

**The MOS SF-36 Health Survey:
A Validation Study With A Turkish Sample**

**Thesis submitted to the
Institute of Social Sciences
in partial fulfillment of the requirements for the degree of**

Master of Arts

in

Psychology

Bogazici University Library



39001100318826

by

Ahmet Cengizhan Demirsoy

Boğaziçi University

1999

ACKNOWLEDGMENTS

This thesis is one of the products of a long and hard, but full of with enthusiasm and pleasure, journey. On my path, I received guidance and warmth from many people. Although it is impossible to acknowledge all of these persons here, I would like to express my gratitude to all of them, especially to Güliz Elal for her kind and "reinforcing" attitude, and to Hamit Fişek, Giovanni Chimienti and Reşit Canbeyli, who taught me much about scientific inquiry.

Another person that requires special consideration is Güler Fişek. She was (is) my teacher and thesis advisor, but she was much more than these; she was always a "holding environment".

İpek, who is now 5 years old, was not present at the beginning. Later on she helped me by becoming a source of joy and by letting me enter the study room (though it was not always easy to get permission from her). And finally Çiğdem. Without her it would be impossible to carry on. She removed all the obstacles and she was always available for emotional and technical support.

ABSTRACT

The MOS SF-36 Health Survey: A Validation Study with a Turkish Sample

by

Ahmet Cengizhan Demirsoy

The main aims of the present study were to assess the reliability and validity of the Turkish translation of the MOS SF-36 Health Survey (SF-36). Reliability assessment yielded high Cronbach alpha values for all the scales, ranging from .75 for the role-emotional scale (RE) to .90 for the mental health scale (MH).

Principle components analysis, a technique of psychometric validation, carried out on the individual item scores identified 7 relevant factors, and, with the exception of physical functioning scale (PF) in which the items were spread to two factors, no scale was spread across more than a single factor. The same analysis conducted on the scale scores identified 2 factors revealing the bi-dimensionality, namely the physical and mental dimensions, of the SF-36 as proposed by its authors. Further, the associations between the scales and factors were found similar to the results obtained from other two studies conducted in the U. S. A and Australia.

In order to test the clinical validity of the SF-36 it was expected that it should differentiate between groups of subjects with different health status (healthy, physically disordered, and mentally disturbed) in the hypothesized manner. It was found that, with two exceptions, all the results were in line with the hypotheses of the present study. In the light of these favorable results it is concluded that the Turkish version of the SF-36 may serve as a useful measure of health status.

ÖZET

SF-36 Sağlık Taraması: Türk Örneklem Grubu ile yapılan bir Geçerlik Çalışması

Ahmet Cengizhan Demirsoy

Bu çalışmanın temel amacı, SF-36 Sağlık Taramasının (SF-36) Türkçe versiyonunun güvenilirlik ve geçerliğini araştırmaktır. Güvenirlik ölçümü, dağılım emosyonel-rol ölçeği (RE) için .75'ten mental sağlık ölçeği (MH) için .90'a olmak üzere, tüm ölçeklerde yüksek Cronbach alfa değerleriyle sonuçlanmıştır.

Psikometrik geçerliği araştırma tekniklerinden biri olan Ana Bileşenler Analizi (Principle Components Analysis) item puanlarına uygulanmasıyla 7 faktör ortaya çıkmıştır. İtemleri 2 faktöre dağılan fiziksel fonksiyon ölçeği (PF) dışında hiçbir ölçeğin itemleri birden fazla faktöre dağılmamıştır. Aynı tekniğin ölçek puanlarına uygulanması 2 faktörün varlığını ortaya çıkararak yazarlarının iddia ettiği gibi SF-36'nın iki boyutluluğunu (fiziksel ve mental) ortaya koymuştur. Ayrıca, faktörler ve ölçekler arasındaki korelasyonlar A.B.D ve Avustralya'da yapılan iki çalışmanın sonuçlarıyla genelde benzerlik göstermiştir.

SF-36'nın klinik geçerliği, ölçeklerinin 3 farklı sağlık durumundaki denekleri (sağlıklı, fiziksel bozukluk, ve mental bozukluk) hipotezlere uygun olarak ayırt edip etmediği araştırılarak yapılmıştır. İki istisna dışında tüm bulgular çalışmanın hipotezleri doğrultusunda sonuçlanmıştır. Bu olumlu bulgular ışığında SF-36'nın Türkçe versiyonunun sağlık durumunu değerlendirmede yararlı olabileceği sonucuna varılmıştır.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iii
ABSTRACT	iv
ÖZET	v
LIST OF TABLES	vii
INTRODUCTION	1
AIMS OF THE STUDY	24
METHOD	27
<u>Subjects</u>	27
<u>Measurement Instruments</u>	29
<u>Procedure</u>	30
RESULTS	33
DISCUSSION	38
REFERENCES	44
APPENDICES	59
<u>Appendix A</u>	60
The MOS SF-36 Health Survey (English)	
<u>Appendix B</u>	66
The MOS SF-36 Health Survey (Turkish)	
<u>Appendix C</u>	72
Symptom Checklist	

LIST OF TABLES

TABLE 1	17
Information About SF-36 Health Status Scales and the Interpretation of Low and High Scores.	
TABLE 2	28
Description of the Sample.	
TABLE 3	33
Reliability of the SF-36 Scales as Measured by Chronbach's Alpha.	
TABLE 4	34
Factor Coefficients of Individual Questions After Rotation.	
TABLE 5	35
Hypothesized Associations Between the SF-36 Scales and the Physical and Mental Components, and Results From the American (McHorney et al., 1993), Australian (McCallum, 1995), and Turkish Samples.	
TABLE 6	36
Means and Standard Deviations of SF-36 Scale Scores For the Entire Sample.	
TABLE 7	37
Summary of the Univariate F Tests of Health Condition on Scale Scores.	

INTRODUCTION

Changing constructions of health: absence of disease versus quality of life.

Health is our most desired and sought after state of being. Proverbs such as “Health is the best wealth”, “Health is better than wealth”, “Herşeyin başı sağlık”, and many others, point out this fact. In a study which attempted to determine the most valued end states of existence it was found that subjects unanimously valued health more than any other state (Rokeach, 1973, cited in Kaplan, 1985). Observation of our daily activities also reveal the importance attributed to health. In order to achieve it we devote a significant proportion of our energies and resources; we diet, exercise, and use a wide variety of products. When it fails, we seek services from an expensive and complex health care system.

Despite the centrality of health in our daily lives, the concept itself is not well defined (Kaplan, 1985; Sarafino, 1998; Ware, 1987). For example, McDermott (1969, cited in McDermott, 1981) states that “Health, like happiness, cannot be defined in exact measurable terms because its presence is so largely a matter of subjective judgment.”. When there is a lack of established consensus on a definition, examining the concepts, procedures, and measures used in the related field may provide a better understanding on the subject matter; and it may even be possible to deduce a definition from this examination.

Epidemiology is the study of the frequency and distribution of disease in a population. Disease refers to a set of biochemical and morphologic changes within the body which may or may not give rise to symptoms (McDermott, 1981). In

epidemiologic research, data are gathered about the rates of disease and possible correlates of disease in a large sample or population.

Mortality (death) and morbidity (disease) are the major outcome measures in many studies in epidemiology (Kaplan, 1985). Mortality is expressed in the form of a rate which is the proportion of deaths from a particular cause occurring within some defined time interval (usually per year). Usually mortality rates are expressed as age specific mortality. This means that the number of deaths in a particular age group is divided by the total number of persons in that same age group. A related statistic is the case fatality rate that is the proportion of persons who die of a particular disease divided by the total number with the disease (including those who die and those who live). The other main health status indicator morbidity is typically expressed for a specific condition and reported in terms of prevalence (the proportion of a population that has the disease being studied at a given point or period of time) or incidence rate (the number of new cases of a disease that occur in some period, usually a year). Epidemiological research, by using these indices, contribute to the identification of risk factors for and causes of disease and death..

Traditionally, the care of symptomatic patients in the medical setting primarily involves the establishment of a diagnosis by taking a history, performing an examination, and ordering tests, as well as the alleviation of symptoms by recommending specific pharmacologic and nonpharmacologic therapies. Medical consultations (other than those with mental health and some allied health professionals) traditionally are dominated by the "biomedical model," the perspective that the important forces impinging upon the patient's health are biological (e.g., disease) and behavioral (e.g., diet, alcohol consumption) in nature (Waitzkin, 1991,

cited in Street, Gold & McDowell, 1994). This paradigm focuses heavily on anatomical (e.g., diameter of the tumor) and pathophysiological parameters (e.g., high blood pressure) (McHorney & Tarlov, 1995). These traditional biomedical measures are used to capture the presence, absence, and / or severity of disease (Lerner, Levine, Malspeis & D'Agostino, 1994).

In clinical trials, the above mentioned parameters of morbidity and the other important health statistic mortality are also the traditional indices used to monitor the course of diseases, and to evaluate the effects of treatment interventions. For example, in the case of coronary artery disease, to monitor disease progression, clinical trials often use estimates of coronary artery stenosis, left ventricular ejection fraction, mortality, or a combination of these end points (Spertus, Winder, Dewhurst, Deyo & Fihn, 1994). The evaluation of an intervention such as heart transplantation is currently based on several outcomes including survival, graft function, rejection episodes, and other complications (Rector, Ormaza & Kubo, 1992).

In sum, in the main, studies in epidemiology and health care tend to focus on death and disease as the traditional outcomes. Biochemical and morphologic changes within the body are used in the diagnosis and monitoring of disease, and evaluation of treatment interventions. At the individual level, the severity of these indices and mortality, and at the community level prevalence, incidence, and mortality rates are used to estimate the burden of diseases, and determine health status. In conclusion, this brief overview of the field demonstrates that health is used to be defined primarily in terms of death and the extent of morbidity (Ware, 1987), and it makes one think that health is merely the absence of disease (i.e., biochemical and morphologic changes) and death.

However, this traditional conceptualization of health is not without criticisms.

A radical one claims that definitions that embrace the concept of “the absence of disease” are in reality misleading, for all living things are to some extent diseased - our crops, our lawns, our household pets and ourselves. Even our microbes are themselves diseased with still smaller microbes, the bacteriophages (McDermott, 1969, cited in McDermott, 1981). It is also stated that people are living longer, and the aging population is growing, and thus, medical care is increasingly concerned with the management of chronic disease, for which cure is impossible but death is a distant eventuality (Deyo, 1991; Lacey, Meier, Krumholz & Gusberg, 1995; McHorney & Tarlov, 1995). In these circumstances, morbidity and mortality statistics will be of limited value in defining the health status of individuals and communities.

It has also been argued that health is a multidimensional concept which embraces more than a physical component (Ware, 1987), the health domain ranges from negatively valued aspects of life, including death, to the more positively valued aspects such as role function or happiness (Guyatt, Feeny & Patrick, 1993), and for this reason, traditional clinical measures of health status concentrating solely on the improvement or deterioration of physiological and physical measurements describe only limited aspects of the patient’s life and do not provide a comprehensive assessment of health (Jette & Downing, 1994; Ruta, Garratt, Wardlaw & Russell, 1994).

Those who advocate a multidimensional model for conceptualizing health, refer to the definitions of health provided by the World Health Organization (WHO) and dictionary. The WHO defined health as “a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity” (WHO, 1948,

cited in Ware, 1987). Dictionary definitions also emphasize both physical and mental dimensions of health. The former pertains to the body and bodily needs, the latter refers to emotional and intellectual status. Health also connotes completeness - nothing is missing from the person - and proper function - all is working efficiently. Well-being -including soundness and vitality - is also part of the dictionary definition (Ware, 1987). Thus, both the WHO and dictionary definitions emphasize two important features of health; namely, the dimensionality of health, particularly the distinction between physical and mental components, and the full spectrum of health states ranging from disease to well-being. These definitions indicate that health is more than being diseased, and even much more than its physical aspect.

