Special Article

Measuring quality of life in clinical trials: a taxonomy and review

Gordon H. Guyatt, MD
Sander J.O. Veldhuyzen Van Zanten, MD
David H. Feeny, PhD
Donald L. Patrick, PhD, MSPH

Measurement of quality of life is becoming increasingly relevant to controlled clinical trials. Two basic types of instrument are available: generic instruments, which include health profiles and utility measurements based on the patient's preferences in regard to treatment and outcome; and specific instruments, which focus on problems associated with individual diseases, patient groups or areas of function. The two approaches are not mutually exclusive; each has its strengths and weaknesses and may be suitable under different circumstances. We surveyed 75 randomized trials published in three medical journals in 1986 and categorized them according to the importance of quality of life as a measure of outcome and the extent to which quality of life was actually measured. Although a number of the investigators used quality-of-life instruments in a sophisticated manner, in only 10 of 55 trials in which the measurement had been judged to be crucial or important were instruments with established validity and responsiveness used. We conclude that although accurate measurement of quality of life in randomized trials is now feasible it is still not widely done. Using the framework we have outlined, investigators can choose generic or specific instruments according to the purpose and the focus of their trial.

From the departments of Clinical Epidemiology and Biostatistics, Medicine and Economics (Centre for Health Economic and Policy Analysis), McMaster University, Hamilton, Ont., and the Department of Health Services, School of Public Health and Community Medicine, University of Washington, Seattle

Dr. Guyatt is a Career Scientist of the Ontario Ministry of Health.

Reprint requests to: Dr. Gordon H. Guyatt, Department of Clinical Epidemiology and Biostatistics, Rm. 3H7, McMaster University Health Sciences Centre, 1200 Main St. W, Hamilton, Ont. L8N 3Z5

Il devient de plus en plus impérieux, dans les essais cliniques comparatifs, de savoir déterminer la qualité de la vie. On dispose à cette fin de deux catégories de moyens. Les moyens génériques comprennent les profils sanitaires et les échelles dites d'utilité qui reflètent les préférences du malade quant au traitement et à son issue. Les moyens spécifiques sont axés sur les difficultés que présentent telles maladies, tels groupes de malades, telles fonctions. Ces deux groupes de moyens ne s'excluent pas l'un l'autre; chacun possédant ses avantages et ses lacunes, il peut être précieux selon les circonstances. Nous passons en revue 75 essais comparatifs sur des sujets désignés au hasard, parus en 1986 dans trois revues médicales, en cherchant à savoir quelle importance on y a accordée à la qualité de la vie dans le jugement des résultats et jusqu'à quel point cette qualité a bien été déterminée. Beaucoup des chercheurs se servent à bon escient des méthodes dont nous avons parlé. Mais parmi les 55 travaux où l'on reconnaissait l'importance plus ou moins grande de cette détermination, dans 10 seulement a-t-on eu recours à des méthodes dont la validité et la sensibilité avaient déjà été démontrées. Nous croyons donc qu'en dépit de sa possibilité la détermination précise de la qualité de la vie ne se fait pas souvent dans ce genre d'essai. Le cadre que nous proposons permet au chercheur de choisir les moyens convenant aux buts de son essai et aux aspects qu'il y privilégie.

During the last decade the importance of measuring aspects of health status related to patients' function and subjective experience has become increasingly recognized. "Quality of life" has appeared as a label for the measurement of physical and emotional function.[1] Of course, quality of life is influenced by many factors other than one's health, but health researchers are interested in health-related quality of life. In this

review we use the term to refer to the wide variety of subjective experiences related to health, such as symptoms, physical function and emotional function.

The focus of interest may be the impact of a disease or condition on quality of life,[2,3] the profile of dysfunction in a population[4] or the relation between quality of life and prognosis.[5] For clinicians the impact of medical interventions on how patients feel and how they function is a crucial area. Readers of clinical journals are starting to see trials in which quality of life is the primary outcome.[6-8]

In this article we suggest a taxonomy for quality-of-life measures, review the approaches to measurement in clinical trials and their relative merits, and survey the use of measures in recent trials. Our discussion, which is built on the contributions of previous authors,[1,9-14] focuses on the empirical performance of quality-of-life measures in clinical trials. We present the taxonomy because we have found it to be a useful conceptual model for those beginning to learn about the issues in measurement of quality of life. A summary of the strengths and weaknesses of the types of measures is presented in Table I.

Necessary attributes of quality-of-life measures

Three attributes are essential for any instrument to be useful as a measure of outcome in clinical trials: reproducibility, validity and responsiveness.

Reproducibility

A measure is reproducible if it yields the same results when repeated in stable subjects. Reproducibility is best measured by repeated administration of an instrument to subjects whose status has not changed.

Validity

The instrument must be measuring what it is supposed to measure.[15-17] Because there is no gold standard for measuring quality of life the validity is established by specifying the domain or dimension to be measured and the expected relations between that domain and other variables.

