tenth of calls require paramedical skills. The London Ambulance Service already trains more paramedics than any other service in Britain and loses some of them to provincial services after training.

Only one third of hospitals served by the London Ambulance Service provide clinical training placements for paramedics, and even to meet current needs this will require a substantial increase. To achieve the Department of Health's target will require a substantial increase in training at considerable cost. Is this likely to be economic? How will most paramedics maintain their skills when only a small proportion of calls require them?¹⁰ Will the damage to the morale of staff who are not paramedics, seen in other services with the development of an elite, occur in London? The London Ambulance Service has avoided many of these problems by ensuring that some extended skills, such as provision of defibrillation and nebuliser treatment, are taught to every qualified ambulance person. Evidence also exists that ambulance staff who are not paramedics can safely identify and treat hypoglycaemia with glucagon.¹¹ If implementing a dispatch system responsive to medical priority shows that resources are being effectively targeted according to clinical need the aim of having a paramedic in every front line vehicle would need to be reconsidered.

Historically, informal links with individual clinicians and specialty groups have facilitated many changes in ambulance training, operations, and equipment. As the range and complexity of prehospital care increases, however, the mechanisms for medical input to the London Ambulance Service and the audit of outcomes will need review. Currently, there is official medical input to extended training in the form of the paramedic steering committee as required by the NHS Training Directorate, but in other matters of policy the only other medical forum, the medical advisory group, has no operational responsibilities. Recent initiatives that have required close liaison are the development of joint training for ambulance and medical incident officers in the management of major incidents,¹² proposals for increasing direct communication between ambulances and accident and emergency departments, and the development of alternatives to obstetric flying squads.¹³ The input of the medical advisory group to long term planning and development would be valuable.

Unified medical input and systematic audit¹⁴ are particularly necessary in view of the differing medical opinions and prejudices regarding the philosophies of "scoop and run"

versus "treat in the street." How best to coordinate medical advice for the London Ambulance Service should be a matter for debate. In particular, the potential benefit to London of the consultant medical director model—popular in the United States and currently being pioneered in the Scottish Ambulance Service—will need careful assessment.

Public interest in the London Ambulance Service is obvious, but public confidence, shaken by the 1989 ambulance strike and by the problems of last year, needs to be restored. Although response times have the highest profile—being easily measured—the quality of clinical care delivered is also crucially important, and its absence from the topics included in the service's annual corporate review is regrettable.¹⁵

The London Ambulance Service's new management must make an unequivocal commitment to both basic and extended training and the establishment of systematic clinical audit to build on the progress made to date. As a consumer, the public has a part to play in making the service more efficient, and a health education campaign supported by the government is long overdue. Management commitment is likely to produce better response times over the next three years, but this will be a hollow achievement without advances in clinical care.

ROBERT A COCKS

Consultant in accident and emergency

Hammersmith Hospital,
London W12 0HS

EDWARD GLUCKSMAN

Consultant in accident and emergency

King's College Hospital,
London SE5 9RS

- 1 Warden J. London's ambulance service fails. *BMJ* 1992;305:1116.
- 2 South West Thames Regional Health Authority. *Report of the inquiry into the London ambulance service*. London: Communications Directorate, SWT RHA, 1993.
- 3 London ambulance service board scrapped. *BMJ* 1993;306:950.
- 4 Department of Health. *ORH good practice guide*. London: DoH, 1992. (EL 92(19).)
- 5 Gardner GJ. The use and abuse of the emergency ambulance service: some of the factors affecting the decision whether to call an emergency ambulance. *Arch Emerg Med* 1990;7:81-9.
- 6 Morris DL, Cross AB. Is the emergency ambulance service abused? *BMJ* 1980;281:121-3.
- 7 Clawson JJ, Hauert S, Martin RL. *Advanced EMD course manual*. Salt Lake City: Medical Priority Consultants, 1992.
- 8 Kellerman AL, Hackman BB, Somes G. Dispatch-assisted cardiopulmonary resuscitation—validation of efficacy. *Circulation* 1989;80:1231-9.
- 9 Department of Health. *The health of the nation: key area handbook: accidents*. London: DoH, 1993.
- 10 Walters G, Glucksmann E. Retention of skills by advanced trained ambulance staff: implications for monitoring and retraining. *BMJ* 1989;298:649-50.
- 11 Steel JM, Allwinkle J, Moffat R, Carrington DJ. Use of Lucozade and glucagon by ambulance staff for treating hypoglycaemia. *BMJ* 1992;304:1283-4.
- 12 Cocks R. On site medical services at major incidents. *BMJ* 1992;305:1015.
- 13 Royal College of Obstetricians and Gynaecologists. *The future of emergency domiciliary obstetric services ("flying squads")*. London: RCOG, 1990.
- 14 Polsky SS, Weigand JV. Quality assurance in emergency medical service systems. *Emerg Med Clin North Am* 1990;8:75-84.
- 15 South West Thames Regional Health Authority. *Revised accountability arrangements for the London Ambulance Service*. London: Communications Directorate, SWTRHA, 1993.

