
Volume 21, issue 2, June 2011
ISSN 0958-2592
ELSEVIER
The Foot
The International Journal of Clinical Foot Science
Official Journal of
EDUCATION
ADVOCACY
RESEARCH
The American College of Foot & Ankle Orthopedics & Medicine
Now indexed by MEDLINE
Original articles
Review articles
Case reports

The Foot 21 (2011) 84–87

Contents lists available at ScienceDirect

The Foot

journal homepage: www.elsevier.com/locate/foot

Responsiveness of the foot and ankle ability measure (FAAM) in individuals with diabetes

Benjamin R. Kivlan [a,*], RobRoy L. Martin [a,b,e], Dane K. Wukich [c,d,1]

[a] John G. Rangos Sr., School of Health Sciences, Duquesne University, Pittsburgh, PA 15282, United States
[b] Department of Physical Therapy, Duquesne University, Pittsburgh, PA 15282, United States
[c] University of Pittsburgh School of Medicine, Pittsburgh, PA 15203, United States
[d] UPMC Comprehensive Foot and Ankle Center, Roesch-Taylor Bldg. Ste 7300, 2100 Jane St., Pittsburgh, PA 15203, United States
[e] UPMC Sports Medicine Center, Pittsburgh, PA 15203, United States

ARTICLE INFO

Article history:
Received 4 March 2011
Received in revised form 9 April 2011
Accepted 11 April 2011

Keywords:
Foot and ankle ability measure
Responsiveness
Diabetes

ABSTRACT

Background: The impact of diabetes on physical function pose a challenge in assessing clinical outcomes.
Objective: The purpose of this study was to provide evidence of responsiveness for the foot and ankle ability measures (FAAM) in individuals with diabetes mellitus.
Methods: The two most recent FAAM scores of 155 diabetic patients treated for foot/ankle pathology were analyzed. Based on physical component summary (PCS) scores of the SF-36, subjects were categorized as improved (>7-point positive change), worsened (>7-point negative change), or unchanged (<7-point change). Analyses of the worsened and improved groups were compared to the unchanged group using two-way repeated measures ANOVAs and ROC curve analyses.
Results: The ANOVAs demonstrated a significant difference between groups ($P = 0.001$). ROC curves analysis for detecting an improvement or decline in status were 0.73 (95% CI 0.62–0.84) and 0.70 (95% CI 0.59–0.81), respectively. An increase in FAAM score of 9 points represented the minimal clinically important difference (MCID) with 0.64 sensitivity and 0.78 specificity. A decrease in FAAM score of 2 points represented a MCID with 0.65 sensitivity and 0.61 specificity.
Conclusions: The FAAM demonstrated responsiveness to change in individuals with orthopedic foot and ankle dysfunction complicated by diabetes and can be used to measure patient outcomes over a 6-month period.

1. Introduction

Diabetes mellitus (DM) affects 23.6 million Americans with 1.6 million new cases diagnosed each year [1]. Based on the incidence of DM in the general population, we expect that nearly one in twelve patients will have diabetes [1]. The impact of DM can affect multiple dimensions of physical, social, and psychological function. Specifically, DM can cause severe and debilitating conditions of the foot and ankle including ulcers, infections, and amputations [2–4]. The prevalence of DM and its impact on physical function can pose a challenge in assessing meaningful clinical outcomes. This is particularly true in patients with foot and ankle involvement complicated by diabetic neuropathy. This study aimed to assess the responsiveness of the foot and ankle ability measure (FAAM) outcome instrument for use among individuals with DM.

Several disease-specific instruments exist to assess patients with DM [5–19]. However, disease-specific instruments may not comprehensively address the functional impact of foot and ankle impairment. Thus an instrument that can be used to assess foot and ankle function in patients with DM is needed. The FAAM is an instrument developed to measure region-specific function of the foot and ankle. It is comprised of two individually scored subscales; a 21-item activities of daily living and 8-item sports subscale [20]. It is easy for patients to complete and uncomplicated for clinicians to calculate. The FAAM was developed with use of the item response theory and has undergone advanced psychometric testing that provides evidence of reliability, validity, and responsiveness in a general orthopedic population [20]. Recently, the ADL sub-scale of the FAAM has demonstrated evidence for validity among patients with foot and ankle dysfunction and co-existing DM [21].