Another outstanding criticism involves the absence of the patient's point of view. It is claimed that whether it be because of the shortness of time of visits (Rubenstein et al., 1989); or the implicit assumption that the traditional parameters will be enough to explain the current status of the other aspects of patients' lives (Schipper, 1983), or the common belief that information obtained from the patients will be "soft" and inherently inferior to "hard" (i.e., objective) data obtained from the laboratory or endoscopy suite (Deyo, 1991), or as a result of training that focuses on organic pathology (Reifler, Kessler, Bernhard, Leon & Martin, 1996), the patients' point of view is neglected and clinicians have been the sole authority in evaluating the health status of patients. While individuals want their physicians to assess their functional performance and emotional well-being as part of medical care, by patients' reports this occurs infrequently (Schor, Lerner & Malspeis, 1995). It is claimed that traditional clinical measures of health status reflect clinicians' perception of disease (Ruta et al., 1994), may be of limited interest to patients (Guyatt et al., 1993), and

correlate poorly with patients' perceptions of health (Jette & Downing, 1996). Hence, a rheumatologist has commented that physicians "may all too easily spend years writing 'doing well' in the notes of a patient who has become progressively more crippled before their eyes" (Smith, 1983, cited in Deyo, 1991).

Although these criticisms did not provide a new, agreed upon definition of health, they gave impetus to an emerging conceptualization of health which is far broader than that is indicated by the traditional morbidity and mortality statistics. It is now considered that life has two dimensions: quantity and quality (Ware, 1987), and there is an emerging consensus that the goals of health care should be twofold: to increase the duration of life and to improve the quality of life (Deyo, 1991; Hayes, Morris, Wolfe & Morgan, 1995; Kaplan, 1985; McDermott, 1981). It is claimed that, while morbidity and mortality may be of value in terms of the duration of life, these indicators have little value in understanding how well people live in developed countries, and thus, the measurement of the second dimension of life requires another set of indicators (Ware, 1987).

Thus, with the increasing recognition that the quantity of life must be defined in the context of its quality (Lacey et al., 1995), in recent years, quality of life (QOL) has become a frequently mentioned concept in the health literature (Bergner 1989; Gill & Feinstein, 1994). The term is often assumed to be equivalent to and used interchangeably with concepts such as health-related quality of life (HRQOL), health status, and functional status (Bergner, 1989; Deyo, 1991; Gill & Feinstein, 1994; Guyatt et al., 1993; Kaplan 1985; Ware, Gandek, & The IQOLA Project Group 1994), and the same methodology will be adopted in this text. While the term QOL is traditionally used in a much broader sense which encompasses, in addition to health,

standard of living, the quality of housing, the neighborhood in which one lives, job satisfaction, and many other factors (Campbell, Converse, & Rogers, 1976, cited in Bergner, 1989; Lohr, 1992; Ware, 1987), in the health literature it has been used to describe the nature and extent of functional and psychosocial impairment in various disease states, to monitor quality of care, and to compare the effects of different interventions (Jambon & Johnson, 1997; Kaplan 1985).

Another important health care development that has taken place in recent years is an increasing consensus regarding the centrality of the patients' point of view (Geigle & Jones, 1990). In spite of the fact that the goal of medical care for most patients today is to obtain a more effective life (McDermott, 1981) and to preserve functioning and well-being (Ellwood, 1988; Public Policy Committee, American College of Physicians, 1988; Schroeder, 1987; Tarlov, 1983); and the patient's ultimate goal in seeking treatment and adhering to prescribed restorative, palliative, and preventive interventions is improvement in health status (Jette & Downing, 1994); and although the patient is usually the best judge of whether these goals have been achieved (Rubenstein, et al., 1989; Rothwell, McDowell, Wong & Dorman, 1997; Ware & Sherbourne, 1992; Ware, 1994), data concerning a patient's experiences of disease and treatment are not routinely collected in clinical research or medical practice (Lohr, 1992; Nelson & Berwick, 1989; Schor et al., 1995; Ware & Sherbourne 1992).

It has been claimed that, traditionally, physicians are not trained to think in terms of patient function (Rubenstein et al., 1989) and a study by Burns, Moskowitz, Ash, et al., (1992, cited in Lohr, 1992) which compared hospital records and the Index of Activities of Daily Living (ADL; a QOL measure which includes scales that

describe six functions: bathing, dressing, toileting, transfer, continence and feeding (Kaplan 1985)), have found that between 20% and 50% of information on specific ADLs was missing from hospital records and 10% of records had no functional status information at all. Likewise, another study (Schor et al., 1995) has revealed that the majority of physicians rarely or never ask about the extent to which patients' health limits their ability to perform everyday activities; neither do they inquire about limitations imposed by emotional problems, thus suggesting that the content of health care may be less comprehensive than physicians believe to be the case. Furthermore, when presented with the same set of QOL material, patients and clinicians have been found to differ in their assessments of the relative importance of different elements of QOL (Rothwell et al., 1997) clinicians could not adequately measure the the patients' QOL (Calkins et al., 1991; Nerenz, Repasky, Whitehouse & Kahkonen, 1992; Slevin, Plant, Lynch, Drinkwater & Gregory, 1988) and there was a considerable variability in results between different clinicians (Slevin et al., 1988).

It is also suggested that physicians tend to infer or extrapolate QOL information from traditional clinical indices (Deyo, 1991). QOL, however, may not always parallel these outcomes. For example, controlled blood glucose levels in diabetes patients, which is intended as a standard in diabetes care programs, may be associated with lower ratings in various QOL dimensions (Nerenz et al., 1992) or may not be associated with QOL scores (Tilly, Belton & McLachlan, 1995; Weinberger et al., 1994); hypertension may not be associated with negative effects on QOL measures (Jette & Downing, 1996; Krousel-Wood & Re, 1994; Stewart et al., 1989), and antihypertensive drugs may control blood pressure but worsen sexual functioning (Croog et al., 1986); and in chronic lung disorders, disease-specific outcomes such as

forced expiratory volume in one second may correlate poorly with broader indicators of disease impact on patients' lives, i.e., namely QOL measures (Guyatt, Berman, Townsend et al., 1987, cited in Viramontes & O'Brien, 1994).

Thus, whatever the reasons are, until recently, assessment of the effects of disease and their treatments on patients' lives has usually been, at the best, intuitive (Rubenstein et al., 1989) and informal (Wagner, Keller et al., 1995). At the worst, the case is as it is confessed by a clinician "Today, I know the MRI and other test results. But I don't know what happened to the last 100 patients we used the treatment on." (Geigle & Jones, 1990).

However, a new era is on the way in the health section for a period of time (Ellwood, 1988; Geigle & Jones, 1990; Tarlov, 1983). In recent years, there has been a shift in health care evaluation away from the purely biomedical model of ill health to a more holistic view of health that includes aspects of emotional and social well-being and functioning (Keoghane et al., 1996; McCallum, 1995; McDermott, 1981; O'Boyle, McGee, Hickey & O'Malley, 1992). It is now recognized that traditional morbidity and mortality data will not suffice to encapsulate these lately recognized aspects of health which are summated under the concept of QOL (Street et al., 1994). The fact that the patient is usually the best source of QOL information (Calkins et al., 1991) and recognition of the fact that medical outcomes belong first and foremost to patients (Paterson, 1996), re-located the patient at the center of the health practices. As a consequence, there has been increasing interest in the measurement of health from the patient's perspective (Aaronson et al., 1992; Keoghane et al, 1996). Finally, in this new era, traditional clinical measures and QOL pursuits are considered to be

complementary and synergistic (Schipper, 1983), and requisite for a comprehensive health status assessment (Ellwood, 1988; Aaronson et al., 1992).

Measurement of QOL

The past decade has seen a surge in the development of the science of QOL, and ways of measuring it (Brazier, Jones & Kind, 1993; Gill & Feinstein, 1994; Jambon & Johnson, 1997; Ware, 1995). In 1989, the U.S. Congress passed the Patient Outcome Research Act, which called for the establishment of a broad-based, patient-centered outcomes research program. In addition to traditional measures of survival, clinical endpoints, and disease- and treatment-specific symptoms and problems, the law mandated measures of "functional status and well-being and patient satisfaction" (Ware, 1995). Measures of QOL increasingly have become the instruments used by policy analysts and health care providers to compare the costs and benefits of competing ways of organizing, delivering, and financing health care services (Geigle & Jones, 1990; Lansky, Butler & Waller, 1992; Ware, 1995; Ware et al., 1986). Clinical investigators who are evaluating the effects of specific treatments and technologies, and practicing clinicians who are trying to achieve the best possible outcomes also began to utilize this information (Croog et al., 1986; Jambon & Johnson, 1997; Lohr, 1992; Nelson & Berwick, 1989; Solomon, Skobieranda & Genzen, 1995). The primary source of new information on health outcomes is rapidly becoming standardized patient questionnaires (Keoghane et al., 1996; Ware & Sherbourne, 1992; Ware et al., 1994).

In general, two basic approaches to the QOL measurement are available: generic instruments (including single indexes, health profiles, and utility measures) and specific instruments (Guyatt, 1995; Jambon & Johnson, 1997). These approaches are

not mutually exclusive (Guyatt et al., 1993; Solomon et al., 1995). Each approach has its strengths and weaknesses and may be suitable for different circumstances.

The single global question about health in the form of “How do you rate your health: excellent, good, fair or poor?”, which has been consistently used in the Australian Bureau of Statistics surveys (McCallum, 1995), is an example of a single index. Despite the general limitations of reliability and validity for single-item measures (Ware & Karmos, 1976, cited in McCallum, 1995), this self-rated health measure has proved to be a parsimonious measure of general health and it has a direct and independent prediction of survival (McCallum, Shadbolt & Wang, 1994, cited in McCallum, 1995). However, they yield very limited information (Guyatt et al., 1993).

More commonly, QOL instruments are questionnaires made up of a number of items or questions. These items are added up in a number of domains (also sometimes called dimensions). A domain or dimension refers to the area of behavior or experience that we are trying to measure. Domains might include mobility and self-care (which could be further aggregated into physical function), or depression, anxiety, and well-being (which could be aggregated to form an emotional-function domain).

Health profiles (also called general health surveys, health status surveys, or simply health surveys) are such questionnaires which attempt to encapsulate all important aspects of QOL (Guyatt et al., 1993; Jambon & Johnson, 1997). Thus, they contain questions on areas that it is assumed are concern to all (Keoghane et al., 1996). The Sickness Impact Profile is an example of a health profile and includes a physical dimension (with categories of ambulation, mobility, as well as body care and movement); a psychosocial dimension (with categories including social interaction,

alertness behavior, communication, and emotional behavior); and five independent categories including eating, work, home management, sleep and rest, as well as recreations and pastimes (Guyatt et al., 1993). Concepts addressed by the health profiles are not specific to any age, disease, or treatment group, thus they allow comparisons of the relative burden of different diseases and the relative benefits of different treatments (Ware & Sherbourne, 1992). However, they may not focus adequately on area of interest, and may not be responsive, i.e., sensitive to detect any differences in QOL over time or as a result of intervention (Guyatt, 1995).

Utility measures are derived from economic and decision theory and they reflect patient preferences in terms of both treatment and likely outcome (Guyatt et al., 1993). Utility scores reflect both health status of the patient and the value assigned to that status, and provide a summary score of the QOL gains from treatment minus the burdens of side effects (Jambon & Johnson, 1997). While utility measures allow cost-utility analysis, however, they may involve difficulty in determining utility values, do not allow examination of effect on different domains of QOL, and may not be responsive (Guyatt, 1995).

Specific instruments focus on aspects of health status that are specific to the area of primary interest. The instrument may thus be specific to a disease (such as asthma), to a particular patient population (such as the frail elderly), to a certain function (such as sexual function), or to a problem (such as back pain) (Guyatt et al., 1993). An example is The Clinical Dyspepsia Questionnaire which is devised on the basis of questions commonly used in the clinical assessment of patients with dyspepsia and other symptoms suggestive of a duodenal or gastric ulcer (Garratt et al., 1996). Although a generic health profile such as the SF-36 (Ware & Sherbourne, 1992) and

the above mentioned disease-specific instrument both contain items concerning pain, in the SF-36 they imply any pain whereas in the latter they imply only stomach pain. While specific questionnaires are clinically relevant, and they may be more responsive than the generic instruments, however, they have disadvantages including the inability to draw comparisons across conditions, and the failure to detect effects on health dimensions other than those that are being focused on and that impact on an individual's QOL (Guyatt, 1995).