Measuring patients' views: the optimum outcome measure

SF 36: a valid, reliable assessment of health from the patient's point of view

The debate about whether and how to measure medical outcomes is heating up as those who fund medical care put providers under greater pressure to control their costs. With little known about which features of medical services improve outcomes, policies that reduce costs will be deemed successful unless data on outcomes show otherwise. Hence the need to study outcomes.

The ideal outcome of treatment is a return to the normal or usual quality of life for a given age and medical condition.¹ Since the first health interview surveys of sickness and disability, the hunt has been on for a reliable measure of how people perceive their health.² Great strides have recently been

made in standardising self reported measures of functioning and wellbeing and using them to monitor outcomes. With the standardised self report survey emerging as the best method of measuring outcomes from the patient's point of view, now is a good time to take stock of these advances.

To date the evaluation of such surveys has mainly been controlled by the researchers who developed them, but health care policy makers, clinical investigators, and providers should now get involved. The contributions of Jenkinson and colleagues (p 1437)³ and Garratt and colleagues (p 1440)⁴ in this week's journal and an earlier contribution by Brazier *et al*⁵ provide useful information about the performance of the

American short form (SF 36) health survey questionnaire adapted for British use. These papers show the ease of administering this self report survey, and suggest that other English speaking countries such as Australia and Canada may be able to use it with minimal changes.

These papers also show the value of two different strategies for interpreting health status—approaches based on norms and criteria.^{3,4} Norms allow individual scores and group averages to be interpreted according to where they lie in the distribution of scores for a general population. They are particularly useful when no single criterion or “gold standard” for evaluating health exists. Because men and women of different ages report different levels of health, separate norms can be calculated for these groups and for different occupational groups.³ These norms will greatly help to estimate the burden of specific conditions. The inclusion of criterion variables such as medical conditions, severity of illness, and use of health care services is useful in interpreting the surveys.^{6,7,8}

Garratt and colleagues show that SF 36 scores tend to correlate with doctors’ assessment of the severity of illness, which is encouraging. But this holds true only on average: functioning and wellbeing vary greatly even among patients with the same condition.^{6,7} Some of this variation may be due to treatment and to factors that change over time. Follow up studies planned by Garratt and colleagues and more widespread use of self report measures in clinical trials should improve our understanding of why outcomes vary and how to improve them.

After extensive evaluations of the content of the SF 36, it is encouraging that these investigations conclude that only minor changes in six items were required to adapt the SF 36 for British use.^{3,5} For example, “walking one block” was changed to “walking 100 yards” and “full of pep” was changed to “full of life.” These changes are unlikely to affect the reliability and validity of the scores. When it comes to the grouping of items and tests of scoring assumptions British research seems to replicate the findings of American research.^{6,8-10} Given the overall consistency of the British and American results, other findings in one country are likely to generalise to the other.⁸

Although simple and brief self report surveys risk a loss of precision and comprehensiveness, studies suggest that the substantial reduction in the burden they placed on respondents is accompanied by only a small drop in precision.^{8,10} In clinical practice, however, where decisions are based on assessments of individual patients, even a small drop in the reliability of the score greatly increases the “noise level”: less reliable surveys will require larger changes in patient scores to register. To guide clinical use, confidence intervals for the SF 36 scales have therefore been estimated for individual patients in the United States.⁸ Another trade off affects comprehensiveness: short forms require that some concepts are left out. But, as the authors of the British papers show, a “core” survey that requires only 5 to 10 minutes to complete provides ample room for questions to be added.