Responsiveness

Investigators want to detect any clinically important changes in quality of life, even if they are small. Responsiveness refers to the instrument's ability to do this and is determined by two properties:[18] reproducibility and changeability (i.e., it must register changes in score when a subject's quality of life improves or deteriorates). If the responsiveness is unproved and the results of a controlled trial in which the instrument was used are negative, either the treatment is not effective or the instrument is not responsive. Thus, in the beginning of a trial a questionnaire that has been proven to be responsive in previous related investigations should be used.

Taxonomy

Generic instruments

Generic instruments are applicable in a wide variety of populations because they cover the complete spectrum of function, disability and distress that is relevant to quality of life. They can be divided into two major classes: health profiles and utility measures.

Health profiles: These are single instruments that measure different aspects of quality of life. They share a scoring system and can be aggregated into a few scores or sometimes even one (an index). As generic measures health profiles are

Table I — Strengths and weaknesses of available measures of quality of life in clinical trials

Measure	Strengths	Weaknesses
Generic instruments*		
Health profile	Is a single instrument	May not focus adequately on area of interest
	Has established reliability and validity	May not be responsive
	Detects differential effects on different aspects of health status	
	Allows comparison between interventions or conditions	
Utility measurement	Provides single number representing net impact on quality of life	May involve difficulty in determining utility values
	Allows cost–utility analysis	Does not allow examination of effect on different aspects of quality of life
		May not be responsive
Specific instruments†	Are clinically sensible	Do not allow comparison between conditions
	May be more responsive than generic instrument	May be limited in terms of populations and interventions

*Applicable to a wide variety of populations.
†Focus on problems associated with specific diseases, patient groups or areas of function.

designed for use in a wide variety of conditions. For example, one of the most popular health profiles, the Sickness Impact Profile (SIP),[19] contains 12 categories, which can be aggregated into two dimensions and five independent categories as well as into a single overall score. The SIP has been used in studies of cardiac rehabilitation,[20] total hip joint arthroplasty[21] and treatment of back pain.[8] Other health profiles include the Nottingham Health Profile,[22] the McMaster Health Index Questionnaire[23] and a collection of related instruments developed by the Rand Corporation, Santa Monica, California, for its study of health insurance.[24]

Health profiles offer a number of advantages. For example, they allow determination of the effects of an intervention on different aspects of quality of life without the need for multiple instruments and thus save both the investigator and the patient time, and because they are designed for a wide variety of conditions they can be used to compare the effects of interventions in different diseases.

Health profiles also have limitations. They may not focus on the aspects of quality of life of specific interest. Inadequate focus will likely result in an unresponsive instrument that may miss small but clinically important changes in quality of life.[25,26] However, if the intervention is likely to affect aspects of quality of life included in a health profile the responsiveness may be adequate. For example, at least some of the SIP dimensions have detected differences in outcome between patients in intervention and control groups in randomized trials of cardiac rehabilitation[20] and between those who underwent either amputation or limb-sparing surgery for soft-tissue sarcoma.[27]

Utility measures: These are derived from economic and decision theories and reflect the preferences of patients for treatment process and outcome. Quality of life is measured as a single number along a continuum, death being 0.0 and full health 1.0. The use of these measures in clinical trials requires measurement of the patient's quality of life throughout the study.

There are two fundamental approaches to utility measurement. One is to ask patients a number of questions about their function and to classify the patients into one of a number of categories on the basis of their responses. Each category has a value assigned to it that has been established in previous ratings by another group (e.g., a random sample of the general population). This approach characterizes the widely used quality of well-being scale.[10,11,15]

The second approach is to ask patients to make a single rating of all aspects of their quality of life.[12] This can be accomplished in many ways, such as the standard gamble, in which subjects are asked to choose between their own health state and a gamble that they may die immediately or achieve full health for the remainder of their lives; the quality of life is determined by the choices made as the probabilities of immediate death or full health are varied. A simplified, more widely used technique is the time trade-off: subjects are asked how many years in their present health state they would be willing to trade off for a shorter life span in full health.[12]

A major advantage of utility measures is their amenability to cost–utility analysis, in which the cost of an intervention is related to the number of quality-adjusted life-years (QALYs) gained. But there are limitations. The measurements can vary depending on how they are obtained, and thus the validity of any single measurement is questioned;[28,29] however, the differences between scores obtained from standard gamble methods and those from time trade-off methods are seldom great. Utility measures do not allow the determination of what aspects of quality of life are responsible for any changes in utility. Subjects provide a rating that accounts for both treatment and side effects. Finally, utility measures at least potentially share the disadvantage of health profiles in that they may not be responsive to small yet clinically important changes.

Utility measures have been found to be responsive in at least two randomized clinical trials.[7,30] Using the quality of well-being scale Toevs, Kaplan and Atkins[7] showed that a program designed to improve compliance with an exercise regimen among patients with chronic airflow limitation could improve quality of life; the cost of the program was $24 256 for each QALY gained. In a double-blind randomized placebo-controlled trial of auranofin therapy for rheumatoid arthritis both the quality of well-being scale and a measure based on time trade-off were found to be highly responsive (indeed, more so than traditional measures such as the number of tender or swollen joints).[30]

Specific instruments

An alternative approach to quality-of-life measurement is to focus on specific aspects of health status.[17] The rationale for this approach lies in its potential for increased responsiveness, because only important aspects of quality of life are included. The instrument may even focus on problems specific to an individual patient.[31]

The instrument may be specific to the disease (e.g., chronic lung disease or rheumatoid arthritis), to a population of patients (e.g., the elderly), to a certain function (e.g., emotional or sexual function) or to a given condition or problem (e.g., pain) that may be caused by various underlying diseases.

