their survey confirms that the reliability of the SF 36 has remained intact.[1] This surely depends on the interpretation of the term and the methods used. The reliability of an instrument is the degree to which it measures something reproducibly and consistently. Garratt and colleagues use internal consistency as their measure of reliability. Streiner and Norman point out that this should be interpreted with caution as an indication of reliability because it ignores potentially important sources of variation that can occur over time.[4]

When a self rating test is used the usual way of assessing reliability is to administer the test to the same people on at least two occasions a short time apart to avoid appreciable changes in their health status; this gives the test-retest reliability. As Garratt and colleagues' paper is meant to be assessing the SF 36 as an outcome measure, so that the score would be measured before and after treatment in a context in which change in the scores could be attributed to the intervention, the reliability of the SF 36 over time is important. If the scores vary greatly over time in the absence of changes in underlying health it will be harder to identify change due to treatment, the precision of estimated effects of treatment will be reduced, and so larger sample sizes will be required.

Most of the published assessments of reliability have been based on measures of internal consistency. Correlations between scores at various times can be misleading, but only one of the studies that assesed test-retest reliability used the more appropriate method of Bland and Altman, which plots the difference in score against the average of the two readings. Brazier *et al* showed that in the general population the limits of agreement between test and retest values two weeks apart are quite large.[5] For example, in physical functioning the limits of agreement ranged from −9 to 10, a range large enough to span a clinically important change in morbidity. It is a pity that Garratt and colleagues did not use this opportunity to report similar tests in morbid populations.

The reliability of the instrument must be examined properly in a range of clinical and organisational contexts. Until these results are available the instrument's statistical properties as an outcome measure remain uncertain.

TREVOR SHELDON

Centre for Health Economics,
University of York,
York YO1 5DD

1 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. (29 May.)
2 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. (29 May.)
3 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.
4 Streiner DL, Norman GR. *Health measurement scales.* Oxford: Oxford University Press, 1989.
5 Brazier J, Harper R, Jones N, Thomas K, Westlake L. Validating the SF 36. *BMJ* 1992;**305**:646.

Definitions of health status should be rationalised

EDITOR,—In their paper on the short form 36 (SF 36) health survey questionnaire Andrew M Garratt and colleagues conclude that a comprehensive portfolio of outcome measures may be needed.[1] The first issue of *Outcomes Briefing*, a quarterly newsletter from the United Kingdom Clearing House on Health Outcomes, contains a typology that might be used to classify such a portfolio and illustrates the difficulties and inconsistencies that might arise.[2]

The typology mixes concepts such as the World Health Organisation's classic definition of health (physical, social, and mental wellbeing); the international classification of impairment, disability, and handicap[3]; and quality of life related to health. Some aspects of health are wrongly placed: impairment should appear under "mental" as well as "physical"; disability (the impact of disease on functional performance) is inappropriately placed under "social"; and disease specific measures should not be confined to "physical" as some deal with mental problems and may also tackle social aspects of disease. Functional status and disability both appear; for some diseases these two are nearly indistinguishable, while for others functional status means handicap. Finally, the inclusion of measures of both quality of life and multidimensional health status may create confusion; many measures of quality of life are measures of multidimensional health status rather than measures of satisfaction with life or wellbeing.

The most appropriate model for outcomes of disease would permit measurements that are likely to be affected by interventions. WHO's definition of health is too broad. Measures of quality of life may not be useful for more specific interventions. We suggest that the international classification of impairment, disability, and handicap should be used as it has several advantages. It allows matching of the expected level at which an intervention works and the outcome used. It can be used for psychiatric and physical illness. It avoids the need for value judgments about the importance of different consequences of disease and highlights the importance of considering a range of outcomes of treatment systematically and specifically—for example, in heart failure only weak correlations exist between cardiac output (impairment), exercise tolerance (disability), and customary walking (an aspect of handicap).[4] Finally, it promotes clearer recognition of the impact of disease on patients, which may facilitate the delivery of care that matches their goals.

We hope that the Outcomes Clearing House will consider revising its typology to take this proposal into account. It is important that the context within which a portfolio of outcome measures is developed and used is clear.