Whether favourable estimates of reliability and validity from cross sectional studies will hold true for longitudinal studies remains to be proved. According to psychometric theory, measures that discriminate well among different groups of patients at one particular time tend to do well in capturing changes over time. Evidence supporting this has been found for the SF 36.^{8,11} Changes in SF 36 scores in one group over time can usually be measured with greater precision than differences between two groups.⁸

To protect the standardisation of the SF 36, ensure comparability across studies and countries, and assure its widespread availability without any fee the copyright for

SF 36 and the trademark were placed with the Medical Outcomes Study Trust, a non-profit making organisation.¹² The trust ensures that the “SF 36” label is used only for forms and scoring algorithms that are comparable to the published SF 36 standard. Those who depart from the standard are required to use a different label so that others will know that data may not be comparable. An international team of investigators is currently developing authorised translations of the SF 36 through the international quality of life assessment project.¹³ Sponsors of the project, including Glaxo and Schering-Plough Research, have agreed to place all translations with the Medical Outcomes Study Trust along with the British version of the SF 36 described here.^{3,5}

There has been considerable debate over the choice between disease specific and generic measures of outcome in clinical trials and studies of cost effectiveness, and numerous research teams funded by the American government are administering both. In North America the SF 36 is currently being used in more than 200 clinical trials, where investigators are interested in the impact of treatment on the quality of life as well as on more traditional medical outcomes. Measures of generic and specific outcomes usually prove to be more useful than either alone.

Much cooperation is needed to define and preserve a measurement standard and to advance understanding of the scores. The challenge is greatest when many groups are involved and funding resources limited. In the United States development of measures in the medical outcomes study began with support from philanthropic sources, which have been joined by several federal agencies.¹⁴ In Britain the King’s Fund has funded studies with the SF 36, including those by the Public Health Research Centre in Salford, the Department of Public Health in Nottingham, and the North West Thames Regional Health Authority. The contributions by Jenkinson and colleagues and Garratt and colleagues in this week’s journal mark an important step forwards in monitoring and improving outcomes.

JOHN E WARE
Senior scientist

Health Institute,
New England Medical Center,
750 Washington Street,
Boston, MA 02111,
USA

- 1 Silver GA. Paul Anthony Lembecke, MD, MPH: A pioneer in medical care evaluation. *Am J Psych* 1990;80:342-8.
- 2 Katz S, Akpom CA, Papsidero JA, and Weiss ST. Measuring the health status of populations. In: Berg RL, ed. *Health status indexes*. Chicago: Hospital Research and Educational Trust, 1973.
- 3 Jenkinson C, Coulter A, Wright L. Short form 36 (SF 36) health survey questionnaire: normative data for adults of working age. *BMJ* 1993;306:1437-40.
- 4 Garratt AM, Ruta DA, Abdalla MI, Buckingham JK, Russell IT. The SF 36 health survey questionnaire: an outcome measure suitable for routine use within the NHS? *BMJ* 1993;306:1440-4.
- 5 Brazier JE, Harper R, Jones NMB, O’Cathain A, Thomas KJ, Usherwood T, et al. Validating the SF 36 health survey questionnaire: new outcome measure for primary care. *BMJ* 1992;305:160-4.
- 6 McHorney CA, Ware JE, Raczek AE. The MOS 36-item short-form health status survey (SF 36). II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care* 1993;31:247-63.
- 7 Stewart AL, Greenfield S, Hays RD, Rogers WH, Berry SD, McGlynn EA. Functional status and well-being of patients with chronic conditions: results from the medical outcomes study. *JAMA* 1989;262:914-9.
- 8 Ware JE, Snow KK, Kosinski M, Gandek B. *SF 36 Health survey manual and interpretation guide*. Boston, MA: New England Medical Center, Health Institute, 1993.
- 9 McHorney CA, Ware JE, Lu JFR. The MOS 36-item short-form health survey (SF 36). III. Tests of data quality, scaling assumptions and reliability across diverse patient groups. *Med Care* (in press).
- 10 McHorney CA, Ware JE, Rogers W, Raczek A, Lu JFR. The validity and relative precision of MOS short and long form health status scales and Dartmouth COOP charts: results from the medical outcomes study. *Med Care* 1992;30:MS253-65.
- 11 Katz JN, Larson MG, Phillips CB, Fossel AH, Liang MH. Comparative measurement sensitivity of short and longer health status instruments. *Med Care* 1992;30:17-25.
- 12 Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF 36). I. Conceptual framework and item selection. *Med Care* 1992;30:473-83.
- 13 Aaronson NK, Acquadro C, Alonso J, Apolone G, Bucquet D, Bullinger M, et al. International quality of life assessment (IQOLA) project. *Quality of Life Research* 1992;1:349-51.
- 14 Stewart AL, Ware JE, eds. *Measuring functioning and well being: the medical outcomes study approach*. Durham, NC: Duke University Press-Raleigh, 1992.