In addition to the likelihood of being more responsive than generic instruments, specific measures focus closely on areas routinely explored by physicians. For example, in chronic lung disease a disease-specific measure of quality of life focuses on dyspnea, day-to-day activities, fatigue and areas of emotional dysfunction, such as frustration and impatience.[31]

Disease-specific measures have been developed for many conditions, such as cardiovascular disease,[32] chronic lung disease,[31,33] arthritis[34,35] and cancer,[36,37] and can be constructed to reflect a "single state" (How tired have you been? Very tired, somewhat tired or full of energy?) or a "transition" (How has your tiredness been? Better, the same or worse?).[38]

Disease-specific instruments have been proven to be useful in clinical trials.[6,31,32] The disadvantage is that they are (deliberately) not comprehensive and cannot be used for comparison between conditions or, at times, even between programs.

Use of multiple quality-of-life measures

Clinical investigators are not restricted to using a single instrument. Much remains to be learned about optimal ways of measuring quality of life, and investigators may wish to see how different instruments perform. Aside from this sort of inquiry, which focuses on the instruments rather than the intervention, an investigator may conclude that a single instrument will not yield all the relevant information. For example, utility and disease-specific measures generate quite different sorts of data, and an investigator may wish to use one of each. Determining the relative merits of different approaches will require further investigation. Direct comparisons, as conducted by Bombardier, Ware and Russell[30] in a trial of oral gold therapy for rheumatoid arthritis, are likely the best way to sort out the optimal methods for measuring quality of life in clinical trials.

Another, somewhat different way to use multiple instruments is to administer a battery of specific ones. For example, in a double-blind, randomized trial of three antihypertensive agents in the treatment of primary hypertension.[39] The investigators identified five dimensions of health they wished to measure: the sense of well-being and satisfaction with life, the physical state, the emotional state, the intellectual functioning, and the ability to perform in social roles and the degree of satisfaction gained from those roles. Even within these dimensions additional components were identified. Patients taking one of the three drugs scored better in general well-being, work performance and satisfaction with life. Thus, a clinician's choice of drug can affect not only the length but also the quality of the patient's life.

This approach, although comprehensive, has limitations. First, investigators must find a valid, responsive instrument for every attribute they wish to measure. Second, likely only some of the instruments chosen will show differences between the treatments. Unless one of the instruments has been designated as the primary measure of outcome before the trial starts, interpretation may be difficult because of different results from the different measures. The greater the number of instruments used the more likely one or more will favour one treatment or the other, even if the treatments are equally effective. Thus, the probability of finding an apparent difference between treatments even though they are equally effective (α error) increases with each new instrument used. Although this problem may be dealt with through statistical adjustment for the number of instruments used, such adjustment is seldom done.[40]

If only a few of the instruments favour an intervention the clinician may be unsure of how to interpret the results. For example, in a controlled trial in which patients with recent myocardial infarction were randomly assigned to receive standard care or to follow an exercise program or a counselling program Mayou and associates[41] rated work, leisure, sex, satisfaction with outcome, compliance with advice, quality of leisure and work, psychiatric symptoms, cardiac symptoms and general health. For almost all of these variables there was no difference between the three groups. However, patients in the exercise program were more satisfied with outcome than the others, families in the counselling group were less protective, and patients in the counselling group had a greater number of work hours and frequency of sexual intercourse at follow-up 18 months later. We agree with the conclusion of Mayou and associates[41] that the study did not support the effectiveness of rehabilitation in improving quality of life. However, some might argue that if even some of the ratings favour treatment the intervention is worth while. The use of multiple instruments opens the door to such potential controversy.

Review of quality-of-life measurement in randomized trials

Methods

To determine the extent and nature of quality-of-life measurement in randomized trials we examined such trials published in 1986 in the *Annals of Internal Medicine*, the *New England Journal of Medicine* and the *American Journal of Medicine*. Studies were included if the subjects were randomly allocated to one of two or more alternative treatments. Studies published in supplements were excluded. We evaluated each trial according to the level of importance (as we saw it) of quality of life as a measure of outcome by using the following criteria.

- Crucial: Physician cannot make rational treatment decision without information on the effect of intervention on quality of life.
- Important: Information on quality of life likely to aid physician in making optimal treatment decision.
- Secondary: Information on quality of life may be of interest but will not likely affect treatment decision.
- Irrelevant or not feasible.

The extent to which quality of life was actually measured was evaluated according to the following criteria.

- At least one instrument of demonstrated reproducibility, validity and responsiveness was used.
- An ad hoc or untested instrument was used to measure quality of life.
- No attempt was made to measure quality of life.