EDWARD DICKINSON
PATRICK GOMPERTZ
ROWAN HARWOOD
SHAH EBRAHIM

Department of Public Health and Primary Care,
Royal Free Hospital School of Medicine,
London NW3 2PF

1 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. (29 May.)
2 Issues in outcome measurement. *Outcomes Briefing* 1993;spring: 9-13.
3 World Health Organisation. *International classification of impairment, disability, and handicap.* Geneva: WHO, 1980.
4 Cowley AJ, Fullwood LJ, Muller AF, Stainer K, Skene AM, Hampton JR. Exercise capacity in heart failure: is cardiac output important after all? *Lancet* 1991;**337**:771-3.

The term outcome is ambiguous

EDITOR,—Measurement of outcomes is, as John E Ware implies in his editorial, an important consideration in evaluations of the quality of care.[1] Considerable ambiguities, however, hinder understanding of this important topic.

A key ambiguity is what constitutes an outcome. Ware and others seem to use the term to refer to the effects of health care. A further major factor influencing outcome, however, is the natural course of disease. Ware implicitly acknowledges the relation between outcome of treatment and a patient's medical condition. The most relevant measure of quality of care is the difference (at any time point) between the natural course of the disease and the course of disease with treatment —the net treatment effect. We are aware that separating the influences on health state of the natural course of disease and of the net treatment effect is difficult owing to a lack of detailed knowledge of the natural course of many diseases. To the degree that the natural course of a disease remains undefined, however, the component of outcome attributable to health care remains uncertain.

A second ambiguity concerns measurement of health state without a definition of when the measurement is made. The time point(s) and intervals at which health state is measured should be explicit and appropriate to the natural course of the disease in question. Thus the natural course of disease is of major importance not only in being a considerable component of the health state at any given time but also in influencing the time point(s) at which it is appropriate to assess health state.

A third uncertainty is the outcome to be aimed for. Ware suggests that the ideal outcome is "a return to the normal or usual quality of life for a given age and medical condition." Such an ideal outcome may be bettered—for example, after cardiac transplantation a patient might exceed his or her previous usual quality of life. A more relevant outcome is that which is realistically attainable given current technology and available resources. That realistically attainable health state should be the explicit objective of high quality care at institutional level, in setting contracts, and in the care of individual patients. Making explicit the objectives of treatment would have advantages in focusing postgraduate education, improving clinical audit, developing and refining clinical guidelines, and ensuring continuity of care after the reduction in junior doctors' hours of work. We suggest that explicit declaration of the objectives of treatment should form part of the routine patient record and should be seriously considered for incorporation in the electronic patient record project of the NHS information management and technology strategy.[2]

We suggest that people should stop using the term outcome since outcome is an ambiguous, composite measure. More helpful terms would be "health state" or "health status" to refer to a patient's state of health at a given time point and "net treatment effect" to refer to that component of the patient's health state that may be attributed to treatment. "Health gain" is less desirable since it presupposes that health care will always benefit a patient's health state, which is not the case. Clear thinking in assessment of the quality of health care is likely to improve patients' care by improving targeting of scarce resources to areas where the greatest net treatment effect is achievable.

J KILGOUR-CHRISTIE
A H WATT

Medicines Assessment Research Unit,
University of Aberdeen,
Aberdeen AB9 2ZD

1 Ware J. Measuring patients' views: the optimum outcome measure. *BMJ* 1993;**306**:1429-30. (29 May.)
2 Information Management Group, NHS Management Executive. *An information management and technology strategy for the NHS in England. Facilitating person-based information systems; an overview.* London: Department of Health, 1993.

SF 36 is suitable for elderly patients

EDITOR,—In their paper reporting normative data for the short form 36 (SF 36) health survey questionnaire for adults of working age[1] Crispin Jenkinson and colleagues repeat doubts previously expressed by Brazier *et al* about the suitability of this instrument for measuring health status in elderly populations.[2] Mid Downs Health Authority and Crawley and Horsham NHS Trust recently carried out a postal survey in which the SF 36 was used as the basis of the questionnaire, and our experience does not support that view.