A wide variety of such measures exist, of which the most frequently used are health profiles (Keoghane et al, 1996). These general health status measures have been used in a wide range of applications including descriptions of health profiles for patients differing in diagnosis, disease severity and treatment regimen; evaluations of the relative benefits of different treatments; comparisons of health outcomes across different health care delivery systems; assessments of health policy initiatives and measurement of general population health (McHorney & Tarlov, 1995; Ware, 1995). The widely used surveys cited by Ware (1995) are: Quality of Well-Being Scale (QWB) (Patrick, Bush & Chen, 1973), the Sickness Impact Profile (SIP) (Bergner et al., 1976), Health Insurance Experiment surveys (HIE) (Brook, Ware, Davies-Avery et al., 1979); Nottingham Health Profile (NPH) (Hunt & McEwen, 1980), Dartmouth Function Charts (COOP) (Nelson et al., 1987), European Quality of Life Index (EUROQOL) (EuroQOL Group, 1990), the MOS Functioning an Well-Being Profile (MOS-FWBP) (Stewart & Ware, 1992), and the MOS 36-Item Short-Form Health Survey (MOS SF-36 or SF-36) (Ware & Sherbourne, 1992).

It is claimed that standardized health surveys have the potential to become the new laboratory tests of medical practice (Ware, 1995). However, while routine use of

general health status measures in clinical settings cries out for the briefest measures possible, psychometric theory demands the exact opposite (McHorney, Ware, Rogers, Raczek & Lu, 1992). Of the above mentioned measures, MOS-FWBS, SIP, QWB, and HIE are lengthy instruments containing 149, 136, 107, and 86 items, respectively. NPH is shorter -38 items- but in contrast to the WHO definition of health, it defines perfect health and well-being as the absence of functional limitations and/or specific symptoms and problems (Ware, 1995), and thus does not satisfy the content validity (Ware, 1987). Moreover, it has been found to be unable to detect lesser health conditions, such as minor illnesses or disability (Brazier et al., 1992).

EUROQOL and COOP are short instruments both with 9 items (Ware, 1995). While there are few reports on the EUROQOL (Anderson, Aaronson & Wilkin, 1993), the results of a study (Brazier et al., 1993) suggest that it will be of limited value in descriptive studies of patient groups with less than major morbidity, or in evaluative research where small to moderate clinically significant changes (deterioration or improvement) in health are expected. As for the COOP charts, the conclusion is that this instrument has the ability to detect moderate effects in physical and emotional functioning, but how far it goes beyond this capability will rest on further research (Anderson et al., 1993). Thus, it seems that, of the above mentioned measures, only the SF-36 remains as a possible candidate which may take place in routine medical practice.

The SF-36 Health Survey: A Brief Measure of QOL

The SF-36 (Ware & Sherbourne, 1992) is one of several health status questionnaires developed in the U.S. by the Medical Outcomes Study (MOS). As the acronym suggests, it is a 36 item short form (SF) of a longer questionnaire.

The MOS can usefully be regarded as a development of an earlier U.S. project, the Health Insurance Experiment (HIE), a comprehensive evaluation of alternative methods of financing health care in the U.S. (Brook, Ware, Davies-Avery et al. 1979, cited in Dixon, Heaton, Long & Warburton, 1994). MOS was a 2-year observational study designed to help understand how specific components of the health care system affect the outcomes of care (Tarlov et al., 1989). Using as a point of departure the instruments which had proved useful during the HIE and other studies, MOS researchers selected, adapted, and developed the 149-item MOSFWBP, in order to measure health status. However, the other purpose of the MOS was to develop more practical tools for monitoring patient outcomes in routine practice (Tarlov et al., 1989). Initially an 18-item, and then, a 20-item short form version was developed from the original battery (Dixon et al., 1994).

Then, SF-36 was constructed, with the intent to broaden the health concepts measured and improve measurement precision over that achieved by the SF-20 (Ware & Sherbourne, 1992). To achieve breadth of measurement, measures of the most frequently studied functional status and well-being concepts described in accepted definitions of health status (Ware, 1987), are included. And to achieve precision in measuring each concept (depth), a short, multi-item scale from a subset of items shown to best reproduce a full-length measurement of proven validity, is constructed.

As summarized in Table 1, SF-36 includes one multi-item scale measuring each of eight health concepts: 1) physical functioning (PF); 2) role limitations because of physical health problems (RP); 3) bodily pain (BP); 4) social functioning (SF); 5) general mental health (psychological distress and psychological well-being) (MH); 6) role limitations because of emotional problems (RE); 7) vitality (energy/fatigue) (VT);

and 8) general health perceptions (GH). Each scale is scored from 0-100; a higher score represents better QOL. In addition, there is one item measuring self-reported health transition, although it is not used to score any of the scales. SF-36 produces a profile of health and/or summary measures for physical and mental health; it is not designed to generate a single index.

SF is suitable for self-administration, computerized administration, or administration by a trained interviewer in person or by telephone, to persons age 14 and older (Ware et al., 1994). It can be completed in 5-10 minutes by most respondents. Disease-specific symptoms and problems are not included in the SF-36 because it is a generic measure; however, because of its brevity, there is room to add these and other specific measures in the same questionnaire without exceeding an acceptable length (see, for example [Kurtin, Davies, Meyer, DeGiacomo & Kantz, 1992; Martin, Marquis & Bonfils, 1994]).

Table 1. Information About SF-36 Health Status Scales and the Interpretation of Low and High Scores.

Concepts	No. of Items	Meaning of Scores	
		Low	High
Physical functioning (PF)	10	Limited a lot in performing all physical activities including bathing or dressing	Performs all types of physical activities including the most vigorous without limitations due to health
Role limitations due to physical problems (RP)	4	Problems with work or other daily activities as a result of physical health	No problems with work or other daily activities as a result of physical health, past 4 weeks
Social functioning (SF)	2	Extreme and frequent interference with normal social activities due to physical and emotional problems	Performs normal social activities without interference due to physical or emotional problems, past 4 weeks
Bodily pain (BP)	2	Very severe and extremely limiting pain	No pain or limitations due to pain, past 4 weeks
General mental health (MH)	5	Feelings of nervousness and depression all of the time	Feels peaceful, happy, and calm all of the time, past 4 weeks
Role limitations due to emotional problems (RE)	3	Problems with work or other daily activities as a result of emotional problems	No problems with work or other daily activities as a result of emotional problems, past 4 weeks
Vitality (VT)	4	Feels tired and worn out all of the time	Feels full of pep and energy all of the time, past 4 weeks
General health perceptions (GH)	5	Believes personal health is poor and likely to get worse	Believes personal health is excellent

Empirical studies of the SF-36 Health Survey

One of the objectives in constructing the SF-36 was achievement of high psychometric standards (Ware & Sherbourne, 1992). Extensive psychometric testing has been conducted in the United States, United Kingdom, and many other countries (Ware et al., 1994).

Measuring instruments are usually tested on 3 major criteria: reliability, validity and responsiveness (Dixon et al., 1994). Reliability refers to the stability of the measurements produced by an instrument, and validity refers to the extent an instrument measures what it was intended to measure (Fişek, 1993). Responsiveness refers to an instrument's ability to detect change (Guyatt et al., 1993).

In a general sample, SF-36's scales test-retest reliability coefficients, item-scale correlations, and Cronbach's α (Cronbach, 1951, cited in Jenkinson, Coulter & Wright, 1993) values ranged from .60 to .81, .74 to .93, and .73 to .96, respectively, with SF occupying the lowest value in all three measurements (Brazier et al., 1992). In a small subsample ($n=35$) of cardiac rehabilitation patients, test-retest reliability coefficients of the scales ranged from .70 to .84, with the exception of .37 which was obtained for RE (Jette & Downing, 1994). In the same study, Cronbach's α ranged from .72 to .85. In a broad patient sample, Garratt, Ruta, Abdalla, Buckingham & Russell (1993) found that item-scale correlation coefficients ranged from .55 to .78, thus satisfying Kline's criterion of .4 (Kline, 1986, cited in Garratt et al., 1993). The correlation coefficients between items and the remainder of their own scale were all higher than those with other scales, providing further evidence of internal consistency. For all eight scales, Cronbach's α exceeded .80. In a sample of elderly outpatients with major depression, all item-scale correlations exceeded .38, items correlated more

highly with their own scale than with competing scales, and all Cronbach's α statistics exceeded .64 ($r=.64-.91$) (Beusterien, Steinwald & Ware, 1996; Heiligenstein et al., 1995).

There are other studies which provide favorable reliability statistics, but only in terms of Cronbach's α coefficient. An α of .50 or above is usually deemed acceptable for group comparisons (Helmstadter, 1964, cited in Jette & Downing, 1994), although Nunnally (1978, cited in Jenkinson et al., 1993) recommends values of .70 or above. In practice, however, well used scales should ideally gain α values in excess of .80 (Carmines & Zeller, 1979, cited in Jenkinson et al., 1993). Ranges of α values for the SF-36 scales obtained in these studies are listed below with information regarding the kind of sample, and the type of version used wherever necessary: .76-.90 in a large community sample (Jenkinson et al., 1993), with the anglicised version; .77-.89 in an Australian general sample with the Australian version (McCallum, 1995); .83-.94 in an elderly population (Lyons, Perry & Littlepage, 1994); .62-.90 in dialysis outpatients (Kurtin et al., 1992); .77-.91 in diabetic patients (Anderson, Fitzgerald, Wisdom, Davis & Hiss, 1997; Jacobson, de Groot & Samson, 1994) and .63-.91 in diabetic Pima Indians (Johnson, Nowatzki & Coons, 1996); .64-.93 in black and .73-.93 in white patients with acute chest pain (Johnson et al., 1995); .80 or above except for SF (.70) and GH (.76) in hypertensives (Krousel-Wood & Re, 1994); .64-.91 in patients with Parkinson's disease (Jenkinson, Peto, Fitzpatrick, Greenhall & Hyman, 1995); .42-.98 in asthmatics (Bousquet, Knani et al., 1994) and .77-.91 in patients with allergic rhinitis (Bousquet, Bullinger et al., 1994), with the French version; .77-.94 in manic-melancholic patients (Thunedborg, Black & Bech, 1995), with the Danish translation; .74-.89 in patients with Hodgkin's disease (van

Tulder, Aaronson & Bruning, 1994), with the Dutch translation; and above .70 with the exception of GH, ranging from .64 to .75, in healthy and patient samples (Bullinger, 1995), with the German translation.

Cronbach's α has also been measured for different modes of administration. For example, in a national survey, values ranged from .63 to .93 in the mail survey and from .77 to .93 in the telephone survey. Reliability estimates between the mail and telephone surveys were statistically different for five of the eight items (McHorney, Kosinski & Ware, 1994). However, in a comparison of telephone, face-to-face, and self-administration in patients, Cronbach's alphas (with ranges .67-.92, .75-.89, and .77-.93, respectively) did not differ by the method of administration (Weinberger, Oddone, Samsa & Landsman, 1996).

The SF-36 also has been evaluated for content validity, and scale scores have been tested empirically for construct and criterion validity. Content validity was established by comparing the content of the SF-36 with that of other widely used surveys (Ware, 1995) and published standards (Ware, 1987). The construct validity (the extent to which a measure correlates with variables in a manner consistent with theory) and criterion validity, particularly in terms of clinical criteria (the extent to which a measure corresponds to accepted clinical standards for determining the presence of a condition or its severity) of the SF-36 have been demonstrated in numerous studies, some of which are mentioned below.

SF-36 has been found to discriminate a wide variety of patient populations from healthy controls or general population norms, some of which are: those with shoulder disease (Matsen, Ziegler, DeBartolo & Seattle, 1995), trauma (McCarthy et al., 1995), migraine (Osterhaus, Townsend, Gandek & Ware, 1994), Parkinson's

disease (Jenkinson, Peto et al., 1995), backpain (Bullinger, 1995), lower urinary tract symptoms (Hunter, McKee, Black & Sanderson, 1995), and depression (McCallum, 1995). Further, SF-36 could also discriminate patients with different diseases (Garratt, Ruta et al., 1993; Lyons, Lo & Littlepage, 1993; McCallum, 1995; McHorney, Ware & Raczek 1993), and, those with and without comorbid conditions (Johnson et al., 1996; McCallum, 1995; McCarthy et al., 1995; McHorney et al., 1992). As expected, SF-36 has demonstrated poorer QOL in the very elderly than the younger ones and among hospital outpatients compared with general practice patients (Hayes et al., 1995); in the long-standing illness reporters compared to those who do not report (Jenkinson et al., 1993); in those who had consulted a physician recently compared to those who had not (Brazier et al., 1992; Jenkinson et al., 1993); in lower social class respondents in comparison to higher social class respondents (Brazier et al., 1992; Jenkinson et al., 1993); and in frequent health service users compared to infrequent users (Wagner, Phillips, Radford & Hornsby, 1995).

