Jefferson.

Thomas Jefferson University
Jefferson Digital Commons

College of Population Health Faculty Papers

Jefferson College of Population Health

March 2002

Measuring health status and quality of life in disease management programs

Dea Belazi
Thomas Jefferson University

Let us know how access to this document benefits you

Follow this and additional works at: http://jdc.jefferson.edu/healthpolicyfaculty

Part of the Health Services Research Commons

Recommended Citation
Belazi, Dea, "Measuring health status and quality of life in disease management programs" (2002). *College of Population Health Faculty Papers.* Paper 19.
http://jdc.jefferson.edu/healthpolicyfaculty/19

This Article is brought to you for free and open access by the Jefferson Digital Commons. The Jefferson Digital Commons is a service of Thomas Jefferson University's Center for Teaching and Learning (CTL). The Commons is a showcase for Jefferson books and journals, peer-reviewed scholarly publications, unique historical collections from the University archives, and teaching tools. The Jefferson Digital Commons allows researchers and interested readers anywhere in the world to learn about and keep up to date with Jefferson scholarship. This article has been accepted for inclusion in College of Population Health Faculty Papers by an authorized administrator of the Jefferson Digital Commons. For more information, please contact: JeffersonDigitalCommons@jefferson.edu.

DISEASE MANAGEMENT
Volume 5, Number 1, 2002
© Mary Ann Liebert, Inc.

Review

Measuring Health Status and Quality of Life in Disease Management Programs

DEA BELAZI, Pharm.D., P.A.H.M.

ABSTRACT

Current medical practice is largely disease and problem focused, concentrating on the diagnosis and treatment of anatomic or physiologic problems. This approach assumes that physiologic parameters accurately reflect the patient's status of health. It also assumes that, when a patient's disease is ameliorated, his or her quality of life will improve. Many health services researchers find these parameters to be imperfect determinates for good outcomes of patients' perceived health status. However, there are many valid and reliable methods to measure patients' health status and quality of life. Recently available survey tools that measure quality of life are based on modern psychometric techniques. Disease management programs should take advantage of these valid and reliable tools to measure outcomes of their patients and assess the quality of their programs.

INTRODUCTION

DISEASE MANAGEMENT PROGRAMS which collect significant amounts of information and data from patients regarding disease, treatments, adverse events and many other characteristics, can be classified as a form of outcomes research.[1–3] Outcomes are categorized as clinical (i.e., HbA_{1C}, blood pressure), humanistic (i.e., patient satisfaction, quality of life), or economic (i.e., medical costs). However, many programs do not measure the appropriate endpoints that determine the effectiveness of treatment and improvement of the patient.[4–6] For the most part, the practice of medicine is largely disease and problem focused, concentrating exclusively on clinical measures such as lab tests and radiographic evidence, rather than patients' perceived severity of illness and state of health.

Many clinicians consider patient reported outcomes to be too subjective to measure and use accurately. In fact, there are many reliable and valid tools available to measure patients' perceived health.[7] The most recognized form of patient reported outcomes are health status and quality of life measurements. Health status measures are predominantly used in psychiatry, rheumatology, and geriatrics, where they are often referred to as scores, instruments, tools, or assessments. The medical interest in quality of life was stimulated by the success in prolonging life and the realization that this may be a mixed blessing: patients want to live, not merely to survive.

Office of Health Policy and Clinical Outcomes, Thomas Jefferson University, Philadelphia, Pennsylvania; Department of Global Epidemiology and Outcomes Research, Abbott Labs, Abbott Park, Illinois; and Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland.

Using health status surveys is one highly effective way healthcare providers can gather additional comprehensive data on patients' health. The intent of this paper is to describe how quality of life may be measured and to introduce a few examples of surveys and their use and limitations. The paper will also explore Dynamic Health Assessments, a new method of health status surveying based on modern psychometric theories, and describe how this can benefit disease management programs.

Health status and quality of life surveys

A recent important development in the healthcare field has been recognition of the importance of the patient's point of view in monitoring the quality of health outcomes.[8] The ultimate goal of medical care for patients is the achievement of a more productive life and the preservation of function and well-being.[9–13] However, physicians often fail to recognize problems such as functional disability in ambulatory care settings.[14–18] For example, a physician may prescribe an appropriate dose of insulin for a diabetic patient but fail to recognize that the patient is unable to administer his or her own medication or prepare his or her own meals. Several advances have been made over the past decade in the development of methods to assess patient perspectives of functional status and well-being, leading to the use of health status and quality of life surveys. Measuring quality of life longitudinally can help improve the quality of the patient's treatment and outcomes.