Two of us (G.H.G. and S.J.O.V.V.Z.) evaluated each article independently. Agreement was quantified by means of a measure of chance-corrected agreement. Both κ and weighted κ, which accounts for partial agreement, were calculated.[42] Quadratic weights were used in estimating the weighted κ. Any disagreement was resolved through discussion between the two raters.

Results

Twelve randomized trials were found in the *Annals of Internal Medicine*,[43-54] 53 in the *New England Journal of Medicine*[8,39,55-105] and 10 in the *American Journal of Medicine*.[30,106-114] For measuring agreement on the importance of quality-of-life measurement the κ was 0.65 and the weighted κ 0.78; for the quality of the measurement used the κ was 0.71 and the weighted κ 0.86. Quality of life was crucial or important in 55 of the trials (Table II); however, in only 10 of these was a validated, responsive measure used. Trials involving established measures used different approaches: generic instruments included the health profile (the SIP) in a trial of the effect of different durations of bed rest on low back pain[8] and various utility measures in a study of oral gold therapy for rheumatoid arthritis;[30] specific instruments included a single one used in a trial of antacids and cimetidine in the treatment of nonulcer dyspepsia[63] and a battery of such instruments in a trial of benzodiazepine withdrawal.[91]

In 11 trials that had a quality-of-life rating of crucial no attempt at measurement had been made.[46,47,49,51,74,77,80,82,87,93,97] In six of these trials the investigators did not measure the effects on quality of life that would justify treatment decisions but, rather, chose "substitute" endpoints. For example, spirometry results were used rather than the measurement of dyspnea in daily activities among patients with chronic airflow limitation,[87] findings on physical examination rather than pain and function in daily activities among children with juvenile rheumatoid arthritis,[74] blood pressure rather than postural symptoms among patients with orthostatic hypotension,[51] exercise capacity in the laboratory rather than chest pain during daily activities among patients with angina[46] and response rates and disease-free survival times rather than quality of life among cancer patients for whom chemotherapy did not alter survival.[80] In each case choosing treatment on the basis of study results requires acceptance of the substitute endpoint. Direct measurement of the relevant quality-of-life variables would have greatly strengthened the basis for treatment recommendations from these studies.

The trials in which quality of life was not measured despite a rating of crucial included one in which early discharge of very-low-birth-weight infants was shown to be safe and less costly than longer hospital stays.[93] However, no attempt was made to measure the extent to which early discharge disrupted the parents' lives and its impact on parental anxiety and well-being. In another trial Mitsuyasu and colleagues[49] demonstrated that treatment of donor bone marrow with anti-T-cell antibody reduced the risk of graft-versus-host disease but decreased engraftment and increased the risk of relapse; utility measures could have played a crucial role by providing a common denominator for different sorts of adverse outcomes and balancing the potential effects of increased death versus decreased suffering on the survivors.

The use of untested, ad hoc measures in 18 studies in which quality of life was crucial[50,52,56-58,60,65,71,84,90,94,96,99,102,104,105,108,109] and another 3 in which it was important[43,48,103] is unfortunate, because the authors clearly felt that quality-of-life measurement was important but for some reason did not select a tested measure.

The use of an ad hoc measure may generate uncertainty concerning the significance of the results. For example, Koopmans Summers and coworkers[96] showed improvements in cognitive function with oral tetrahydroaminoacridine therapy among patients with Alzheimer's disease. However, their reliance on qualitative descriptions of change in daily function rather than on a number

Table II — Review of characteristics of quality-of-life measures in 75 randomized clinical trials[8,30,39,43-114] by level of importance of quality of life as a measure of outcome

	No. of trials			
	Characteristic			
Level of importance	Valid and responsive	Untested measure	Quality of life not measured	Total
Crucial	10	18	11	39
Important	–	3	13	16
Secondary	–	–	8	8
Irrelevant	–	1	11	12
Total	10	22	43	75

of instruments designed to measure behaviour or functional status leaves doubt as to the clinical importance of their findings. One might also question the clinical importance of interferon therapy for condylomata acuminata.[94]

Finally, there have been trials in which quality-of-life measurement was judged to be important (rather than crucial) but was not attempted. These include studies of chemotherapy and radiotherapy[48,111] and of antibiotic therapy for the prevention or treatment of infectious complications of immunosuppression.[110,113,114] Fortunately, the value of quality-of-life measurement in these areas is becoming increasingly recognized, and a number of specific validated measures are now available.[36,37]

Conclusions

A number of instruments for measuring quality of life in clinical trials are now available. Each instrument and study approach has its strengths and weaknesses, and none is suitable for all situations. The relative merits of the different approaches must be further evaluated. Nevertheless, instruments that provide accurate, clinically important information are available for most health problems for which randomized trials are conducted, and there are various guidelines for selecting the appropriate instruments.[115]

Although quality-of-life measures have been used with increasing sophistication in some studies, they are generally underused. When investigators do measure quality of life they often use untested, ad hoc measures. Before beginning a trial researchers should ask themselves if the measurement of quality of life is important. If so, they should carefully consider the optimal approach and seek an instrument with established reproducibility, validity and responsiveness.