In many disease states, SF-36 has been found to correlate well with the routinely used established clinical indices. For example, the frequency and severity of complications in diabetic patients (Anderson et al., 1997; Jacobson et al., 1994); Hamilton Depression Scale in depressive patients (Heiligenstein et al., 1995; Thunedborg et al., 1995); the Hoehn and Yahr Scale in patients with Parkinson's disease (Jenkinson, Peto et al., 1995); clinical grading and pulmonary function tests in asthmatics (Bousquet, Knani et al., 1994); clinical ratings of dyspnea (Mahler & Mackowiak, 1995), and the Oxygen Cost Diagram and clinical severity classification (Viramontes & O'Brien, 1994) in patients with chronic lung disease; the Lysholm Knee Rating Scale (Katz, Harris et al., 1992) and the Iowa knee scores (Marsh, Smith

& Do, 1995) in patients with knee osteoarthritis; and clinical data items in heart valve patients (Phillips & Lansky 1992) correlated significantly with many of the scale scores of the SF-36.

Responsiveness of the SF-36 also has been demonstrated in a variety of conditions. SF-36 has been shown to be sensitive in detecting improvements due to drug or surgical treatments or specialized care in patients with depression (Beusterien et al., 1995; Coulehan, Schulberg, Block, Madonia & Rodriguez, 1997; Heiligenstein et al., 1995; Walker et al., 1995); asthma (Okamoto, Noonan, DeBoisblanc & Kellerman, 1995; Vollmer et al., 1997); renal failure (Benedetti et al., 1994); sciatica (Patrick et al., 1995); spinal deformity (Albert, Purtill, Mesa, McIntosh & Balderston, 1995); total hip arthroplasty (Katz, Larson, Phillips, Fosse & Liang, 1992); gallbladder disease, hemorrhoids, inguinal and incisional hernias, or severe obesity (Temple et al., 1995); migraine (Jhingran et al., 1994; Solomon et al., 1995); meniscal tears (Katz, Harris et al., 1992); and, heart diseases (Phillips & Lansky, 1992; Rector et al., 1993).

SF-36 has been found to correlate well with other generic QOL questionnaires such as the SIP (Katz, Larson et al., 1992; Weinberger et al., 1991), the EUROQOL (Brazier et al., 1993), and the COOP charts (Jenkinson, Lawrence, McWhinnie & Gordon, 1995). SF-36 also has been found to correlate well with specific QOL questionnaires (see, for example, [Garratt, Macdonald et al., 1993; Marquis, Fayol, Joire & Leplege, 1995; Patrick et al., 1995; Ruta et al., 1995]). Indeed, SF-36 is considered as "an established gold standard measure of health status" (Ruta et al., 1994) to validate subsequently developed QOL measures.

In sum, favorable outcomes regarding the reliability, validity and/or responsiveness of SF-36 have been documented in a wide variety of conditions. These results and the practicality (i.e., its applicability in different modes of administration, and its brevity) of the SF-36 have contributed to its widespread use, particularly in clinical settings.

Given the widespread usefulness of the SF-36, demand for validated translations have increased in recent years. In response, the International Quality of Life Assessment (IQOLA) Project, a 4-year project to translate and adapt the SF-36 in 15 countries, with the objectives of validating, norming, and documenting the new translations in as many countries as possible, was launched in 1991 (Aaronson et al., 1992; Ware, Gandek, Keller & IQOLA Project Group, 1995). Although there is much interest and demand for the SF-36 (Kaynar, 1996; Ünalán, 1996; personal communications), Turkey was not included in the project (Ware et al., 1994). However, Fişek (1996, personal communication) has translated the SF-36 into Turkish, tested the item scaling assumptions in a pilot study, and this translation has been approved by the Medical Outcomes Trust (Gandek, 1997).

AIMS OF THE STUDY

The main aims of the present study were to assess the Turkish version of the SF-36 in terms of reliability and validity. In this study Cronbach's alphas were measured to assess the reliability of each SF-36 scale. Reliability is considered acceptable for group comparisons when alpha exceeds 0.7 (Nunnally, 1978, as cited in Jenkinson et al., 1993).

Tests of validity can be divided into tests based on psychometric criteria and tests based on clinical and social criteria (Garratt, Ruta et al., 1993). Principle components analysis, a technique of psychometric validation, assesses the agreement between hypothetical factors that go to make up the measure and the scales designed to assess those factors. Within such an assessment a factor is considered relevant only if its "eigenvalue" (a statistical measure of its power to explain variation between subjects) exceeds 1.1 (Jolliffe, 1986, cited in Garratt, Ruta et al., 1993).

In the present study principle components analysis was conducted twice; on the items and on the scales. In the first case, it was expected that, if the SF-36 is a valid measure for use in Turkey, the scales defined by its authors should emerge from the analysis of the Turkish sample data, and items relating to a particular scale should be grouped together within a single factor.

In the second case, it was expected that the analysis should result in two factors to reveal the bi-dimensionality, namely the physical and mental dimensions, of the SF-36 as proposed by its authors. In addition, it was also expected that this analysis should replicate the results found with an U.S. sample (McHorney et al., 1993) and an Australian sample (McCallum, 1995). In these studies, the associations between scales and factors were hypothesized and found as follows: for the PF, RP

and BP, strong correlations with the physical component but weak correlations with the mental component; the reverse for the MH and RE; for the SF, a strong correlation with the mental component and a moderate correlation with the physical component; and for the VT and GH, moderate correlations with both physical and mental health components (See Table 5).

In order to test the clinical validity of the SF-36, it was expected that it should differentiate between groups of subjects with different health status (healthy, physically disordered, and mentally disturbed) in the hypothesized manner. The hypotheses and their respective rationales are presented below:

1. In other studies, it was found that back pain, arthritis, or sciatica patients scored significantly lower than the healthy subjects on each of the 8 parameters measured by the SF-36 (Bullinger, 1995; Garratt, Ruta et al., 1993; Lyons, Lo et al., 1994). Similarly, most of the studies conducted with patients who have a mental disturbance demonstrated that they also scored significantly lower than the healthy subjects on each scale of the SF-36 (Reifler et al., 1996; Lyons, Lo et al., 1994; Coulehan et al., 1997; Philbrick, Connelly & Wofford, 1996). Thus, it is hypothesized that both the physically disordered and the mentally disturbed groups will score significantly lower than the healthy group on each of the SF-36 scales ($p < 0.05$).

2. In two studies which compared back pain patients with general population norms or healthy subjects, the largest differences were found usually in the PF, RP and BP scales (Bullinger, 1995; Garratt, Ruta et al., 1993), which were proposed as the scales that best reflect physical health (McHorney et al., 1993; Ware, Kosinski & Keller, 1996). In two studies which compared depressed elderly with national norms,

it was found that the scores of the depressed elderly were particularly low in MH, RE, SF, and VT (Beusterien et al., 1996; Heiligenstein et al., 1995).

However, in one study, it was found that RP was one of the scales in which depressive patients especially scored lower than population norms (Coulehan et al., 1997), and in another study, anxiety patients scored lower than back pain patients (Lyons, Lo et al., 1994). While these results cause confusions in terms of hypotheses about the expected differences between the physically disordered and the mentally disturbed subjects, nevertheless, in the present study expectations were based on the findings and claims stated in the previous paragraph, and thus, the hypotheses were formulated as follows:

- a. PF, RP, and BP scale scores of the physically disordered group will be significantly lower than the scores of the mentally disturbed group ($p < 0.05$),
- b. MH, RE, SF, and VT scale scores of the mentally disturbed group will be significantly lower than the scores of the physically disordered group ($p < 0.05$).

METHOD

Subjects:

In the present study 238 students from Istanbul University served as subjects. Their ages ranged from 17 to 29 years with a mean of 20.9, and a standard deviation of 2.24. These subjects fell into three exclusive groups, namely, the healthy group, the physically disordered group, and the mentally disturbed group. The criteria used to locate subjects into these groups are described below in the procedure section. Statistics regarding age and gender of subjects in each group are given in Table 2. Although information regarding occupational and marital status were also asked, these were not considered since almost none of the subjects reported occupational status, and all subjects, except 5, reported as unmarried.

Table 2. Description of the Sample.

		<u>Total</u>	<u>Healthy</u>	<u>Physical</u>	<u>Mental</u>
	No.	238	101	71	66
	%		42.4	29.8	27.7
<u>Total</u>	Age R.	17-29	17-25	18-25	18-29
	Mn. Age	20.9	19.9	21.2	21.9
	SD	2.24	1.94	1.63	2.64
	No	96	41	28	27
	%	40.3	40.6	39.4	40.9
<u>Female</u>	Age R.	18-29	18-24	19-25	18-29
	Mn. Age	20.7	19.9	21.1	21.6
	SD	2.20	2.03	1.52	2.65
	No	142	60	43	39
	%	59.7	59.4	60.6	59.1
<u>Male</u>	Age R.	17-28	17-25	18-25	18-28
	Mn. Age	21	19.9	21.3	22.2
	SD	2.27	1.89	1.71	2.64

Note: The terms "healthy", "physical", and "mental" denote to the healthy, physically disordered, and mentally disturbed subjects, respectively.

Measurement Instruments:

In this study, subjects were given two questionnaires. The first questionnaire was a self-reported symptom checklist which was a modified version of the one that was used by Osterhaus et al., (1994). It consisted of 2 parts (see Appendix C). In the first part, subjects were asked "Has a doctor EVER told you that you had any of the following conditions," followed by a list consisting hypertension, recent myocardial infarction, congestive heart failure, diabetes, angina, cancer (except skin cancer), recent psychiatric disorder (e.g., depression in the last year), and others (e.g., multiple sclerosis, goiter). This question was asked with the intent to evaluate whether or not the subjects' past health status had an influence on their present health status.

In the second part, subjects were asked "Do you NOW have (in the last month) any of the following conditions," followed by a list that included arthritis or rheumatism, sciatica or chronic back problems, headache, chronic lung disease, a gastrointestinal disorder (such as gastritis, ulcer, enteritis, or colitis), dermatitis and other chronic skin rash, a psychological problem (e.g., sorrow, depression, anxiety, stress, alcohol overconsumption, etc., which is severe enough to make you think of consulting a psychiatrist or psychologist), frequent or continuous drug use (in that case name of the drug), and others. In this second section, it was stated that subjects should mark an item, only if, in the preceding month, a given complaint has been severe enough to make them think of consulting a doctor. The aim was to identify subjects' present health status.

The second questionnaire used in the study was the Turkish translation of SF-36 (see Appendix B). As summarized in Table 1, SF-36 includes one multi-item scale measuring each of eight health concepts: 1) physical functioning (PF); 2) role

limitations because of physical health problems (RP); 3) bodily pain (BP); 4) social functioning (SF); 5) general mental health (psychological distress and psychological well-being) (MH); 6) role limitations because of emotional problems (RE); 7) vitality (energy/fatigue) (VT); and 8) general health perceptions (GH). Each scale is scored from 0-100; a higher score represents better quality of life (QOL). In addition, there is one item measuring self-reported health transition, although it is not used to score any of the scales. SF-36 produces a profile of health and/or summary measures for physical and mental health; it is not designed to generate a single index.

The Turkish translation has been obtained by using a forward-backward method, and this translation has been approved by the Medical Outcomes Trust (Fişek, 1996). The Turkish translation has been tested in terms of its item scaling assumptions in a pilot study, and while it was suggested that the translation of some items be reexamined, it was concluded that data appeared to meet scaling assumptions in a satisfactory manner overall (Gandek, 1997).

Procedure:

In each case, researchers announced that they were collecting data on health status of individuals for a masters thesis, told the individuals that their names were not required, and asked if they would like take part in this study. Upon their verbal consent, investigators administered the study material. The subjects were instructed first to fill in the first page which consisted the demographic data (gender, age, educational level, occupational and marital status) and the symptom checklists, and then to fill in the second page which consisted of the answer sheet for the SF-36.

For the selection of healthy subjects, the present researcher entered some of the courses of İstanbul University for which he could get permission from the course instructor. After the standard announcement, he delivered the questionnaires for self-completion to those individuals who agreed. Upon completion, the researcher collected them back. Those who did not mark any of the items in the symptom checklist were chosen as the subjects for the healthy group, and those who marked any of the items were excluded from the study.