The overall concept of quality of life consists of a number of key domains: physical status and abilities, psychological status and well-being, and social functioning. Table 1 lists examples of quality of life and health status surveys and their domains. There are two classifications for these types of surveys: generic and disease-specific measures. Generic measures can be used in almost any population irrespective of the underlying condition. Because generic measures apply to a wide variety of populations, they allow for broad comparisons of the relative impact of various diseases. However, generic assessments may be less responsive to changes when compared to disease-specific surveys, yet disease-specific assessments can only be used in the population from which they have been tested and designed to elicit responses.

Table 1. Examples of Health Status and Quality of Life Surveys, and Their Domains

Surveys	*Domains*
Datmouth COOP Charts	Physical Emotional Daily activities Social activities Pain Overall health
EuroQOL	Mobility Self-care Usual activity Pain/discomfort Anxiety/depression
McMaster Health Index	Physical function Social function Emotional function
Nottingham Health Profile	Physical Social Emotional
SF-36, SF-12, SF-8 (Health Surveys)	General health Role physical Physical functioning Bodily pain Vitality Social functioning Role emotional Mental health
Sickness Impact Profile	Physical Psychosocial Independent (daily living)

Like most pharmaceuticals, surgical procedures, and diagnostic tests, quality of life and health status surveys undergo extensive testing to prove their validity and reliability. Validity is the extent to which an instrument provides the information it was designed to provide. There are many different types of validity, including content, criterion, and construct. Content validity refers to how adequate the questions (of the survey) represent what they are intended to measure. Criterion validity is derived from testing the instrument against a gold standard. Testing the validity and logic of the domains for predictive associations is known as construct validity. Construct validity is more frequently used when there is no gold standard against which to compare the instrument. Reliability is the extent to which the instrument produces the same results consis-

tently over time in the same situation. The validity and reliability of health status tools can be evaluated using evidence from the published literature. All of the surveys listed in Table 1 have been studied and shown to be valid and reliable.[19–26] This list is just a sample of generic health status and quality of life measures that are available, with more new disease specific tools on the horizon.

Deciding on the appropriate assessment tool for a disease management program can be difficult and sometimes confusing. Whether to use a generic or a disease-specific survey or both depends upon many factors such as time that patients have with providers, personnel to manage the collected surveys, and the burden of surveys for the patient. In a perfect world, where resources are unlimited, using both generic and disease-specific surveys would be best. Modern prevailing methodologies and psychometric theories that have recently found their way into measuring health could resolve this dilemma.

Modern psychometric theory (Item Response Theory)

Psychometric theory is defined as a measurement or procedure used to assign a numerical score to subjective judgments.[27] Widely accepted health status measures have been developed using traditional psychometric techniques. Modern psychometric techniques, such as Rasch models and Item Response Theory, were developed in quantitative psychology.[28,29] These more recent techniques have the potential to attain valid, precise, efficient, and reliable health status measurement for use at the individual patient level.[30] The purpose of these techniques is to tell us how likely patients or respondents at a particular level of health are apt to choose a response category over another.[30]

Item Response Theory and Rasch analysis build upon a statistical model of the dynamic between a patient's answer to a multiple-choice question and his or her own score on the concept being measured (e.g., a physical or emotional domain).[30,31] For health assessments, health status is an independent variable which can predict the probability of choosing each of the questionnaire response categories. Figure 1 demonstrates predictions based on the so-called partial credit model for the three response choices offered from the SF-36® physical functioning scale. This particular scale measures the function of climbing several flights of stairs. The horizontal axis in Figure 1 indicates the level (or score) of physical functioning. The three curves in the figure manifest the probability of selecting each response choice at each level of physical functioning. For instance, a patient with a score of 50 has a probability of 0.76 of choosing *not limited,* a little less