The family health services authority's population register was used to select 3600 people aged 65 and over. Considerable attention was paid to the layout of the questionnaire, which had large type to aid visually impaired people. It was also made clear that help could be obtained to complete the

questionnaire if necessary, and in 12% of cases someone completed the questionniare on the recipient's behalf. The response rate thus achieved (78% overall and 85% when those found to be ineligible were excluded) was better than that reported for general population samples.[1-3] The age and sex distributions of the respondents, when compared with census data, showed no evidence of decreased response with increasing age—indeed, the reverse was the case.

The reliability of the responses was as good as that reported previously,[1-3] as indicated by the values for Cronbach's α obtained for the different scales in the SF 36. In addition to the SF 36's core questions the postal questionnaire contained some questions on help received with activities of daily living and on visits from the general practitioner and district or practice nurses in the past three months. The level of support indicated by responses to these questions showed a strong relation with severity based on scores on the scale of physical functioning (table), suggesting that this particular scale is valid. An interview survey of a subsample of the respondents to the postal survey is nearing completion and should provide much more information on the validity of the instrument for this age group.

Amount of help received with activities of daily living related to severity of physical problems (according to score on SF 36's scale of physical functioning). Figures are numbers (percentages) of respondents

	Problems with physical function				
Help received	Very severe	Quite severe	Moderate	Minor	Not significant
None	18 (6)	112 (29)	328 (54)	627 (75)	500 (89)
Occasional	63 (20)	66 (17)	107 (18)	103 (12)	35 (6)
Regular	239 (75)	204 (53)	171 (28)	111 (13)	27 (5)
Total	320	382	606	841	563

There are occasions when a well tested instrument of proved reliability and validity across a wide age range provides a core of basic questions. The SF 36 seems to be a good candidate in this respect, although more work, such as the interview survey we are conducting, is needed to prove its validity. Its link to a range of questionnaires designed to measure outcomes for a variety of specific conditions—for example, stroke—is an added advantage.

NICOLA SINGLETON
ANN TURNER

Research Unit,
Department of Public Health,
Mid Downs Health Authority,
Haywards Heath,
West Sussex RH16 4BE

1 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. (29 May.)
2 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.
3 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. (29 May.)

SF 36 may reinforce ageism

EDITOR,—I wish to add a caution regarding use of the short form 36 (SF 36) health survey questionnaire.[1] This "optimum outcome measure" has the potential to bias allocation of resources away from elderly people.

At a given level of disability patients' satisfaction improves with increasing age and duration of disease.[2] Thus subjective components of the SF 36, such as its "general health perception" and "mental health" indices, tend to underrate impairment in elderly or chronically sick people. Although it is recognised that elderly people play down their symptoms, this effect may be forgotten in a complex outcome statistic, which, as a result, emphasises the needs of other groups.

Use of normative data (which unfortunately do not yet include data on elderly people)[3] is imperative to ensure that comparisons of outcome are valid and that the SF 36 cannot be used to justify and reinforce ageism.

ANTONY JOHANSEN

Whipps Cross Hospital,
London E11

1 Ware JE. Measuring patients' views: the optimum outcome measure. *BMJ* 1993;**306**:1429-30. (29 May.)
2 Ross CK. The role of expectations and preferences in health care satisfaction of patients with arthritis. *Arthritis Care and Research* 1990;**3**(2):92-8.
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. (29 May.)

Diagnosing meningococcal infection

Don't delay giving antibiotics

EDITOR,—I was intrigued by Marcel van Deuren and colleagues' study of aspiration or biopsy of skin lesions as a diagnostic procedure in acute meningococcal infection.[1] The data in their table I confirm the common clinical impression that not all cases of acute meningococcal infection will be confirmed bacteriologically on culture of blood or cerebrospinal fluid, and this cannot always be accounted for by prior antibiotic treatment.

It seems that patients with negative results of culture of blood and cerebrospinal fluid were included in the study if meningococci were discovered on examination of skin lesions and excluded otherwise. The study population was thus partly defined by the result of this technique, and this in turn will have spuriously increased the technique's apparent sensitivity. This distortion is liable to be stronger still in the subgroup in which Gram staining of cerebrospinal fluid initially gave an inconclusive result. While it is clearly difficult to calculate sensitivity when a true, independent denominator is not available, it might have been helpful if the authors had indicated how many patients had episodes of clinical meningococcal infection in which all available bacteriological techniques yielded negative results during the study.