In order to identify subjects with only physical disorder, the present researcher cooperated with the physician of the Physical Medicine and Rehabilitation outpatient clinic of the İstanbul University Student Health Service. The physician referred to the researcher the students whom he diagnosed as having a musculoskeletal disorder of at least one month duration. After their verbal consent, the researcher wrote down the physician's diagnosis on the symptom checklist page and then delivered the questionnaires to the students for self-completion. Those who marked any of the items concerning a psychiatric disorder or a psychological problem from the symptom checklists were excluded from the study. Those who marked only physical complaints (whether it be the one congruent with the diagnosis and/or others) were included in the study as the subjects of the physical disorder only group. In this group, the most common diagnoses were low back pain (n=30), myalgia (a disorder associated with pain and tenderness of muscle and adjacent connective tissue) (n=18), arthralgia (joint pain) (n=7), and soft tissue injury (n=6).

Psychiatrists (one female and one male) of the same health service also participated in this study. The psychiatrists examined the students consulting them, and they asked for the verbal consent of those whom they diagnosed as having a

psychiatric disorder. Then they wrote down the diagnosis on the symptom checklist page. The symptom checklists were either filled in by the patients themselves or by the psychiatrists. In the second case, the psychiatrist asked the questions and marked the appropriate items if he/she decided that the patients' answers were not related to the psychiatric diagnosis, and did not mark if the decision was that the answers were related to the psychiatric diagnosis. Then the psychiatrist asked for the self-completion of SF-36. While those patients who had a mark on any of the physical complaints were excluded from the study, those who had a mark only on the psychiatric disorder and/or the psychological problem item were included in the study as the subjects of the mental disorder only group. In this group, the most common diagnoses were depression (n=24), obsessive compulsive disorder (n=19), social phobia (n=9), and generalized or panic anxiety disorders (n=8).

RESULTS

The coefficient alpha values for the scales are given in Table 3. Scales had high internal consistency measures with Cronbach's alphas ranging from 0.75 for RE to 0.90 for MH.

Table 3. Reliability of the SF-36 Scales as Measured by Chronbach's Alpha.

<u>Scale</u>	<u>Alpha</u>
PF	.87
RP	.82
BP	.87
GH	.82
VT	.84
SF	.79
RE	.75
MH	.90

Principle components analysis carried out on the 35 items resulted with the extraction of 7 factors (See Table 4). With the exception of PF, in which the items were grouped under the second and the fourth factor, no scale was spread across more than a single factor. While items of the GH, RP, RE and BP scales grouped separately under the third, fifth, sixth and seventh factors, respectively, items of the MH, SF and VT scales grouped under the same factor, namely the first factor.

Table 4. Factor Coefficients of Individual Questions After Rotation.

F1	F2	F3	F4	F5	F6	F7
<u>MH</u>	<u>PF</u>	<u>GH</u>	<u>PF</u>	<u>RP</u>	<u>RE</u>	<u>BP</u>
9b .57	3a .65	1 .64	3g .59	4a .71	5a .67	7 .79
9c .75	3b .74	11a .69	3h .84	4b .54	5b .75	8 .78
9d .82	3c .77	11b .75	3I .87	4c .77	5c .75	
9f .77	3d .56	11c .68	3j .45	4d .65		
9h .86	3e .67	11d .65				
	3f .63					
<u>SF</u>						
6 .52						
10 .60						
<u>VI</u>						
9a .82						
9e .73						
9g .70						
9I .52						

To test the hypothesized dimensionality of the SF-36 scales, principle components analysis was carried out on the scales. The analysis resulted with the extraction of two factors. The first principle component accounted for 51.5% of the total measured variance and correlated highly with all eight scales (range = 0.57 for bodily pain to 0.83 for vitality, median = 0.72). Extraction of the second component increased the percentage of total variance explained from 51.5% to 67.6%.

The rotated factor matrix (see Table 5) revealed that the Physical Functioning (PF) and Bodily Pain (BP) scales correlated most highly with the second rotated component, and correlated weakly with the first component. The Role-Physical (RP) scale also correlated highly with the second component, but at the same time had moderate correlation with the first component.

The Mental Health (MH) and Role-Emotional (RE) scales correlated highly with the first rotated component and weakly with the second component. The Vitality (VT) and Social Functioning (SF) scales also correlated highly with the first component, but also had moderate correlation with the second component. Finally, the general health perceptions scale had moderate correlations with both components. Based on these patterns of correlations, the first and second components are interpreted as “mental” and “physical” health dimensions, respectively.

Table 5. Hypothesized Associations Between the SF-36 Scales and the Physical and Mental Components, and Results From the American (McHorney et al., 1993), Australian (McCallum, 1995), and Turkish Samples.

	<u>Hypothesized Association</u>		<u>Rotated Principle Components</u>					
	<u>Physical</u>	<u>Mental</u>	<u>Physical</u>			<u>Mental</u>		
			<u>USA</u>	<u>Aus.</u>	<u>Tur.</u>	<u>USA</u>	<u>Aus.</u>	<u>Tur.</u>
PF	+++	+	.88	.76	.84	.04	.09	.15
RP	+++	+	.78	.72	.73	.30	.21	<u>.37</u>
BP	+++	+	.77	.76	.82	.24	.11	.08
MH	+	+++	.12	.19	.14	.90	.85	.91
RE	+	+++	.19	-.04	.01	.81	.81	.74
SF	++	+++	.44	.33	.42	.71	.72	.73
VT	++	++	.59	.47	.31	.57	.64	<u>.82</u>
GH	++	++	.68	.64	.55	.32	.20	.42

Note: (+) Weak association ($r \leq .30$)

(++) Moderate to substantial association ($.30 < r < .70$)

(+++) Strong association ($r \geq .70$)

Average scores for each of the eight scale were computed and used for further analyses. Means and standard deviations of eight scale scores are presented in Table 6 for the entire sample.

Table 6. Means and Standard Deviations of SF-36 Scale Scores For the Entire Sample.

<u>Scale</u>	<u>Condition</u>					
	<u>Healthy</u>		<u>Physical</u>		<u>Mental</u>	
	<u>Mean</u>	<u>Sd</u>	<u>Mean</u>	<u>Sd</u>	<u>Mean</u>	<u>Sd</u>
RE	55.78	38.04	49.77	40.97	23.74	35.45
VT	59.16	16.70	47.39	18.44	33.86	18.46
MH	65.74	15.80	57.01	17.72	39.52	20.06
GH	71.50	17.17	54.93	21.27	47.86	21.02
BP	74.58	18.96	42.30	17.90	64.89	22.64
SF	78.22	18.34	61.80	21.74	48.30	27.90
RP	90.35	18.02	52.46	35.90	48.48	42.26
PF	95.99	6.60	71.97	21.85	82.88	18.31

To assess construct validity, a multivariate analysis of variance was carried out on the eight scale scores as a function of sex and health condition. It was found that there was no significant interaction between sex and health condition on their influence on scale scores [$F(16,452)=.79, p>.05$]. There was no significant multivariate main effect of sex [$F(8,225)=.55, p>.05$], but there was a significant effect of health condition on scale scores [$F(16,452)=17.14, p<.001$].

Table 7. Summary of the Univariate F Tests of Health Condition on Scale Scores. D. F. (2,232).

<u>SCALE</u>	<u>F</u>	<u>Sig. of F</u>
BP	53.09	.000
GH	28.98	.000
MH	41.78	.000
PF	44.18	.000
RE	14.25	.000
RP	42.69	.000
SF	34.40	.000
VT	37.22	.000

The univariate F tests showed that there was a main effect of health condition on all eight scale scores ($p < .001$). See Table 7. Tukey's HSD test was carried out on each scale score to compare all possible pairs of means. With the exception of the RE scale, it was found, as hypothesized, that the scores of the healthy group were significantly higher than those of the physically and the mentally disordered group on all scales ($p < .05$). For the RE scale, although the healthy group scored higher than the physically disordered group, the difference was not significant ($p > .05$).

As hypothesized, the physically disordered group scored significantly lower than the mentally disordered group on the BP and PF scales ($p < .05$), however, for the RP scale score the difference was not significant ($p > .05$). Again as hypothesized for the MH, RE, SF and VT scales, it was the mentally disordered group which scored significantly lower than the physically disordered group ($p < .05$). As for the GH scale, it was found that the scores of the physically disordered group and that of the mentally disturbed group did not differ significantly ($p > .05$).

DISCUSSION

The main aims of the present study were to assess the reliability and validity of the Turkish version of the SF-36. For the reliability assessment, Cronbach's alpha measures were used for each SF-36 scale. Validity was assessed in terms of psychometric and clinical criteria. For the psychometric validation, principle components analysis was carried out on the individual items and on the scales. As for the clinical validation, hypotheses concerning the differentiation of healthy, physically disordered, and mentally disturbed subjects were tested by carrying out multivariate analysis of variance and Tukey's HSD test on scale scores.

The reliability assessment of the SF-36 yielded highly satisfactory results. Although an alpha of .50 or above is usually deemed acceptable for group comparisons (Helmstadter, 1964, cited in Jette & Downing, 1994), in the present study, for all eight scales internal consistency measured by Cronbach's alpha exceeded .75, thus satisfying Nunnally's criterion of .70 (Nunnally, 1978, cited in Jenkinson et al., 1993). In practice, however, it is expected that well used scales should ideally have alpha values in excess of .80 (Carmines & Zeller, 1979, cited in Jenkinson et al., 1993). In the present study, with the exception of .75 for the RE and .79 for the SF scale, alpha values ranged from .82 to .90, thus satisfying this expectation.

The results of the principle components analysis carried out on the 35 items were also satisfactory overall, since, with the exception of the PF scale items which grouped under two factors, items relating to a particular scale grouped together within a single factor. In a postal survey of a broad sample of patients suffering from four clinical conditions (Garratt, Ruta et al., 1993), the principle components analysis

identified five relevant factors. In that study, it was found that no scale was spread across more than a single factor. However, although the PF, GH, and RE scales were grouped separately under the factors one, four, and five, respectively, the second factor represented the MH and VT scales, and the third factor represented the RP, SF and BP scales. The authors reported these results as a precise correspondence between factors and scales which is rare in factor analysis.

However, the case of the PF scale items is problematic since it does not readily lend itself to a probable explanation. For example, while item 3j ("Yıkanmak ya da giyinmek" "Bathing or dressing yourself") was related to the fourth factor, item 3f ("Eğilmek, diz çökmek, yerden birşey almak" "Bending, kneeling, or stooping") was related to the second factor although these behaviors can also be seen as part of the content of the previous item.

Principle components analysis carried out on the scale scores revealed the bi-dimensionality, namely the physical and mental dimensions, of the SF-36 as proposed by its authors. Further, the analysis also yielded highly similar results with the two studies that were carried out with a U.S. sample (McHorney et al., 1993) and an Australian sample (McCallum, 1995). The results of the present study differed from the results of these studies only in two respects: the RP scale correlated moderately with the mental component and the VT scale correlated highly with the mental component, in contrast to the other two studies in which the correlations were weak and moderate, respectively.

In the case of the RP scale, the correlation with the mental component that was found in the U.S study was .30, which was proposed as the upper limit value for

the weak association. Further, this value is close to the .37 which was found in the present study. In the case of the VT scale, in the Australian study it was found that the correlation of this scale with the mental component was .64, a value which is much above .47, the correlation between this scale and the physical component, and nearer a strong correlation. Thus it can be claimed that the results regarding these two scales did not deviate much from those found in the above mentioned two studies.

Further, although minor, the differences between the results of the present study and those of the two studies may be attributed to the differences in the sample characteristics. The samples of the earlier studies were older, and had a higher female percentage than the sample of the present study. Whatever the reason for the differences, it can be concluded that the present study and the other two studies yielded highly similar results in general.

The multivariate analysis of variance revealed that there was a main effect of health condition on all eight scale scores. Comparison of mean scale scores utilizing Tukey's HSD test revealed that (a) in line with hypothesis I, the mentally disturbed group scored significantly lower than the healthy subjects on each of the SF-36 scales; (b) in line with hypothesis I, the physically disordered subjects scored significantly lower than the healthy subjects on the PF, RP, BP, MH, SF, VT and GH scales; (c) but, as opposed to hypothesis I, on the RE scale, the physically disordered subjects did not score significantly lower than the healthy subjects; (d) in line with hypothesis IIa, the physically disordered subjects scored significantly lower than the mentally disturbed subjects on the PF and BP scales; (e) but as opposed to hypothesis IIa, on the RP scale, the physically disordered subjects did not score significantly lower than the

mentally disturbed subjects; and, (f) in line with hypothesis IIb, the mentally disturbed subjects scored significantly lower than the physically disordered subjects on the MH, RE, SF and VT scales. Thus, of the above findings, most were in the hypothesized direction and only two results deviated from the hypotheses of the present study.