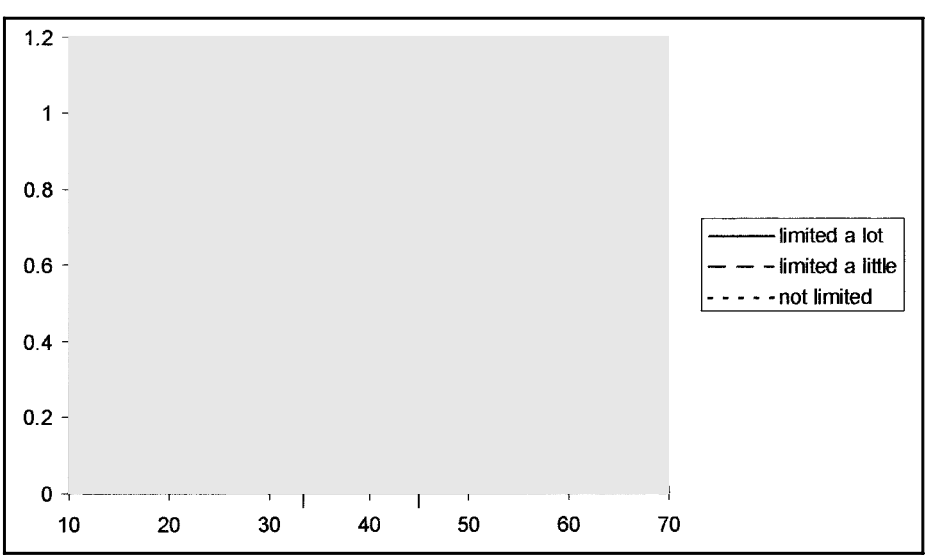

FIG. 1. Partial credit model trace lines for SF-36 item physical functioning (PF04). Climbing several flights of stairs.

than 0.24 probability of selecting *limited a little,* and less than 0.01 probability of choosing *limited a lot.* For those patients functioning at higher levels, above a score of 70, the probability of choosing not limited approaches one (or almost 100%). Those who have a very low physical functioning, a score below 20, have a probability approaching one (or almost 100%) of choosing *limited a lot.*

The three curves in Figure 1 describe item characteristics that hold true regardless of the health status of the population.[30,31] Two vertical lines, drawn between scores of 30 and 40 and 40 and 50, show the two points where the probability curves for adjacent response categories intersect. At a physical functioning score of approximately 34, the probability of choosing *limited a lot* is equal to that of choosing *limited a little* (hence their intersection). The probability of selecting *limited a little* is equal to that of selecting *not limited* at a score of about 45. These two score values are important because they show the item thresholds of the physical functioning scale in the partial credit model. As patients' physical function passes any of these thresholds, they tend to choose the response category above the threshold rather than the category below.[32–34] To estimate the probability of any pattern of item responses for any given level of health, the probabilities from these item characteristic curves can be combined.[30,35] This occurs when there are multiple questions or items for a specific domain. In practice, one does not know the physical functioning of any patients but can estimate the physical functioning for each pattern of responses. While many of the classical health status or quality of life surveys ask multiple questions for each domain, modern psychometrics can limit the number of questions without encroaching on the validity and precision.

Computer adaptive testing (Dynamic Health Assessments)

Most health status questionnaires measure the same core domains, but the questions presented are worded slightly differently. These variations prohibit comparisons of results across health measures. However, using modern psychometric methods, items from different health surveys that measure the same concept or dimension can be scored on a common level by equating the scales.[36] Using classical measurement methods and psychometrics, the best precision and validity can be achieved only by administering many questions, which will increase the burden of the respondent or patient.[30] The alternative is to use algorithms that extend from the calibration of item characteristics or questions that will allow the most precise responses. These algorithms, which are developed from modern psychometric methods, drive or power dynamic assessments. Of course, this requires the use of a computer, or some other electronic (or technologically capable) device that can execute these algorithms.

The process of the algorithm begins with an initial estimate of the respondent's score, which is usually the population average for that domain or category.[31] This estimate is used to select the most informative optimal item or question. The answer from the item is used to re-estimate the score. Next, the computer determines a specific confidence interval. Then the computer determines if the score has been estimated within a preset standard of precision based on the confidence interval. If the estimate is not precise enough, another question is given to the patient, and the cycle is repeated.[31]

Using a computer allows input from a patient to be stored directly into a database, and scoring of the survey is almost instantaneous. This decreases the resources usually needed in the administration and scoring of all paper forms of health status or quality of life surveys. The use of health status or quality of life surveys via a computer is known as computer adaptive testing or Dynamic Health Assessments. Computer adaptive testing is not a new approach but has been used in psychology and standardized testing for many years.