We thank Dr. Claire Bombardier for suggesting the topic of this article and Debbie Maddock and Christine Bolduc for preparing the manuscript.

This work was supported in part by the Ontario Ministry of Health.

References

1. Bergner M: Measurement of health status. *Med Care* 1985; 23: 696-704
2. Prigatano GP, Wright EC, Levin D: Quality of life and its predictors in patients with mild hypoxemia and chronic obstructive pulmonary disease. *Arch Intern Med* 1984; 144: 1613-1619
3. Guyatt GH, Townsend M, Berman LB et al: Quality of life in patients with chronic airflow limitation. *Br J Dis Chest* 1987; 81: 45-54
4. Nelson E, Conger B, Douglass R et al: Functional health status levels of primary care patients. *JAMA* 1983; 249: 3331-3338
5. Siegrist J: Impaired quality of life as a risk factor in cardiovascular disease. *J Chronic Dis* 1987; 40: 571-578
6. Guyatt GH, Townsend M, Pugsley SO et al: Bronchodilators in chronic airflow limitation: effects on airway function, exercise capacity and quality of life. *Am Rev Respir Dis* 1987; 135: 1069-1074
7. Toevs CD, Kaplan RM, Atkins CJ: The costs and effects of behavioral programs in chronic obstructive pulmonary disease. *Med Care* 1984; 22: 1088-1100
8. Deyo RA, Diehl AK, Rosenthal M: How many days of bed rest for acute low back pain? A randomized clinical trial. *N Engl J Med* 1986; 315: 1064-1070
9. Fletcher AE, Hunt BM, Bulpitt CJ: Evaluation of quality of life in clinical trials of cardiovascular disease. *J Chronic Dis* 1987; 40: 557-566
10. Kaplan RM, Bush JW: Health-related quality of life measurement for evaluation research and policy analysis. *Health Psychol* 1982; 1: 61-80
11. Torrance GW: Measurement of health state utilities for economic appraisal. *J Health Econ* 1986; 5: 1-30
12. Spitzer WO: State of science 1986: quality of life and functional status as target variables for research. *J Chronic Dis* 1987; 40: 465-471
13. Deyo RA: Measuring functional outcomes in therapeutic trials for chronic disease. *Controlled Clin Trials* 1984; 5: 223-240
14. Patrick DL, Deyo RA: Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989; 27: 217-232
15. Kaplan RM, Bush JW, Berry CC: Health status: types of validity and the index of well-being. *Health Serv Res* 1976; 11: 478-507
16. Kirshner B, Guyatt GH: A methodologic framework for assessing health indices. *J Chronic Dis* 1985; 38: 27-36
17. Guyatt GH, Bombardier C, Tugwell PX: Measuring disease-specific quality of life in clinical trials. *Can Med Assoc J* 1986; 134: 889-895
18. Guyatt GH, Walter S, Norman G: Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987; 40: 171-178
19. Bergner M, Bobbitt RA, Carter WB et al: The Sickness Impact Profile: development and final revision of a health status measure. *Med Care* 1981; 19: 787-805
20. Ott CR, Sivarajan ES, Newton KM et al: A controlled randomized study of early cardiac rehabilitation: the Sickness Impact Profile as an assessment tool. *Heart Lung* 1983; 12: 162-170
21. Liang MH, Larson MG, Cullen KE et al: Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. *Arthritis Rheum* 1985; 28: 542-547
22. Hunt SM, McKenna SP, McEwen J et al: A quantitative approach to perceived health status: a validation study. *J Epidemiol Community Health* 1980; 34: 281-286
23. Sackett DL, Chambers LW, MacPherson AS et al: The development and application of indices of health: general methods and a summary of results. *Am J Public Health* 1977; 67: 423-428
24. Ware JE, Brook RH, Davies-Avery A et al: *Conceptualization and Measurement of Health for Adults in the Health Insurance Study*, vol 1 of *Model of Health and Methodology*, Rand Corporation, Santa Monica, Calif, 1980
25. MacKenzie CR, Charlson ME, DiGioia D et al: Can the Sickness Impact Profile measure change? An example of scale assessment. *J Chronic Dis* 1986; 39: 429-438
26. Deyo RA, Centor RM: Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. Ibid: 897-906
27. Sugarbaker PH, Barofsky I, Rosenberg SA et al: Quality of life assessment of patients in extremity sarcoma clinical trials. *Surgery* 1982; 91: 17-23
28. Sutherland HJ, Dunn V, Boyd NF: Measurement of values for states of health with linear analog scales. *Med Decis Making* 1983; 3: 477-487
29. Llewellyn-Thomas H, Sutherland HJ, Tibshirani R et al: The measurement of patients' values in medicine. *Med Decis Making* 1982; 2: 449-462
30. Bombardier C, Ware J, Russell IJ et al: Auranofin therapy and quality of life in patients with rheumatoid arthritis:

results of a multicenter trial. *Am J Med* 1986; 81: 565-578
31. Guyatt GH, Berman LB, Townsend M et al: A measure of quality of life for clinical trials in chronic lung disease. *Thorax* 1987; 42: 773-778
32. Olsson G, Lubsen J, van Es G et al: Quality of life after myocardial infarction: effect of long term metoprolol on mortality and morbidity. *Br Med J* 1986; 292: 1491-1493
33. Mahler DA, Weinberg DH, Wells CK et al: The measurement of dyspnea. Contents, interobserver agreement, and physiologic correlates of two new clinical indexes. *Chest* 1984; 85: 751-758
34. Meenan RF, Anderson JJ, Kazis LE et al: Outcome assessment in clinical trials. *Arthritis Rheum* 1984; 27: 1344-1352
35. Fries JF, Spitz P, Kraines RG et al: Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980; 23: 137-145
36. Spitzer WO, Dobson AJ, Hall J et al: Measuring the quality of life of cancer patients. *J Chronic Dis* 1981; 34: 585-597
37. Schipper H, Clinch J, McMurray A et al: Measuring the quality of life of cancer patients: the Functional Living Index-Cancer: development and validation. *J Clin Oncol* 1984; 2: 472-483
38. MacKenzie CR, Charlson ME: Standards for the use of ordinal scales in clinical trials. *Br Med J* 1986; 292: 40-43
39. Croog SH, Levine S, Testa MA et al: The effects of antihypertensive therapy on the quality of life. *N Engl J Med* 1986; 314: 1657-1664
40. Pocock SJ, Hughes MD, Lee RJ: Statistical problems in the reporting of clinical trials. *N Engl J Med* 1987; 317: 426-432
41. Mayou R, MacMahon D, Sleight P et al: Early rehabilitation after myocardial infarction. *Lancet* 1981; 2: 1399-1401
42. Cohen J: Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* 1968; 70: 213-220
43. Hoofnagle JH, Davis GL, Pappas SC et al: A short course of prednisolone in chronic type B hepatitis. *Ann Intern Med* 1986; 104: 12-17
44. Lieberman DA, Keeffe EB: Treatment of severe reflux esophagitis with cimetidine and metoclopramide. Ibid: 21-26
45. Shapiro MF, Ware JE, Sherbourne CD: Effects of cost sharing on seeking care for serious and minor symptoms: results of a randomized controlled trial. Ibid: 246-251
46. Jordan RA, Seth L, Casebolt P et al: Rapidly developing tolerance to transdermal nitroglycerin in congestive heart failure. Ibid: 295-298
47. Taylor SG, Gelman RS, Falkson G et al: Combination chemotherapy compared to tamoxifen as initial therapy for stage IV breast cancer in elderly women. Ibid: 455-461
48. Bonadonna G, Valagussa P, Santoro A: Alternating non-cross-resistant combination chemotherapy or MOPP in stage IV Hodgkin's disease: a report of 8-year results. *Ann Intern Med* 1986; 105: 739-746
49. Mitsuyasu RT, Champlin RE, Gale RP et al: Treatment of donor bone marrow with monoclonal anti-T-cell antibody and complement for the prevention of graft-versus-host disease. *Ann Intern Med* 1986; 105: 20-26
50. Wharton JM, Coleman DL, Wofsy CB et al: Trimethoprim-sulfamethoxazole or pentamidine for *Pneumocystis carinii* pneumonia in the acquired immunodeficiency syndrome: a prospective randomized trial. Ibid: 37-44
51. Hoeldtke RD, Cavanaugh ST, Hughes JD et al: Treatment of orthostatic hypotension with dihydroergotamine and caffeine. Ibid: 168-173
52. Bell RC, Mullins RC, West LG et al: The effect of almitrine bismesylate on hypoxemia in chronic obstructive pulmonary disease. Ibid: 342-346
53. Shen EN, Keung E, Huycke E et al: Intravenous propafenone for termination of reentrant supraventricular tachycardia. Ibid: 655-661
54. Duke RJ, Bloch RF, Turpie AGG et al: Intravenous heparin for the prevention of stroke progression in acute partial stable stoke: a randomized controlled trial. Ibid: 825-828