As for the first case, i.e. clause (c), while in the majority of the previous research the results were in the direction of the hypothesis of the present study, there was one study which resulted similarly with this one. In a postal survey (McCallum, 1995) it was found that subjects who reported minor medical conditions did not score lower on the RE scale than the subjects with no condition, however, in contrast to the present study, it was also the case with the MH scale.

As for the second case, i.e. clause (e), the analysis further revealed that, although the difference was not significant, mean score for the mentally disturbed group on the RP scale (48.4) was lower than that of the physically disordered group (52.4). However, although this result is in contrast with the hypothesis of the present study, it is in line with the results of the two studies which are mentioned in the Aims of the Study section (Coulehan et al., 1997; Lyons, Lo et al., 1994). While statistical significance were not stated in these studies, in the first study (Coulehan et al., 1997) it was found that the RP was one of the scales in which depressive patients especially scored lower than the population norms, and in the second study (Lyons, Lo et al., 1994) it was found that anxiety patients scored lower than back pain patients on this scale.

Since both results involve role limitations, a probable explanation for the above results may lie in the difference between the physically disordered and the

mentally disturbed subjects in terms of their causal attributions towards role limitations. It may be argued that, the physically disordered subjects, realistically, attributed their role limitations to physical factors, but not to emotional factors. Thus, they scored significantly lower than the healthy subjects on the RP scale, but not on the RE scale, and they scored significantly higher than the mentally disturbed subjects on this scale. The mentally disturbed subjects, however, attributed their role limitations to emotional factors, and also to physical factors.

The explanation proposed above may raise two questions: (1) why the physically disordered subjects did not respond in a manner consistent with the hypothesis of the present study and the results of the majority of previous research mentioned above; and (2) why the mentally disturbed subjects responded in a manner consistent with the results of the two studies mentioned above, but inconsistent with the hypothesis of the present study.

As for the first question, a probable explanation for the deviation may be the differences between samples of the present study and those of the previous research in terms of the severity and duration of the disorders. Although these characteristics were not assessed in the present study and were not reported in previous research, in the present study the subjects were young university students participating in their courses at the time they consulted the physician. Thus, it can be speculated that, bearing in mind the common link between enduring pain and emotional disturbance (Eimer and Freeman, 1998), they did not suffer severely enough to live emotional disturbance and hence did not attribute their role limitations to emotional factors.

As for the second question a probable explanation may be that the expectation formulated in the hypothesis regarding the RP scale was not appropriate. A tentative explanation may be that, the mentally disturbed subjects phenomenologically experience their disturbance without differentiating between the emotional and physical aspects, and hence attribute their role limitations both to the emotional and physical factors.

In conclusion, in terms of the reliability and validity of the Turkish version of the SF-36, the present study yielded satisfactory results overall. However, further study with various samples are needed in order to generalize these results to the Turkish population. If such studies can be realized, the Turkish version of the SF-36 may come forward as a useful Quality of Life measure and, along with the traditional measures of health, it can achieve widespread use in clinical and epidemiological studies.

REFERENCES

- Aaronson, N. K., Acquadro, C., Alonso, J., Apolone, G., Bucquet, D., Bullinger, M., Bungay, K., Fukuhara, S., Gandek, B., Keller, S., Razavi, D., Sanson-Fisher, R., Sullivan, M., Wood-Dauphinee, S., Wagner, A., & Ware, J. E. (1992). International quality of life (IQOLA) project. Quality of Life Research, 1, 349-351.
- Albert, T. J., Purtill, J., Mesa, J., McIntosh, T., & Balderston, R. A. (1995). Health outcome assessment before and after adult deformity surgery. A prospective study. Spine, 20 (18), 2002-2004.
- Anderson, R. T., Aaronson, N. K., & Wilkin D. (1993). Critical review of the international assessments of health-related quality of life. Quality of Life Research, 2 (6), 369-395.
- Anderson, R. M., Fitzgerald, J. T., Wisdom, K., Davis, W. K., & Hiss, R. G. (1997). A comparison of global versus disease-specific quality-of-life measures in patients with NIDDM. Diabetes Care, 20 (3), 299-305.
- Benedetti, E., Matas, A. J., Hakim, N., Fasola, C., Gillingham, K., McHugh, L., & Najarian, J. S. (1994). Renal transplantation for patients 60 years of older. A single-institution experience. Annals of Surgery, 220 (4), 445-458.
- Bergner, M. (1989). Quality of life, health status, and clinical research. Medical Care, 27 (3), S148-156.
- Beusterien, K. M., Steinwald, B., & Ware, J. E. (1996). Usefulness of the SF-36 Health Survey in measuring health outcomes in the depressed elderly. Journal of Geriatric Psychiatry and Neurology, 9, 13-21.

- Bousquet, J., Bullinger, M., Fayol, C., Marquis, P., Valentin, B., & Burtin, B. (1994). Assessment of quality of life in patients with perennial allergic rhinitis with the French version of the SF-36 Health Status Questionnaire. Journal of Allergy and Clinical Immunology, 94 (2), 182-188.
- Bousquet, J., Knani, J., Dhivert, H., Richard, A., Chicoye, A., Ware, J. E., & Michel, F. B. (1994). Quality of life in asthma. I. Internal consistency and validity of the SF-36 questionnaire. American Journal of Respiratory Critical Care Medicine, 149, 371-375.
- Brazier, J. E., Harper, R., Jones, N. M. B., O'Cathain, A., Thomas, K. J., Usherwood, T., & Westlake, L. (1992). Validating the SF-36 health survey questionnaire: New outcome measure for primary care. British Medical Journal, 305, 160-164.
- Brazier, J. E., Jones, N. M. B., & Kind, P. (1993). Testing the validity of EuroQOL and comparing it with the SF-36 health survey questionnaire. Quality of Life Research, 2, 169-180.
- Bullinger, M. (1995). German translation and psychometric testing of the SF-36 health survey: Preliminary results from the IQOLA Project. Social Science and Medicine, 41 (10), 1359-1366.
- Calkins, D. R., Rubenstein, L. V., Cleary, P. D., Davies, A. R., Jette, A. M., Fink, A., Kosecoff, J., Young, R. T., Brook, R. H., & Delbanco, T. L. (1991). Failure of physicians to recognize functional disability in ambulatory patients. Annals of Internal Medicine, 114 (6), 451-454.
- Coulehan, J. L., Schulberg, H. C., Block, M. R., Madonia, M. J., & Rodriguez, E. (1997). Treating depressed primary care patients improves their physical,

- mental, and social functioning. Archives of Internal Medicine, 157, 1113-1120.
- Croog, S. H., Levine, S., Testa, M. A., Brown, B., Bulpitt, C. J., Jenkins, C. D., Klerman, G. L., & Williams, G. H. (1986). The effects of antihypertensive therapy on the quality of life. New England Journal of Medicine, 314 (26), 1657-1664.
- Deyo, R. A. (1991). The quality of life, research, and care. Annals of Internal Medicine, 114 (8), 695-696.
- Dixon, P., Heaton, J., Long A., & Warburton, A. (1994). Reviewing and applying the SF-36. Outcomes Briefing, 4, 3-25.
- Eimer, B. N., & Freeman, A. (1998). Pain management psychotherapy. A practical guide. John Wiley & Sons, Inc.
- Ellwood, P. M. (1988). Outcomes management: a technology of patient experience. New England Journal of Medicine, 318 (23), 1549-1556.
- Fişek, H. (1993). Lecture notes for a course on social/behavioral science research methods. Part 6. Boğaziçi University.
- Gandek, B. (1997). Psychometric analysis of the SF-36 Health Survey: Turkish data. Unpublished report.
- Garratt, A. M., Macdonald, L. M., Ruta, D. A., Russell, I. T., Buckingham, J. K., & Krukowski, Z. H. (1993). Towards Measurement of Outcome for Patients with Varicose Veins. Quality in Health Care, 2, 5-10.
- Garratt, A. M., Ruta, D. A., Abdalla, M. I., Buckingham, J. K., & Russell, I. T. (1993). The SF-36 health survey questionnaire: an outcome measure suitable for routine use within the NHS? British Medical Journal, 306, 1440-1444.

- Garratt, A. M., Ruta, D. A., Russell, I. T., Macleod, K., Brunt, P., McKinlay, A. Mowat, A., & Sinclair, T. (1996). Developing a condition-specific measure of health for patients with dyspepsia and ulcer-related symptoms. Journal of Clinical Epidemiology, 49 (5), 565-571.
- Geigle, R., & Jones, S. B. (1990). Outcomes measurement: A report from the front. Inquiry, 27, 7-13.
- Gill, T. M., & Feinstein A. R. (1994). A critical appraisal of the quality-of-life measurements. The Journal of the American Medical Association, 272 (8), 619-626.
- Guyatt, G. H., Feeny, D. H., & Patrick, D. L. (1993). Measuring health-related quality of life. Annals of Internal Medicine, 118 (8), 622-629.
- Guyatt, G. H. (1995). A taxonomy of health status instruments. The Journal of Rheumatology, 22 (6), 1188-1190.
- Hayes, V., Morris, J., Wolfe, C., & Morgan, M. (1995). The SF-36 health survey questionnaire: is it suitable for use with older adults? Age and Ageing, 24 (2), 120-125.
- Health and Public Policy Committee, American College of Physicians. (1988). Comprehensive functional assessment for elderly patients. Annals of Internal Medicine, 109 (1), 70-72.
- Heiligenstein, J. H., Ware, J. E., Beusterien, K. M., Roback, P. J., Andrejasich, C. & Tollefson, G. D. (1995). Acute effects of fluoxetine versus placebo on functional health and well-being in late-life depression. International Psychogeriatrics, 7, 125-137.

- Hunter, D. J., McKee, M., Black, N. A., Sanderson, C. F. (1995). Health status and quality of life of British men with lower urinary tract symptoms: results from the SF-36. Urology, 45 (6), 962-971.
- Jacobson, A. M., de Groot, M., & Samson J. A. (1994). The evaluation of two measures of quality of life in patients with type I and type II diabetes. Diabetes Care, 17 (4), 267-274.
- Jambon, B., & Johnson, K. I. (1997). Individual quality of life and clinical trials. Quality of Life Newsletter, 17, 1, 2, 16, 17.
- Jenkinson, C., Coulter, A., & Wright L. (1993). Short form 36 (SF-36) health survey questionnaire: normative data for adults of working age. British Medical Journal, 306, 1437-1440.
- Jenkinson, C., Lawrence, K., McWhinnie, D., & Gordon, J. (1995). Sensitivity to change of health status measures in a randomized controlled trial: comparison of the COOP charts and the SF-36. Quality of Life Research, 4 (1), 47-52.
- Jenkinson, C., Peto, V., Fitzpatrick, R., Greenhall, R., & Hyman, N. (1995). Self-reported functioning and well-being in patients with Parkinson's disease: comparison of the short-form health survey (SF-36) and the Parkinson's Disease Questionnaire (PDQ-39). Age and Ageing, 24 (6), 505-509.
- Jette, D. U., & Downing, J. (1994). Health status of individuals entering a cardiac rehabilitation program as measured by the medical outcomes study 36-item short-form survey (SF-36). Physical Therapy, 74 (6), 521-527.
- Jette, D. U., & Downing, J. (1996). The relationship of cardiovascular and psychological impairments to the health status of patients enrolled in cardiac rehabilitation programs. Physical Therapy, 76 (2), 130-139.