The science of modern psychometric theory applied to measuring health offers four distinct advantages over conventional surveys:

1. Dynamic Health Assessments have the brevity of a short form. This means that the tool will basically ask only the questions necessary to precisely rate or score the patient. Therefore, looking back at Figure 1, when a patient responds to an item or ques-

tion with either *not limited* or *limited a lot*, the patient does not require further questions about that particular domain and can move on to another domain of the health status survey. This will reduce response burden of the patient without affecting validity or reliability of the survey.

2. Dynamic Health Assessments maintain the validity and precision of a long form or survey. The structure of the survey is based on a long form survey, so that multiple questions for a particular domain are available if required. For instance, if the patient answers with *limited a little,* the survey will ask more questions to be able to precisely measure the patient's status.
3. Dynamic Health Assessments are able to measure the full range of health status or quality of life. The surveys do not have ceiling or floor effects, which are the inability to measure above a certain score (ceiling effects) or below a certain score (floor effects). This will allow for improved distinction between those who score very well or very poorly on a survey.
4. Dynamic Health Assessments can be compared across other forms such as those in Table 1. The development of the item pool of questions from other validated surveys allows for the Dynamic survey to be compared to other popular surveys.

Dynamic Health Assessments can be very beneficial to disease management programs. The ability of the assessment to be administered through a computer will save resources that were originally needed to administer, input (data into a database) and score the survey. Also, clinicians can use the scored survey to address other issues such as the possibility of depression (if that was not a known issue). For the patient, the response burden is less than that of the other lengthy paper form of surveys.[37] By keeping the time needed to administer the survey to a minimum, more time is allowed for the patient and clinician to interact. Finally, Dynamic Health Assessments in a disease management program will allow for comparisons of outcomes to other programs that use conventional health status and quality of life assessments.

Dynamic Health Assessments are currently available in generic version (that can be used in any population) and a few disease-specific versions that incorporate the generic items. One of the first disease-specific assessments created is the Dynamic Headache Impact Test™ (HIT).[38] There are more assessments currently in development for rheumatoid arthritis, osteoarthritis, congestive heart failure, asthma, and rhinitis.[39] More information about Dynamic Health Assessments can be found on the web at www.qmetric.com and www.amihealthy.com.

Implementing health status and quality of life monitoring in disease management

Disease management program developers and administrators first need to consider how quality of life measurement should be incorporated in an overall initiative. For example, health status and quality of life instruments can be used to screen and confirm diagnoses and diseases. They can also be used as an outcome measure for evaluating a treatment or the entire disease management program. Given the clinical and scientific basis for quality of life assessment, it is important to have in-house and/or consulting expertise when selecting and implementing quality of life tools. Once the objective or purpose of using these tools has been established, one should next consider the use of a generic or disease-specific assessment, or both. Earlier in this article, the benefits and limitations associated with using either generic or disease-specific assessments were discussed.

After a decision has been made to use either a generic or disease-specific assessment or both, an actual instrument needs to be identified. According to the MAPI Research Institute, there are more than 800 generic and disease specific tools from which to choose.[40] A list describing most of the assessments available can be found on the MAPI Research Institute's website (at http://195.101.204.50:8081/). Once an instrument has been selected, one should review the published literature to confirm its validity and reliability and examine its use in research and clinical settings. Just about every health status or quality of life assessment requires either a license or some type of permission for use. A disease management program

should obtain a license or written permission to avoid any potential legal consequences.

When choosing a survey instrument, there are other various factors and issues to consider. One factor is whether personnel are available at the site to perform the tasks needed to administer, collect, and analyze the surveys. Some assessments, such as the SF-36, can be scanned into a computer or even administered on a computer, which can save on resources of the staff.[41] For traditional paper surveys, personnel resources must be allocated for entering the data into a computer database. Also, the staff would have to be trained in statistical software use to score the surveys. There are survey administration and scoring services available where data could be gathered and/or tabulated and interpreted, such as by faxing a completed survey to a scoring vendor.[42,43] Clinicians who plan on using survey results must consider the time required to collect and score the survey on an ongoing basis. If the clinician would like to review the results at the time of the visit to modify treatment plans or discuss new potential concerns, then a system where collection and scoring of the survey is almost instantaneous would be best. A survey that is administered on a computer, such as Dynamic Health Assessments, could be ideal when results are needed during the patient visit. In any case, the use of internal resources versus outside vendors should be considered with respect to the logistics and resources inherent in the implementation scenario.