Additional psychometric testing is needed to determine if the FAAM is responsive to changes in physical function of the foot and ankle in a diabetic population. The purpose of this study was to assess responsiveness of the FAAM in patients with DM. Responsiveness was assessed at the group and individual level. It was

* Corresponding author.
E-mail address: bkivlan@zoominternet.net (B.R. Kivlan).
[1] Tel.: +1 412 586 1546; fax: +1 412 586 1544.

0958-2592/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.foot.2011.04.004

Table 1
Subject characteristics.

	Number of patients	Percentage of patients
Diagnosis[a]		
Polyarthropathy	62	40%
Arthropathy	12	8%
Ulcer	16	10%
Ankle fracture	12	8%
Metatarsal fracture	11	7%
Tibia-Fibula fracture	1	<1%
Osteomyelitis	8	5%
Cellulitis	7	4%
Pes planus	7	4%
Deformity of the toe	5	3%
Plantar fasciitis	2	1%
Osteoporosis	6	4%
Osteochondrosis	2	1%
Osteoarthritis	2	1%
Amputation		
Foot	4	3%
Leg	1	<1%
Ankle instability	3	2%
Ankle sprain	2	1%
Insulin dependent		
Yes	87	56%
No	68	44%
Diabetic complications		
Yes	122	79%
No	33	21%
Surgical status		
Non-surgical	29	19%
Surgical	126	81%

[a] Subjects may have more than one diagnosis.

hypothesized that a change in FAAM scores in the groups that either improved or worsened would be greater than the group that did not change. At the individual level, it was hypothesized that ROC curves would offer statistical evidence that the FAAM was responsive to change in physical function. From this analysis the respective minimal clinically important difference (MCID) values that represent meaningful improvement and meaningful decline in functional status were determined.

2. Materials and methods

2.1. Subjects

Subjects consisted of patients being seen by an orthopedic surgeon for a routine clinic visit for pathology related to their foot and ankle. Inclusion criteria included having the diagnosis of DM with FAAM and Short Form-36 (SF-36) scores available. Exclusion criteria included those who were unable to read and understand English. One hundred and fifty-five subjects were included in the study. These subjects had a mean age of 59 years (range 18–80, SD = 10) with 63% females and 47% males. Subjects averaged a duration of 17 years (SD = 14) of DM. Description of the diagnoses, the presence of neuropathy, insulin dependence, and surgical status of the subjects is reported in Table 1.

2.2. Data collection

Subjects prospectively completed the FAAM activities of daily living subscale and SF-36 during each visit. Scores from the two most recent office visits (T1 and T2) were chosen for analysis. Changes in the physical component score (PCS) of the SF-36 score were calculated. Studies have shown that a 7-point change in PCS represents a significant change in status and was used as the criterion for change [22]. Subjects who had at least a 7-point increase or decrease in PCS were, respectively, categorized as improved or worsened. Those with less than a 7-point PCS change were categorized as unchanged.

Table 2
Comparison of FAAM scores based on change in status.

Change in status	Time 1	Time 2
Improved	39.7 (range 1–75 SD 21.8)	56.7 (range 23–84 SD 18.2)
Worsened	44.8 (range 10–83 SD 19.3)	38.4 (range 0–74 SD 19.5)
Unchanged	45.1 (range 1–84 SD 19.5)	48.5 (range 0–84 SD 28.5)

2.3. Data analysis

Evaluation of responsiveness requires comparison of a group that has changed to a group that has not changed. Thus a separate analysis of responsiveness should be determined for the group that improved as well as the group that worsened. This yields an assessment of responsiveness specific to the direction of change in the patient's score. Assessment of responsiveness was done at the group and individual level. Group level assessment of responsiveness was done with two separate two-way repeated measure ANOVAs. One analysis was done to compare change in FAAM scores in the group that demonstrated improvement compared to the group which did not change. A separate analysis was done to compare the group that demonstrated a decline of the PCS score beyond 7 points compared to those who did not change. The a priori alpha level for this analysis was set at 0.05. Statistical analysis was performed with SPSS 17.0 (Chicago, IL) statistical software package.