- Jhingran, P., Cady, R. K., Rubino, J., Miller, D., Grice, R. B., & Gutterman, D. L. (1996). Improvements in health-related quality of life with sumatriptan treatment for migraine. Journal of Family Practice, 42 (1), 36-42.
- Johnson, P. A., Goldman, L., Orav, E. J., Garcia, T., Pearson, S. D., & Lee, T. H. (1995). Comparison of the Medical Outcomes Study Short-Form 36-Item Health Survey in black patients and white patients with acute chest pain. Medical Care, 33 (2), 145-160.
- Johnson, J. A., Nowatzki, T. E., & Coons, S. J. (1996). Health-related quality of life of diabetic Pima Indians. Medical Care, 34 (2), 97-102.
- Kaplan, R. M. (1985). Quality-of-Life Measurement. In P. Karoly (Ed.), Measurement Strategies in Health Psychology (pp. 115-145). New York: Wiley.
- Katz, J. N., Harris, T. M., Larson, M. G., Krushell, R. J., Brown, C. H., Fossel, A. H., & Liang MH. (1992). Predictors of functional outcomes after arthroscopic partial meniscectomy. Journal of Rheumatology, 19, 1938-1942.
- Katz, J. N., Larson, M. G., Phillips, C. B., Fosse, A. H., & Liang, M. H. (1992). Comparative measurement sensitivity of short and longer health status instruments. Medical Care, 30 (10), 917-25.
- Keoghane, S. R., Lawrence, K. C., Jenkinson, C. P., Doll, H. A., Chappel, D. B. & Cranston, D. W. (1996). The Oxford Laser Prostate Trial: sensitivity to change of three measures of outcome. Urology, 47 (1), 43-47.
- Krousel-Wood, M. A., & Re, R. N. (1994). Health status assessment in a hypertension section of an internal medicine clinic. The American Journal of the Medical Sciences, 308 (4), 211-217.

Kurtin, P. S., Davies, A. R., Meyer, K. B., DeGiacomo, J. M., & Kantz, M. E.

(1992). Patient-based health status measures in outpatient dialysis. Early experiences in developing an outcomes assessment program. Medical Care, 30 (5 Suppl), MS136-149.

Lacey, K. O., Meier, G. H., Krumholz, H. M., & Gusberg, R. J. (1995). Outcomes after major vascular surgery: the patients' perspective. Journal of Vascular Nursing, 13 (1), 8-13.

Lansky, D., Butler, J. B. V., & Waller, F. T. (1992). Using health status measures in the hospital setting: From acute care to 'outcomes management'. Medical Care, 30 (5), MS57-73.

Lerner, D. J., Levine, S., Malspeis, S., & D'Agostino, R. B. (1994). Job strain and health-related quality of life in a national sample. American Journal of Public Health, 84 (10), 1580-1585.

Lohr, K. N. (1992). Applications of Health status assessment measures in clinical practice. Overview of The Third Conference on Advances in Health Status Assessment. Medical Care, 30 (5), MS1-14.

Lyons, R. A., Lo, S. V., & Littlepage, B. N. C. (1994). Comparative health status of patients with 11 common illnesses in Wales. Journal of Epidemiology & Community Health, 48 (4), 388-390.

Lyons, R. A., Perry, H. M., & Littlepage, B. N. (1994). Evidence for the validity of the short-form 36 questionnaire (SF-36) in an elderly population. Age and Ageing, 23, 182-184.

- Mahler, D.A. & Mackowiak, J. I. (1995). Evaluation of the short-form 36-item questionnaire to measure health-related quality of life in patients with COPD. Chest, 107 (6):1585-1589.
- Marquis, P., Fayol, C., Joire, J. E., & Leplege, A. (1995). Psychometric properties of a specific quality of life questionnaire in angina pectoris patients. Quality of Life Research, 4 (6), 540-546.
- Marsh, J. L., Smith, S. T., & Do, T. T. (1995). External fixation and limited internal fixation for complex fractures of the tibial plateau. Journal of Bone & Joint Surgery - American Volume, 77 (5), 661-673.
- Martin, C., Marquis, P., & Bonfils, S. (1994). A 'Quality of Life Questionnaire' Adapted to Duodenal Ulcer Therapeutic Trials. Scandinavian Journal of Gastroenterology, 206 (29 Suppl.), 40-43.
- Matsen III, F. A., Ziegler, D. W., DeBartolo, S. E., & Seattle, W. (1995). Patient self-assessment of health status and function in glenohumeral degenerative joint disease. Journal of Shoulder and Elbow Surgery, 4 (5), 345-351.
- McCallum, J. (1995). The SF-36 in an Australian sample: validating a new, generic health status measure. Australian Journal of Public Health, 19 (2), 160-166.
- McCarthy, M. L., MacKenzie, E. J., Bosse, M. J., Copeland, C. E., Hash, C. S., & Burgess, A. R. (1995). Functional status following orthopedic trauma in young women. The Journal of Trauma, 39 (5), 828-836.
- McDermott, W. (1981). Absence of indicators of the influence of its physicians on a society's health: impact of physician care on society. The American Journal of Medicine, 70 833-43.

- McHorney, C. A., Ware, J. E., Rogers, W. H., Raczek, A., & Lu, J. F. R. (1992). The validity and relative precision of MOS Short- and Long-Form Health Status Scales and Dartmouth COOP charts: Results from the Medical Outcomes Study. Medical Care, 30 (5), MS 253-265.
- McHorney, C. A., Ware, J. E., & Raczek, A. E. (1993). The MOS 36-item short-form health survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Medical Care, 31 (3), 247-263.
- McHorney, C. A., Kosinsky, M., & Ware, J. E. (1994). Comparisons of the costs and quality of norms for the SF-36 health survey collected by mail versus telephone interview: Results from a national survey. Medical Care, 32 (6), 551-567.
- McHorney, C. A., & Tarlov, A. R. (1995). Individual-patient monitoring in clinical practice: are available health status surveys adequate? Quality of Life Research, 4 (4), 293-307.
- Nelson, E. C., & Berwick, D. M. (1989). The measurement of health status in clinical practice. Medical Care, 27 (3), S77-90.
- Nerenz, D. R., Repasky, D. P., Whitehouse, F. W., & Kahkonen, D. M. (1992). Ongoing assessment of health status in patients with diabetes mellitus. Medical Care, 30 (5 Suppl), MS112-124.
- O'Boyle, C. A., McGee, H., Hickey, A., & O'Malley, K. (1992). Clinical practice: Individual quality of life in patients undergoing hip replacement. The Lancet, 339, 1088-1091.

Okamoto, L. J., Noonan, M., DeBoisblanc, B. P., & Kellerman, D. J. (1996).

Fluticasone propionate improves quality of life in patients with asthma requiring oral corticosteroids. Annals of Allergy, Asthma, & Immunology, 76, 455-461.

Osterhaus, J. T., Townsend, R. J., Gandek, B., & Ware, J. E. (1994). Measuring the functional status and well-being of patients with migraine headache. Headache, 34, 337-343.

Paterson, C. (1996). Measuring outcomes in primary care: a patient generated measure, MYMOP, compared with the SF-36 health survey. British Medical Journal, 312, 1016-1020.

Patrick, D. L., Deyo, R. A., Atlas, S. J., Singer, D. E., Chapin, A., & Keller, R. B. (1995). Assessing health-related quality of life in patients with sciatica. Spine, 20 (17), 1899-1908.

Philbrick, J. T., Connelly, J. E., & Wofford, A. B. (1996). The prevalence of mental disorders in rural office practice. Journal of General Internal Medicine, 11, 9-15.

Phillips, R. C., & Lansky, D. J. (1992). Outcomes management in heart valve replacement surgery: early experience. Journal of Heart Valve Disease, 1 (1), 42-50.

Reifler, D. R., Kessler, H. S., Bernhard, E. J., Leon, A. C., & Martin, G. J. (1996). Impact of screening for mental health concerns on health service utilization and functional status in primary care patients. Archives of Internal Medicine, 156, 2593-2599.

- Rector, T. S., Ormaza, S. M., & Kubo, S. H. (1993). Health status of heart transplant recipients versus patients awaiting heart transplantation: a preliminary evaluation of the SF-36 questionnaire. Journal of Heart and Lung Transplantation, 12 (6), 983-986.
- Rothwell, P. M., McDowell, Z., Wong, C. K., & Dorman, P. J. (1997). Doctors and patients don't agree: cross sectional study of patients' and doctors' perceptions and assessments of disability in multiple sclerosis. British Medical Journal, 314, 1580-1583.
- Rubenstein, L. V., Calkins, D. R., Young, R. T., Cleary, P. D., Fink, A., Kosecoff, J., Jette, A. M., Davies, A. R., Delbanco, T. L., & Brook, R. H. (1989). Academia and Clinic: Improving patient function: A randomized trial of functional disability screening. Annals of Internal Medicine, 111, 836-842.
- Ruta, D. A., Garratt, A. M., Wardlaw, D., & Russell, I. T. (1994). Developing a valid and reliable measure of health outcome for patients with low back pain. Spine, 19 (17), 1887-1896.
- Ruta, D. A., Garratt, A. M., Chadha, Y. C., Flett, G. M., Hall, M. H., & Russell, I. T. (1995). Assessment of patients with menorrhagia: how valid is a structured clinical history as a measure of health status? Quality of Life Research, 4 (1), 33-40.
- Sarafino, E. P. (1998). Health psychology. Biopsychosocial interactions. John Wiley & Sons, Inc.
- Schipper, H. (1983). Why measure quality of life? Canadian Medical Association Journal, 128, 1367-1369.

- Schor, E. L., Lerner, D. J., & Malspeis, S. (1995). Physicians' assessment of functional health status and well-being. Archives of Internal Medicine, 155, 309-314.
- Schroeder, S. A. (1987). Outcome assessment 70 years later. Are we ready. The New England Journal of Medicine, 316, 160-162.
- Slevin, M. L., Plant, H., Lynch, D., Drinkwater, J., & Gregory, W. M. (1988). Who should measure quality of life, the doctor or the patient? British Journal of Cancer, 57, 109-112.
- Solomon, G. D., Skobieranda, F. G., & Genzen, J. R. (1995). Quality of life assessment among migraine patients treated with sumatriptan. Headache, 35 (8), 449-454.
- Spertus, J. A., Winder, J. A., Dewhurst, T. A., Deyo, R. A., & Fihn, S. D. (1994). Monitoring the quality of life in patients with coronary artery disease. American Journal of Cardiology, 74 (12), 1240-1244.
- Stewart, A. L., Greenfield, S., Hays, R. D., Wells, K., Rogers, W. H., Berry, S. D., McGlynn, E. A., & Ware, J. E. (1989). Functional status and well-being of patients with chronic conditions. Results from the Medical Outcomes Study. Journal of the American Medical Association, 262 (7), 907-913.
- Street, R. L., Gold, W. R., & McDowell, T. (1994). Using health status surveys in medical consultations. Medical Care, 32 (7), 732-744.
- Tarlov, A. R. (1983). The increasing supply of physicians, the changing structure of the health-services system, and the future practice of medicine. New England Journal of Medicine, 308 (20), 1235-1244.

- Tarlov, A. R., Ware, J. E., Greenfield, S., Nelson, E. C., Perrin, E., & Zubkoff, M. (1989). The Medical Outcomes Study: An application of methods for monitoring the results of medical care. Journal of the American Medical Association, 262 (7), 925-30.
- Temple, P., Travis, B., Sachs, L., Strasser, S., Choban, P., & Flancbaum, L. (1995). Functioning and well-being of patients before and after elective surgical procedures. Journal of the American College of Surgeons, 181, 17-25.
- Thunedborg, K., Black, C. H., & Bech, P. (1995). Beyond the Hamilton Depression Scores in long-term treatment manic-melancholic patients: Prediction of recurrence of depression by quality of life measures. Psychotherapy and Psychosomatics, 64, 131-140.
- Tilly, K. F., Belton, A. B., McLachlan, J. F. C. (1995). Continuous monitoring of health status outcomes: experience with a diabetes education program. Diabetes Educator, 21 (5), 413-419.
- van Tulder, M. W., Aaronson, N. K., & Bruning, P. F. (1994). The quality of life of long-term survivors of Hodgkin's disease. Annals of Oncology, 5 (2), 153-158.
- Viramontes, J. L. & O'Brien, B. (1994). Relationship between symptoms and health-related quality of life in chronic lung disease. Journal of General Internal Medicine, 9 (1), 46-48.
- Vollmer, W. M., O'Hollaren, M., Ettinger, K. M., Stibolt, T., Wilkins, J., Buist, A. S., Linton, K. L., & Osborne, M. L. (1997). Speciality differences in the management of asthma. Archives of Internal Medicine, 157, 1201-1208.