55. McCormick JB, King IJ, Webb PA et al: Lassa fever: effective therapy with ribavirin. *N Engl J Med* 1986; 314: 20-26
56. Douglas RM, Moore BW, Miles HB et al: Prophylactic efficacy of intranasal α_2-interferon against rhinovirus infections in the family setting. Ibid: 65-70
57. Hayden FG, Albrecht JK, Kaiser DL et al: Prevention of natural colds by contact prophylaxis with intranasal α_2-interferon. Ibid: 71-75
58. Whitley RJ, Alford CA, Hirsch MS et al: NIAID Collaborative Antiviral Study Group. Vidarabine versus acyclovir therapy in herpes simplex encephalitis. Ibid: 144-149
59. Littenberg B, Gluck EH: A controlled trial of methylprednisolone in the emergency treatment of acute asthma. Ibid: 150-152
60. Shepp DH, Dandliker PS, Meyers JD: Treatment of varicella-zoster virus infection in severely immunocompromised patients: a randomized comparison of acyclovir and vidarabine. *N Engl J Med* 1986; 314: 208-212
61. Pappaioanou M, Fishbein DB, Dreesen DW et al: Antibody response to preexposure human diploid-cell rabies vaccine given concurrently with chloroquine. Ibid: 280-284
62. MacMahon SW, Wilcken DEL, MacDonald GJ: The effect of weight reduction on left ventricular mass: a randomized controlled trial in young, overweight hypertensive patients. Ibid: 334-339
63. Nyren O, Adami H, Bates S et al: Absence of therapeutic benefit from antacids or cimetidine in non-ulcer dyspepsia. Ibid: 339-343
64. Brain Resuscitation Clinical Trial I Study Group: Randomized clinical study of thiopental loading in comatose survivors of cardiac arrest. Ibid: 397-403
65. Dyck PJ, Daube J, O'Brien P et al: Plasma exchange in chronic inflammatory demyelinating polyradiculoneuropathy. Ibid: 461-465
66. Rajfer J, Handelsman DJ, Swerdloff RS et al: Hormonal therapy of cryptorchidism: a randomized, double-blind study comparing human chorionic gonadotropin and gonadotropin-releasing hormone. Ibid: 466-470
67. Austin HA, Klippel JH, Balow JE et al: Therapy of lupus nephritis: controlled trial of prednisone and cytotoxic drugs. Ibid: 614-619
68. Hearst N, Newman TB, Hulley SB: Delayed effects of the military draft on mortality: a randomized natural experiment. Ibid: 620-624
69. Feldt-Rasmussen B, Mathiesen ER, Hegedus L et al: Kidney function during 12 months of strict metabolic control in insulin-dependent diabetic patients with incipient nephropathy. Ibid: 665-670
70. Storb R, Deeg HJ, Whitehead J et al: Methotrexate and cyclosporine compared with cyclosporine alone for prophylaxis of acute graft versus host disease after marrow transplantation for leukemia. Ibid: 729-735
71. O'Neill W, Timmis GC, Bourdillon PD et al: A prospective randomized clinical trial of intracoronary streptokinase versus coronary angioplasty for acute myocardial infarction. Ibid: 812-818
72. Bowden RA, Sayers M, Flournoy N et al: Cytomegalovirus immune globulin and seronegative blood products to prevent primary cytomegalovirus infection after marrow transplantation. Ibid: 1006-1010
73. Canadian Multicentre Transplant Study Group: A randomized clinical trial of cyclosporine in cadaveric renal transplantation: analysis at three years. Ibid: 1219-1225
74. Brewer EJ, Giannini EH, Kuzmina N et al: Penicillamine and hydroxychloroquine in the treatment of severe juvenile rheumatoid arthritis. Ibid: 1269-1276
75. Salzman EW, Weinstein MJ, Weintraub RM et al: Treatment with desmopressin acetate to reduce blood loss after cardiac surgery: a double-blind randomized trial. Ibid: 1402-1406
76. Smith GR, Monson RA, Ray DC: Psychiatric consultation in somatization disorder: a randomized controlled study. Ibid: 1407-1413