An assessment of responsiveness was done at the individual level using ROC curves. Similar to the group level analysis, two separate ROC curves and respective sensitivity and specificity values were calculated to determine the ability of the FAAM to be responsive to an improvement as well as a decline in physical status. A 95% confidence interval (CI) for the area under the ROC curve that does not contain 0.5 is indicative of responsiveness to change in physical function. Each ROC curve analysis determined a MCID value with the greatest sensitivity and specificity to identify meaningful change. The MCID values provide the criterion specific to the direction of change to determine if interventions have resulted in respective improvement or decline of function [23].

3. Results

Using a 7-point change on the PCS as the criterion for change, 34 (22%) subjects improved, 38 (25%) worsened, and 84 (53%) were unchanged. Mean FAAM scores for the groups that improved, worsened, and were unchanged at T1 and T2 are present in Table 2. The average change in the FAAM score between T1 and T2 for the group that improved was 17.0 (range 15–60 SD 19.0). The average change in FAAM scores between T1 and T2 for the group that remained unchanged was 3.3 (range: 35–56 SD 12.7). The average change in FAAM score between T1 and T2 for the group that worsened was −6.4 (range = 37–38 SD 15.3). The group by time interaction was significant when comparing the improved and unchanged groups ($F(1, 115) = 12.73$ $P = 0.001$) and comparing the worsened and unchanged groups($F(1, 119) = 21.63$ $P < 0.0005$). The areas under the ROC curve for detecting an improvement or decline in status were 0.73 (95% CI 0.62–0.84) and 0.70 (95% CI 0.59–0.81), respectively. The 95% CI for the areas under the ROC curves did not contain 0.5, indicating the FAAM was responsive to a change in status. An increase in FAAM score of 9 points represented a MCID with 0.64 and 0.78 sensitivity and specificity values, respectively. A decrease in FAAM score of 2 points represented a MCID with 0.65 and 0.61 sensitivity and specificity values, respectively.

4. Discussion

The results of this study offer evidence for the responsiveness of the FAAM in detecting change in function among patients with orthopedic foot and ankle pathology complicated by DM over a 6-month time frame. The responsiveness of the instrument was demonstrated on both the group and individual levels. At the group level, the instrument demonstrated a significant group by time interaction in 2-way repeated measures ANOVA that was consistent with our original hypothesis. Evidence of responsiveness at the individual level was demonstrated by the results of the ROC. As hypothesized, the area under the curve did not contain 0.5, indicating that the FAAM had the ability to detect a change at an individual level. These results build on existing evidence of the psychometric properties of the FAAM and offer additional evidence for use in detecting changes of physical function among orthopedic patients with complicating DM.

Given the prevalence of foot and ankle dysfunction among patients with DM, it is important to establish tools that have evidence of responsiveness to changes in physical function. This allows confidence that a change in status is the result of an actual change rather than an error of the outcome measure [24]. Evidence for psychometric characteristics may only be established in the context and population in which it has been evaluated [21]. Patients with DM being treated for foot and ankle conditions may experience different complaints than non-diabetic patients. Notably, patients with DM commonly have complicating neuropathy that may limit their complaints of pain. Other region-specific measures of the foot and ankle focus heavily on components of pain [25–29]. For instance, nearly 40% of Foot Function Index [26] score and over 21% of the Foot and Ankle Outcome score [29] is determined by pain. Patients with DM may not experience pain, but still may experience significant impairment to ADL function.

The FAAM specifically assesses the functional abilities of the patient, matching the components that may be of greatest interest to a diabetic population. By establishing evidence for responsiveness in the context of an orthopedic population of patients with DM, the instrument increases its usability for a greater range of patients. It also demonstrates evidence of responsiveness over a clinically relevant period of time. A 6-month interval represents a reasonable time-frame in which patients treated for foot and ankle impairments may be re-evaluated by their orthopedic surgeon. The FAAM is the only region-specific instrument that has established evidence of psychometric properties with content specific to the complaints of patients with DM. The results of this study support utilization of the FAAM among patients with primarily orthopedic foot and ankle dysfunction complicated by DM. However, further study is required to support its use under various conditions as the process of evaluating the usefulness of an instrument is a continual and ongoing process. Current studies are underway to evaluate the test re-test reliability and responsiveness over differing lengths of follow-up.