It is becoming vital for disease management programs to use and measure patients' health status and quality of life. Practicing physicians receive little feedback about the kinds of outcomes they achieve and how these outcomes compare to those of other physicians. The benefits of measuring health status and quality of life for disease management programs lie first in broadening the scope of outcome measures and second in providing a formal means for the patient's judgment to influence treatment. Quality of life measurement is also valuable in comparing treatments that are equivalent in terms of clinical effectiveness. Using these tools, clinicians can screen patients for risk of other problems or diseases and monitor patients' health status over time. Patients are the only reasonable source of information regarding the effects of illness on their quality of life or of subjective evaluations or ratings of their health. Only the patient can tell how severe a symptom feels or how much an illness causes (in human terms) suffering, pain, worry, or concern about health. Amid concerns about the difficulty of measuring these subjective patient data, scientific method and clinical expertise have been used to create a broad array of valid and reliable quality of life assessments, as described in this paper. Whether disease management programs use conventional surveys or Dynamic Health Assessments, measuring health status and quality of life will ultimately benefit the patient and demonstrate the quality of the program.

REFERENCES

1. Mosley C. Coordination of care in disease management: opportunities and financial issues: Semin Dial 2000;13:346–350.
2. Nesse RE, Hagedorn SD, Scheitel SM, et al. Disease management strategies: managing care giving in managed care. Qual Manag Health Care 2000;9:42–48.
3. Abarca J, Armstrong EP. Improving the use of data resources in disease management programs. Dis Manage Health Outcomes 2001;9:459–471.
4. Bodenheimer T. Disease management: promises and pitfalls. N Engl J Med 1999;340:1202–1205.
5. Armstrong EP, Langley PC. Disease management programs. Am J Health Syst Pharm 1996;53:53–58.
6. Holdford DA. Barriers to disease management. Am J Health Syst Pharm 1996;53:2093–2096.
7. Kozma CM. Health-related quality-of-life measures in disease management programs. Manag Care Interface 1998;11:71–72.
8. Geigle R, Jones SB. Outcomes measurement: a report for the front. Inquiry 1990;27:7–13.
9. McDermott W. Absence of indicators of the influence of its physicians on a society's health: impact of physician care on society. Am J Med 1981;70:833–843.
10. Cluff LE. Chronic disease, function and the quality of care. J Chronic Dis 1981;34:299–304.
11. Ellwood PM. Outcomes management: a technology of patient experience (Shattuck lecture). N Engl J Med 1988;318:1549–1556.
12. Schroeder SA. Outcomes assessment 70 years later: are we ready? N Engl J Med 1987;316:160–162.
13. Tarlov AR. The increasing supply of physicians, the changing structure of the health services system, and the future practice of medicine (Shattuck lecture). N Engl J Med 1983;398:1235–1244.
14. Deyo RA, Carter WB. Strategies for improving and