77. A prospective trial of intravenous streptokinase in acute myocardial infarction (ISAM): mortality, morbidity, and infarct size at 21 days. Ibid: 1465–1471
78. Cohn JN, Archibald DG, Ziesche S et al: Effect of vasodilator therapy on mortality in chronic congestive heart failure: results of a Veterans Administration Cooperative Study. Ibid: 1547–1552
79. Gaston MH, Verter JI, Woods G et al: Prophylaxis with oral penicillin in children with sickle cell anemia: a randomized trial. Ibid: 1593–1599
80. Link MP, Goorin AM, Miser AW et al: The effect of adjuvant chemotherapy on relapse-free survival in patients with osteosarcoma of the extremity. Ibid: 1600–1606
81. Boyer KM, Gotoff SP: Prevention of early-onset neonatal group B streptococcal disease with selective intrapartum chemoprophylaxis. Ibid: 1665–1669
82. Yu VL, Goetz A, Wagener M et al: *Staphylococcus aureus* nasal carriage and infection in patients on hemodialysis: efficacy of antibiotic prophylaxis. *N Engl J Med* 1986; 315: 91–96
83. Newburger JW, Takahashi M, Burns JC et al: The treatment of Kawasaki syndrome with intravenous gamma globulin. Ibid: 341–347
84. Gibson RS, Boden WE, Theroux P et al: Diltiazem and reinfarction in patients with non-Q-wave myocardial infarction: results of a double-blind, randomized, multicenter trial. Ibid: 423–429
85. Pizzo PA, Hathorn JW, Hiemenz J et al: A randomized trial comparing ceftazidime alone with combination antibiotic therapy in cancer patients with fever and neutropenia. Ibid: 552–558
86. Livio M, Sci B, Mannucii PM et al: Conjugated estrogens for the management of bleeding associated with renal failure. Ibid: 731–735
87. Easton PA, Jadue C, Dhingra S et al: A comparison of the bronchodilating effects of a beta-2 adrenergic agent (albuterol) and an anticholinergic agent (ipratropium bromide), given by aerosol alone or in sequence. Ibid: 735–739
88. Merritt TA, Hallman M, Bloom BT et al: Prophylactic treatment of very premature infants with human surfactant. Ibid: 785–790
89. Vanderpas JB, Rivera-Vanderpas MT, Bourdoux P: Reversibility of severe hypothyroidism with supplementary iodine in patients with endemic cretinism. Ibid: 791–795
90. Packer M, Lee WH, Yushak M et al: Comparison of captopril and enalapril in patients with severe chronic heart failure. Ibid: 847–853
91. Busto U, Sellers EM, Naranjo CA et al: Withdrawal reaction after long-term therapeutic use of benzodiazepines. Ibid: 854–859
92. Turpie AGG, Levine MN, Hirsh J et al: A randomized controlled trial of a low-molecular-weight heparin (Enoxaparin) to prevent deep-vein thrombosis in patients undergoing elective hip surgery. Ibid: 925–929
93. Brooten D, Kumar S, Brown LP et al: A randomized clinical trial of early hospital discharge and home follow-up of very-low-birth-weight infants. Ibid: 934–939
94. Eron LJ, Judson F, Tucker S et al: Interferon therapy for condylomata acuminata. Ibid: 1059–1064
95. Hull RD, Raskob GE, Hirsh J et al: Continuous intravenous heparin compared with intermittent subcutaneous heparin in the initial treatment of proximal-vein thrombosis. Ibid: 1109–1114
96. Koopmans Summers W, Majovski LV, Marsh GM et al: Oral tetrahydroaminoacridine in long-term treatment of senile dementia, Alzheimer type. Ibid: 1241–1245
97. Siu AL, Sonnenberg FA, Manning WG et al: Inappropriate use of hospitals in a randomized trial of health insurance plans. Ibid: 1259–1266
98. Simberkoff MS, Cross AP, Al-Ibrahim M et al: Efficacy of pneumococcal vaccine in high-risk patients: results of a Veterans Administration Cooperative Study. Ibid: 1318–1327
99. Lung Cancer Study Group: Effects of postoperative mediastinal radiation on completely resected stage II and stage III epidermoid cancer of the lung. Ibid: 1377–1381
100. Laga M, Naamar W, Brunham RC et al: Single-dose therapy of gonococcal ophthalmia neonatorum with ceftriaxone. Ibid: 1382–1385
101. Ettinger B, Tang A, Citron JT et al: Randomized trial of allopurinol in the prevention of calcium oxalate calculi. Ibid: 1386–1389
102. Watt G, Theakston RDG, Hayes CG et al: Positive response to edrophonium in patients with neurotoxic envenoming by cobras (*Naja naja philippinensis*): a placebo-controlled study. Ibid: 1444–1448
103. Kaplan MM, Alling DW, Zimmerman HJ et al: A prospective trial of colchicine for primary biliary cirrhosis. Ibid: 1448–1454
104. Sobel JD: Recurrent vulvovaginal candidiasis: a prospective study of the efficacy of maintenance ketoconazole therapy. Ibid: 1455–1458
105. Hong WK, Endicott J, Itri LM et al: 13-cis-retinoic acid in the treatment of oral leukoplakia. Ibid: 1501–1505
106. Fanta CH, Rossing TH, McFadden ER: Treatment of acute asthma: Is combination therapy with sympathomimetics and methylxanthines indicated? *Am J Med* 1986; 80: 5–10
107. Wohns RN, Colpitts M, Clement T et al: Phenytoin and acute mountain sickness on Mount Everest. Ibid: 32–36
108. Dooley CP, Weiner JM, Larson AW: Endoscopy or radiography? The patient's choice: prospective comparative survey of patient acceptability of upper gastrointestinal endoscopy and radiography. Ibid: 203–207
109. White CJ, Phillips WA, Abrahams LA et al: Objective benefit of nifedipine in the treatment of Raynaud's phenomenon: double-blind controlled study. Ibid: 623–625
110. Winston DJ, Winston GH, Nakao SL et al: Norfloxacin versus vancomycin/polymyxin for prevention of infections in granulocytopenic patients. Ibid: 884–889
111. Koziner B, Myers J, Cirrincione C et al: Treatment of stages I and II Hodgkin's disease with three different therapeutic modalities. Ibid: 1067–1078
112. Peterson CM, Jovanovic L, Chanoch LH: Randomized trial of computer-assisted insulin delivery in patients with type 1 diabetes beginning pump therapy. *Am J Med* 1986; 81: 69–72
113. Karp JE, Dick JD, Angelopulos C: Empiric use of vancomycin during prolonged treatment-induced granulocytopenia: randomized, double-blind, placebo-controlled clinical trial in patients with acute leukemia. Ibid: 237–242
114. Cuttner J, Troy KM, Funaro L: Clotrimazole treatment for prevention of oral candidiasis in patients with acute leukemia undergoing chemotherapy: results of a double-blind study. Ibid: 771–774
115. McDowell I, Newell C: *Measuring Health: a Guide to Rating Scales and Questionnaires*, Oxford U Pr, New York, 1987