There are multiple statistical expressions that may provide evidence of responsiveness of an instrument [30,31]. Of particular importance is to determine the smallest change in score that can be considered a meaningful change, known as the MCID [32]. Most studies that report MCID values examine the response to an intervention with the expectation that the patient will demonstrate improvement. Thus, only the MCID value representing a positive change in status may be explored [33]. It may be equally useful to establish a MCID value to note when a decline in status has occurred. This is particularly true in the study of patients with progressive diseases, such as DM, where a decline in function may be anticipated. Therefore, an instrument utilized to assess outcomes would ideally detect meaningful improvement as well as a decline in function [33]. In our study, the patients that improved and the patients that worsened were individually compared to an unchanged group of patients to determine MCID values specific to the direction of change. This allowed interpretation of the changes in scores for patients that have demonstrated improvement as well as those who worsened according to their PCS scores. The bidirectional analysis of MCID values is recommended but seldom performed [33]. Brunner et al. [34] studied the bidirectional responsiveness of disease specific activity measures on patients with systemic lupus erythematosus. Their results indicated that a relatively small change in activity measures resulted in a meaningful change in status of both the patients who improved and the patients who worsened. However, the magnitude of the MCID values for each direction is not always equal [33]. Hays and Woolley [32] reference a general health perceptions scale that demonstrated an MCID value of 13 points to determine a meaningful change of improvement, while a 34-point regression was necessary to signify deterioration. This illustrates the disparity that may exist in MCID values of improvement versus worsening and further emphasizes the need for bidirectional analysis [32,33]. The results of our analysis indicate that a 9-point improvement in FAAM's ADL subscale is necessary to demonstrate a clinically meaningful improvement.

This is very similar to the findings by Martin et al. [20] indicating an 8-point MCID value among patients in a general orthopedic population. A 2-point regression represented a significant decline in self-reported function. Thus a patient with DM and foot/ankle complaints who has an initial FAAM score of 60 would need to improve to a score of 69 or greater at their 6-month follow-up appointment to be considered to have experienced a significant improvement of physical function. Alternately, the patient's score would need to decrease to 58 or lower to be considered to have a significant decline of function. The individual analysis of responsiveness allows clinical interpretation of individual FAAM scores among the target population that is specific to the direction of the change.

Ceiling and floor effects may limit usefulness of an instrument and negatively influence responsiveness. Ceiling effects occur when patients score very high on the instrument and have no room to demonstrate a change in status. Floor effects occur as the scores from the patients are very low and the items of the instrument are too advanced to capture the level of ability of the patient [35]. According to Martin et al. [21], the ADL subscale did not exhibit evidence of floor or ceiling effects. However, a significant floor effect was noted for the sports subscale scores. Thus the sports subscale may be too advanced to demonstrate change in a group of patients with DM and was excluded from evaluation in this study. Similar to the previous work [21], ceiling and floor effect were not evident in the ADL subscale study as the highest FAAM score was an 85 and less than 5% scored less than 10.

While this study demonstrates the responsiveness of the FAAM, limitations of the study must be recognized. First, the FAAM is a region-specific instrument designed to capture disability of the foot and ankle as reported by the patient. The FAAM does not contain items that are specific to DM nor does it contain observable measures of impairment. Therefore the role of the FAAM in a diabetic population may best be served as a compliment to disease-specific instruments and/or objective physical exam findings to offer the most complete representation of a clinical outcome. The methods of analysis may also serve as a limitation to the study. There are several additional ways to analyze responsiveness including mean differences, effect size, reliable change index, and responsiveness index [36]. Different methods may yield different interpretations [32]. We chose repeated ANOVAs and establishment of MCID values through ROC analysis to assess responsiveness. Utilizing two separate methods of analysis is recommended to allow corroboration of the results and strengthen the credibility of the findings [33,37]. An additional concern is related to the criterion used to establish groupings of patients who improved, worsened, or remained unchanged. Many studies utilize global change scores that may lack validity and reliability to determine perceived improvement of the patient [37]. We attempted to minimize this concern by utilizing a well-established scale (PCS of the SF-36) that has an established strong relationship to the FAAM [20] and has an ability to demonstrate change in patients with DM [22]. A 7-point change in status of the PCS represents two standard errors of measurement and has been previously established as the value to determine a significant change in status among patients with chronic diseases including DM [22,38]. Based on this information we feel the criteria for change used in this study was appropriate. This study builds from previous work in establishing evidence of psychometric properties of the FAAM among patients being treated for orthopedic foot and ankle problems with complicating DM.