The authors state that in meningococcal infection a prompt bacteriological diagnosis is one of the prerequisites for saving lives. I suggest that for patients with clinical sepsis and haemorrhagic lesions the main prerequisite is the prompt institution of vigorous parenteral antibiotic treatment with supportive measures, and this should not be deferred pending bacteriological confirmation.

GEORGE FARMER

Stirling Royal Infirmary NHS Trust,
Stirling FK8 2AU

1 Van Deuren M, van Dijke BJ, Koopman RJJ, Horrevorts AM, Meis JFGM, Santman FW, *et al.* Rapid diagnosis of acute meningococcal infections by needle aspiration or biopsy of skin lesions. *BMJ* 1993;**306**:1229-32. (8 May.)

Authors' reply

EDITOR,—We agree with George Farmer that if meningococcal sepsis is suspected the main prerequisite for saving life is the prompt institution of antibiotics with supportive measures. Early confirmation of the diagnosis, however, helps doctors to estimate the prognosis, anticipate clinical deterioration, and start chemoprophylaxis for the family. One of the important messages of our paper was that early treatment with antibiotics does not affect the bacteriological diagnosis when it is based on examination of skin lesions.

We compared the contributions of Gram stained specimens of cerebrospinal fluid and skin lesions to rapid confirmation of the diagnosis. The word "sensitivity" was used to indicate the proportion of patients with a positive test result in our study population—that is, in patients with proved meningococcal infection. We would be reluctant to classify patients with disease that responds to antibiotics but with negative bacteriological results as having meningococcal infections. Thus we did not study the value of the tests in these patients.

The comment that the study population was partly defined by the result of the tests and that it is difficult to calculate a sensitivity when there is no independent denominator is theoretically correct. In our opinion, however, in patients in whom meningococcal disease is suspected clinically the finding of Gram negative cocci in a skin lesion proves the diagnosis of meningococcal infection even when cultures of cerebrospinal fluid or blood remain negative. Moreover, in only two of the 51 patients was the diagnosis made exclusively on the basis that Gram staining of a skin lesion gave a positive result; in the subgroup of patients with shock and without meningitis (n=18) this occurred in only one patient. When this patient was excluded Gram staining of a skin lesion in this group still gave a positive result in 70%, whereas Gram staining of cerebrospinal fluid gave a positive result in only 23%.

MARCEL VAN DEUREN
ROLAND J J KOOPMAN
ALFONS M HORREVORTS
JACQUES F G M MEIS
FREDERIK W SANTMAN
JOS W M VAN DER MEER

University Hospital,
PO Box 9101,
6500 HB Nijmegen,
Netherlands

BEATRIX J VAN DIJKE

Laurentius Hospital,
Roermond,
Netherlands

Medical angiologists

Britain needs them

EDITOR,—As members of the steering committee of the forum on angiology of the Royal Society of Medicine, we welcome C P Warlow's challenge to prove that medical angiologists are needed in the United Kingdom.[1] In fact, he highlights the case for medical angiologists in the prevention and treatment of venous thromboembolism and peripheral arterial disease. Venous thromboembolism is the commonest preventable cause of death in hospital, and because three quarters of fatal pulmonary emboli occur in medical wards it seems appropriate to designate a physician to coordinate and audit prevention and treatment,[2] which, as Warlow states, is almost entirely medical. Purchasers will not be happy if 1% of their referred patients die of pulmonary embolism through lack of organised prophylaxis,[2] and hospital managers will face potential litigation costs.

Symptomatic peripheral arterial disease is present in almost 5% of the population of the United Kingdom aged 55-75, but these subjects make up only a small proportion of those with detectable atherosclerosis.[3] Warlow appreciates the need for medical management of the one million people in the United Kingdom who have claudication: few will have surgery, but most require medical assessment and prophylaxis to reduce their major risk of heart disease and stroke. Several of us work happily with our vascular