expanding the application of health status measures in clinical settings: a researcher-developer viewpoint. Med Care 1992;30:MS176–MS186.
15. Calkins DR, Rubenstein LV, Cleary PD, et al. Failure of physicians to recognize functional disability in ambulatory patients. Ann Intern Med 1991;114:451–454.
16. Nelson EC, Langraf JM, Hayes RD, et al. The functional status of patients: how can it be measured in physicians' offices? Med Care 1990;28:1111–1126.
17. German PS, Shapiro S, Skinner EA, et al. Detection and management of mental health problems of older patients by primary care providers. JAMA 1987;257: 489–493.
18. Wells KB, Hays RD, Burnham A., et al. Detection of depressive disorder for patients receiving prepaid or fee-for-service care: results from the Medical Outcomes Study. JAMA 1989;262:3298–3302.
19. Ware JE, Brook RH, Davies-Avery A, et al. Conceptualization and measurement of health for adults in the Health Insurance Study: Vol 1. Model of health and methodology. Santa Monica, CA: RAND (R-1987/1-HEW), 1980.
20. Bergner M, Bobbitt RA, Carter WB, et al. The Sickness Impact Profile conceptual formulation and methodology for the development of health status measure. Int J Health Serv 1976;6:393–415.
21. van Agt H, Essink-Bot M, Krabbe PFM, Bonsel GJ. Test-retest reliability of health state valuations collected with the EuroQol questionnaire. Soc Sci Med 1994;39;1537–1544.
22. Brazier J, Jones N, Kind P. Testing the validity of the EuroQol and comparing it to the SF-36 health survey questionnaire. Qual Lif Res 1993;2:169–180.
23. Chambers LW. The McMaster Health Index Questionnaire (MHIQ): methodological documentation and report of the second generation of investigations. Department of Clinical Epidemiology and Biostatistics. Hamilton, Canada: McMaster University, 1982.
24. Fortin F, Kerouac S. Validation of questionnaires on physical function. Nurs Res 1977;26:128–135.
25. Hunt SM, Mckenna SP, Williams J. Reliability of a population survey tool for measuring perceived health problems: a study of patients with osteoarthritis. J Epidemiol Commun Health 1981;35:297–300.
26. Ware JE, Snow KK, Kosinski M, Gandek B. SF-36 health survey manual and interpretation guide. Boston: New England Medical Center, The Health Institute, 1993.
27. McDowell I, Newell C. Measuring health: a guide to rating scales and questionnaires, 2nd ed. New York: Oxford University Press, 1996.
28. Lord FM, Norvick MR. Statistical theories of mental test scores. Reading, MA: Addison-Wesley, 1986.
29. Rasch G. Probabilistic models for some intelligence and attainment tests. Chicago: University of Chicago Press, 1980.
30. Bjorner JB, Ware JE. Using modern psychometric methods to measure health outcomes. Med Outcomes Trust Monitor 1998;3:12–19.
31. Ware JE, Bjorner J, Kosinski M. Dynamic health assessments: the search for more practical and more precise outcomes measures. Quality Life News Lett 1999;21:11–13.
32. Masters GN. A Rasch model for partial credit scoring. Psychometrika 1982;47:149–173.
33. Davies AR, ed. The SF-36 and SF-12 health surveys: how to use them. Woodbridge, NJ: HealthStat Productions, 1997.
34. Wright B, Masters GN. Rating scale analysis—Rasch measurement. Chicago: MESA Press, 1982.
35. Wainer H, Mislevy RJ. Item response theory, item calibration, and proficiency estimation. In: Wainer H, Dorans NJ, Flaugher R, et al., eds. Computerized adaptive testing: a primer. Hillsdale, NJ: Lawrence Erlbaum Associates, 1990:61–100.
36. Fisher WP, Euthbanks RL, Marier RL. Equating the MOS SF-36 and the LSU HIS physical functioning scales. J Outcome Measur 1997;1:329–362.
37. Jenkinson C, Fitzpatrick R, Garratt A, et al. Can item response theory reduce patient burden when measuring health status in neurological disorders? Results from Rasch analysis of the SF-36 physical functioning scale (PF-10). Med Care 2000;71:220–224.
38. QualityMetric Inc. New Assessment Tool Available On-line for Headache Patients and Those Who Treat Them. Available at: www.qmetric.com/pressroom/release_015.shtml. Accessed October 22, 2001.
39. QualityMetric Inc. Disease Impact Surveys. Available at: www.qmetric.com/innohome/indi.shtml. Accessed October 22, 2001.
40. Mapi Research Institute. Quality of Life Instrument Database. Available at: 195.101.204.50:8081/. Accessed December 10, 2001.
41. Assist Technologies. Available at: www.outcomes-analyzer.com. Accessed on December 10, 2001.
42. QualityMetric Inc. SF-36 Health Survey. Available at: www.qmetric.com/innohome/insf36.shtml. Accessed on December 10, 2001.
43. IMF Consultants. Available at: www.imf.com. Accessed on December 10, 2001.

Address reprint requests to:
Dea Belazi, Pharm.D.
Office of Health Policy & Clinical Outcomes
Thomas Jefferson University
1015 Walnut Street
Suite 115
Philadelphia, PA 19107

E-mail: Dea.Belazi@mail.tju.edu