5. Conclusions

The results provide evidence that the FAAM is responsive to change in physical function of the foot and ankle in individuals with diabetes. It is the first region-specific instrument to establish evidence of psychometric properties in a population of patients with DM. The FAAM is a unique tool that can be used to determine if an individual treated primarily for orthopedic foot and ankle dysfunction with complicating DM has improved, remained unchanged, or worsened as a result of treatment over a 6-month period.

Conflict of interest statement

The authors disclose no financial or personal relationships that influenced the work done for this study.

Acknowledgements

There were no other contributions made to this research study other than the authors. No funding was provided for the execution of this study.

References

[1] American Diabetes Association, National Diabetes Fact Sheet, 2007, http://www.diabetes.org/diabetes-basics/diabetes-statistics/.
[2] Margolis DJ, Allen-Taylor L, Hoffstad O, Berlin JA. Diabetic neuropathic foot ulcers and amputation. Wound Repair and Regeneration 2005;13:230–6.
[3] Moulik PK, Mtonga R, Gill GV. Amputation and mortality in new-onset diabetic foot ulcers stratified by etiology. Diabetes Care 2003;26:491–4.
[4] Valk GD, Kriegsman DM, Assendelft WJ. Patient education for preventing diabetic foot ulceration. Cochrane Database System Review 2005:CD001488 [Review 1].
[5] Bann CM, Fehnel SE, Gagnon DD. Development and validation of the Diabetic Foot Ulcer Scale-short form (DFS-SF). Pharmacoeconomics 2003;21:1277–90.
[6] Bastyr 3rd EJ, Price KL, Bril V. Development and validity testing of the neuropathy total symptom score-6: questionnaire for the study of sensory symptoms of diabetic peripheral neuropathy. Clinical Therapeutics 2005;27:1278–94.
[7] Bott U, Muhlhauser I, Overmann H, Berger M. Validation of a diabetes specific quality-of-life scale for patients with type 1 diabetes. Diabetes Care 1998;21:757–69.
[8] Bradley C, Todd C, Gorton T, Symonds E, Martin A, Plowright R. The development of an individualized questionnaire measure of perceived impact of diabetes on quality of life: the ADDQoL. Quality of Life Research 1999;8:79–91.
[9] Bril V, Perkins BA. Validation of the Toronto Clinical Scoring System for diabetic polyneuropathy. Diabetes Care 2002;25:2048–52.
[10] Carey MP, Jorgensen RS, Weinstock RS, Sprafkin RP, Lantinga LJ, Carnrike Jr CL, et al. Reliability and validity of the appraisal of diabetes scale. Journal of Behavioral Medicine 1991;14:43–51.
[11] Cornblath DR, Chaudhry V, Carter K, Lee D, Seysedadr M, Miernicki M. Total neuropathy score: validation and reliability study. Neurology 1999;53:1660–4.
[12] Dhawan V, Spratt KF, Pinzur MS, Baumhauer J, Rudicel S, Saltzman CL. Reliability of AOFAS diabetic foot questionnaire in Charcot arthropathy: stability, internal consistency, and measurable difference. Foot and Ankle International 2005;26:717–31.
[13] Evans AR, Pinzur MS. Health-related quality of life of patients with diabetes and foot ulcers. Foot and Ankle International 2005;26:32–7.
[14] Feldman EL, Stevens MJ, Thomas PK, Brown MB, Canal N, Greene DA. A practical two-step quantitative clinical and electrophysiological assessment for the diagnosis and staging of diabetic neuropathy. Diabetes Care 1994;17:1281–9.
[15] Garratt AM, Schmidt L, Fitzpatrick R. Patient-assessed health outcome measures for diabetes: a structured review. Diabetic Medicine 2002;19:1–11.
[16] Jacobson AM, de Groot M, Samson JA. The evaluation of two measures of quality of life in patients with type I and type II diabetes. Diabetes Care 1994;17:267–74.
[17] Meadows K, Steen N, McColl E, Eccles M, Shiels C, Hewison J, et al. The Diabetes Health Profile (DHP): a new instrument for assessing the psychosocial profile of insulin requiring patients—development and psychometric evaluation. Quality of Life Research 1996;5:242–54.
[18] Shen W, Kotsanos JG, Huster WJ, Mathias SD, Andrejasich CM, Patrick DL. Development and validation of the diabetes quality of life clinical trial questionnaire. Medical Care 1999;37(4) [Supplement AS45-66].
[19] Talbot F, Nouwen A, Gingras J, Gosselin M, Audet J. The assessment of diabetes-related cognitive and social factors: the multidimensional diabetes questionnaire. Journal of Behavioral Medicine 1997;20:291–312.
[20] Martin RL, Irrgang JJ, Burdett RG, Conti SF, Van Swearingen JM. Evidence of validity for the foot and ankle ability measure (FAAM). Foot and Ankle International 2005;26:968–83.
[21] Martin RL, Hutt DM, Wukich DK. Validity of the foot and ankle ability measure (FAAM) in Diabetes mellitus. Foot and Ankle International 2009;30:297–302.
[22] Bayliss EA, Bayliss MS, Ware Jr JE, Steiner JF. Predicting declines in physical function in persons with multiple chronic medical conditions: what we can learn from the medical problem list. Health Quality of Life Outcomes 2004;2:247.
[23] Cook CE. Clinimetrics corner: the minimal clinically important change score (MCID): a necessary pretense. Journal of Manual and Manipulative Therapy 2008;16:E82–83.
[24] Jaeschke R, Singer J, Guyatt GH. Measurement of health status, ascertaining the minimal clinically important difference. Controlled Clinical Trials 1989;10:407–15.
[25] Johanson NA, Liang MH, Daltroy L, Rudicel S, Richmond J. American Academy of Orthopaedic Surgeons lower limb outcomes assessment instruments. Reliability, validity, and sensitivity to change. Journal of Bone and Joint Surgery, American Volume 2004;86-A:902–9.
[26] Budiman-Mak E, Conrad KJ, Roach KE. The Foot Function Index: a measure of foot pain and disability. Journal of Clinical Epidemiology 1991;44:561–70.
[27] Hefferan G, Khan F, Awan N, Riordain CO, Corigan J. A comparison of outcome scores in os calcis fractures. Irish Journal of Medical Sciences 2000;169:127–8.
[28] Karlsson J, Peterson L. Evaluation of the ankle joint function: the use of a scoring scale. The Foot 1991;1:15–9.
[29] Roos EM, Brandsson S, Karlsson J. Validation of the foot and ankle outcome score for ankle ligament reconstruction. Foot and Ankle International 2001;22:788–94.
[30] Beaton DE, Bombardier C, Katz JN, Wright JG, Wells G, Boers M, et al. Looking for important change/differences in studies of responsiveness. OMERACT MCID Working Group. Outcome measures in rheumatology. Minimal clinically important difference. Journal of Rheumatology 2001;28:400–5.
[31] Wells G, Anderson J, Beaton D, Bellamy N, Boers M, Bombardier C, et al. Minimal clinically important difference module: summary, recommendations, and research agenda. Journal of Rheumatology 2001;28:452–4.
[32] Hays RD, Woolley JM. The concept of clinically meaningful difference in health-related quality-of-life research. How meaningful is it? Pharmacoeconomics 2000;18:419–23.
[33] Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. Physical Therapy 2006;86:735–43.
[34] Brunner HI, Higgins GC, Klein-Gitelman MS, Lapidus SK, Olson JC, Onel K, et al. Minimal clinically important differences of disease activity indices in childhood-onset systemic lupus erythematosus. Arthritis Care Research 2010;62:950–9.
[35] Martin RL, Philippon MJ. Evidence of reliability and responsiveness for the hip outcome score. Arthroscopy 2008;24:676–82.
[36] Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR. Methods to explain the clinical significance of health status measures. Mayo Clinic Proceedings 2002;77:371–83.
[37] Cella D, Bullinger M, Scott C, Barofsky I. Group vs individual approaches to understanding the clinical significance of differences or changes in quality of life. Mayo Clinic Proceedings 2002;77:384–92.
[38] Ware Jr JE, Bayliss MS, Rogers WH, Kosinski M, Tarlov AR. Differences in 4-year health outcomes for elderly and poor, chronically ill patients treated in HMO and fee-for-service systems. Results from the Medical Outcomes Study. Journal of the American Medical Association 1996;276:1039–47.