- Wagner, A. K., Keller, S. D., Kosinski, M., Baker, G. A. Jacoby, A., Hsu, M. A. Chadwick, D. W., & Ware, J. E. (1995). Advances in methods for assessing the impact of epilepsy and antiepileptic drug therapy on patients' health-related quality of life. Quality of Life Research, 4 (2), 115-134.
- Wagner, P. J., Phillips, W., Radford, M., & Hornsby, J. L. (1995). Frequent use of medical services. Patient reports of intentions to seek care. Archives of Family Medicine, 4 (7), 594-599.
- Walker, V., Streiner, D. L., Novosel, S., Rocchi, A., Levine, M. A., & Dean, D. M. (1995). Health-related quality of life in patients with major depression who are treated with moclobemide. Journal of Clinical Psychopharmacology, 15 (4 Suppl. 2), 60S-67S.
- Ware, J. E., Brook, R. H., Rogers, W. H., Keeler, E. B., Davies, A. R., Sherbourne, C. D., Goldberg, G. A., Camp, P., & Newhouse, J. P. (1986). Health Insurance: comparison of health outcomes at a health maintenance organisation with those of fee-for-service care. The Lancet, 327, 1017-1022.
- Ware, J. E. (1987). Standards for validating health measures: definition and content. Journal of Chronic Disease, 40 (6), 473-80.
- Ware, J.E., & Sherbourne, C.D. (1992). The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. Medical Care, 30 (6), 473-483.
- Ware, J. E., Gandek, B., & The IQOLA Project Group. (1994). The SF-36 health survey: Development and use in mental health research and the IQOLA Project. International Journal of Mental Health, 23 (2), 49-73.

- Ware, J. E., Gandek, B., Keller, S. D., & The IQOLA Project Group. (1995). Evaluating instruments used cross-nationally: Methods from the IQOLA Project. In B. Spilker (Ed.), Quality of Life and Pharmacoeconomics in Clinical Trials, Second Edition (in press). New York: Raven Press.
- Ware, J.E. (1995). The status of health assessment 1994. Annual Review of Public Health, 16, 327-354.
- Ware, J. E., Kosinsky, M., & Keller, S. D. (1996). A 12-item short-form health survey. Construction of scales and preliminary tests of reliability and validity. Medical Care, 34 (3), 220-233.
- Weinberger, M., Samsa, G. P., Hanlon, J. T., Schmader, K., Doyle, M. E., Cowper, P. A., Uttech, K. M., Cohen, H. J., & Feussner, J. R. (1991). An evaluation of a brief health status measure in elderly veterans. Journal of the American Geriatric Society, 39, 691-694.
- Weinberger, M., Kirkman, M. S., Samsa, G. P., Cowper, P. A., Shortliffe, E. A., Simel, D. L., & Feussner, J. R. (1994). The relationship between glycemic control and health-related quality of life in patients with non-insulin-dependent diabetes mellitus. Medical Care, 32 (12), 1173-1181.
- Weinberger, M., Oddone, E. Z., Samsa, G. P., & Landsman, P. B. (1996). Are health-related quality-of-life measures affected by the mode of administration? Journal of Clinical Epidemiology, 49 (2), 135-140.

APPENDICES

Appendix A**The MOS SF-36 Health Survey****(English)**

SF- 36 HEALTH SURVEY

INSTRUCTIONS: This survey asks for your views about your health. This information will keep track of how you feel and how well you are able to do your usual activities.

Answer every question by marking the answer as indicated. If you are unsure about how to answer a question, please give the best answer you can.

1. In general, would you say your health is:

(circle one)

- | | |
|-----------------|---|
| Excellent | 1 |
| Very good | 2 |
| Good..... | 3 |
| Fair | 4 |
| Poor | 5 |

2. Compared to one year ago, how would you rate your health in general now?

(circle one)

- | | |
|--|---|
| Much better now than one year ago..... | 1 |
| Somewhat better now than one year ago..... | 2 |
| About the same as one year ago..... | 3 |
| Somewhat worse now than one year ago..... | 4 |
| Much worse now than one year ago | 5 |

3. The following items are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much?

(circle one number on each line)

<u>ACTIVITIES</u>	Yes, Limited A lot	Yes, Limited A Little	No, Not Limited At All
	a. Vigorous activities , such as running, lifting heavy objects, participating in strenuous sports	1	2
b. Moderate activities , such as moving a table, pushing a vacuum cleaner, bowling, or playing golf	1	2	3
c. Lifting or carrying groceries	1	2	3
d. Climbing several flights of stairs	1	2	3
e. Climbing one flight of stairs	1	2	3
f. Bending, kneeling, or stooping	1	2	3
g. Walking more than a mile	1	2	3
h. Walking several blocks	1	2	3
i. Walking one block	1	2	3
j. Bathing or dressing yourself	1	2	3

4. During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

(circle one number on each line)

	YES	NO
a. Cut down on the amount of time you spent on work or other activities	1	2
b. Accomplished less than you would like	1	2
c. Were limited in the kind of work or other activities	1	2
d. Had difficulty performing the work or other activities (for example, it took extra effort)	1	2

5. During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?

(circle one number on each line)

	YES	NO
a. Cut down on the amount of time you spent on work or other activities	1	2
b. Accomplished less than you would like	1	2
c. Didn't do work or other activities as carefully as usual	1	2

6. During the past 4 weeks, to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbors, or groups?

(circle one)

- Not at all1
- Slightly2
- Moderately3
- Quite a bit4
- Extremely5

7. How much bodily pain have you had during the past 4 weeks?

(circle one)

- None.....1
- Very mild.....2
- Mild3
- Moderate4
- Severe5
- Very severe6

8. During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?

(circle one)

- Not at all1
- A little bit2
- Moderately3
- Quite a bit4
- Extremely5

9. These questions are about how you feel and how things have been with you during the past 4 weeks. For each question, please give the one answer that comes closest to the way you have been feeling. How much of the time during the past 4 weeks -

(circle one number on each line)

	All of the time	Most of the time	A Good Bit of the Time	Some of the Time	A Little of the Time	None of the Time
a. Did you feel full of pep?	1	2	3	4	5	6
b. Have you been a very nervous person?	1	2	3	4	5	6
c. Have you felt so down in the dumps that nothing could cheer you up?	1	2	3	4	5	6
d. Have you felt calm and peaceful?	1	2	3	4	5	6
e. Did you have a lot of energy?	1	2	3	4	5	6
f. Have you felt downhearted and blue?	1	2	3	4	5	6
g. Did you feel worn out?	1	2	3	4	5	6
h. Have you been a happy person?	1	2	3	4	5	6
i. Did you feel tired?	1	2	3	4	5	6

10. During the past 4 weeks, how much of the time has your physical and emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?

(circle one)

- All of the time.....1
- Most of the time.....2
- Some of the time.....3
- A little of the time.....4
- None of the time.....5

11. How TRUE or FALSE is each of the following statements for you?

(circle one number on each line)

	Definitely True	Mostly True	Don't Know	Mostly False	Definitely False
a. I seem to get sick a little easier than other people	1	2	3	4	5
b. I am as healthy as anybody I know	1	2	3	4	5
c. I expect my health to get worse	1	2	3	4	5
d. My health is excellent	1	2	3	4	5

Appendix B**The MOS SF-36 Health Survey****(Turkish)**

SF- 36 SAĞLIK TARAMASI

YÖNERGE: Bu tarama formu size sağlığınıza ilgili görüşlerinizi sormaktadır. Bu bilgiler sizin nasıl hissettiğinizi ve herzamanki faaliyetlerinizi ne rahatlıkla yapabildiğinizi izlemekte yardımcı olacaktır.

Bütün soruları belirtildiği şekilde cevaplayın. Eğer bir soruyu ne şekilde cevaplayacağınızdan emin olamazsanız, lütfen en yakın cevabı işaretleyin.

1. Genel olarak sağlığınıza nasıl değerlendirirsiniz?

(birinin etrafına daire çizin)

Mükemmel	1
Çok iyi	2
İyi	3
Fena değil	4
Kötü	5

2. Geçen seneyle karşılaştırıldığında, şimdi sağlığınıza nasıl değerlendirirsiniz?

(birinin etrafına daire çizin)

Bir yıl önceye göre çok daha iyi	1
Bir yıl önceye göre daha iyi	2
Hemen hemen aynı	3
Bir yıl önceye göre daha kötü	4
Bir yıl önceye göre çok daha kötü	5

3. Aşağıdakiler normal olarak gün içerisinde yapıyor olabileceğiniz bazı faaliyetlerdir. Su sıralarda sağlığınız sizi bu faaliyetler bakımından kısıtlıyor mu? Kısıtlıyorsa ne kadar?

(Her satırda bir sayının etrafına daire çizin)

<u>FAALİYETLER</u>	Evet, oldukça kısıtlıyor	Evet, biraz kısıtlıyor	Hayır, hiç kısıtlamıyor
a. Kuvvet gerektiren faaliyetler , örneğin ağır eşyalar kaldırmak, futbol gibi sporlarla uğraşmak	1	2	3
b. Orta zorlukta faaliyetler , örneğin masa kaldırmak, süpürmek, yürüyüş gibi hafif spor yapmak	1	2	3
c. Çarşı-pazar torbalarını taşımak	1	2	3
d. Birkaç kat merdiven çıkmak	1	2	3
e. Bir kat merdiven çıkmak	1	2	3
f. Eğilmek, diz çökmek, yerden birşey almak	1	2	3
g. Bir kilometre'den fazla yürümek	1	2	3
h. Birkaç yüz metre yürümek	1	2	3
ı. Yüz metre yürümek	1	2	3
j. Yıkanmak ya da giyinmek	1	2	3

4. Geçtiğimiz bir ay (4 hafta) içerisinde işinizde veya diğer günlük faaliyetlerinizde bedensel sağlığınız nedeniyle aşağıdaki sorunların herhangi biriyle karşılaştınız mı?

(Her satırda bir sayının etrafına daire çizin)

	EVET	HAYIR
a. İş ya da iş dışı uğraşlarınıza verdiğiniz zamanı kısmak zorunda kalmak	1	2
b. Yapmak istediğinizden daha azını yapabilmek (bitmeyen projeler, temizlenmeyen ev gibi...)	1	2
c. Yapabildiğiniz iş türünde ya da diğer faaliyetlerde kısıtlanmak	1	2
d. İş ya da diğer uğraşları yapmaktan zorlanmak	1	2

5. Geçtiğimiz bir ay (4 hafta) içerisinde işinizde veya diğer günlük faaliyetlerinizde duygusal problemleriniz nedeniyle (üzüntülü ya da kaygılı olmak gibi) aşağıdaki sorunların herhangi biriyle karşılaştınız mı?

(Her satırda bir sayının etrafına daire çizin)

	EVET	HAYIR
a. İş ya da iş dışı uğraşlarınıza <u>verdiğiniz zamanı</u> kısmak zorunda kalmak	1	2
b. Yapmak istediğinizden <u>daha azını</u> yapabilmek (bitmeyen projeler, temizlenmeyen ev gibi...)	1	2
c. İş ya da diğer uğraşları her zaman gibi dikkatlice yapamamak	1	2

6. Son bir ay (4 hafta) içerisinde bedensel sağlığınız ya da duygusal problemleriniz, aileniz, arkadaşlarınız, komşularınızla ya da diğer gruplarla normal olarak yaptığınız sosyal faaliyetlere ne ölçüde engel oldu?

(birinin etrafına daire çizin)

- hiç1
 biraz2
 orta derecede3
 epeyce4
 çok fazla5

7. Geçtiğimiz bir ay (4 hafta) içerisinde ne kadar bedensel ağrılarınız oldu?

(birinin etrafına daire çizin)

- hiç1
 çok hafif2
 hafif3
 orta hafiflikte4
 aşırı derecede5
 çok aşırı derecede6

8. Son bir ay (4 hafta) içerisinde, **ağrı** normal işinize (ev dışında ve ev işi) ne kadar engel oldu?

(birinin etrafına daire çizin)

hiç olmadı 1
 biraz 2
 orta derecede 3
 epey 4
 çok fazla 5

9. Aşağıdaki sorular geçtiğimiz bir ay (4 hafta) içerisinde kendinizi nasıl hissettiğinizle ve işlerin sizin için nasıl gittiği ile ilgilidir. Lütfen, her soru için nasıl hissettiğinize en yakın olan cevabı verin. Geçtiğimiz 4 hafta içindeki sürenin ne kadarı :

(Her satırda bir sayının etrafına daire çizin)

	Her zaman	Çoğu zaman	Epeyce	Arada sırada	Çok ender	Hiçbir zaman
a. Kendinizi hayat dolu hissettiniz?	1	2	3	4	5	6
b. Çok sinirli bir kişi oldunuz?	1	2	3	4	5	6
c. Hiç bir şeyin sizi neşelendiremeyeceği kadar moraliniz bozuk ve kötü oldu?	1	2	3	4	5	6
d. Sakin ve huzurlu hissettiniz?	1	2	3	4	5	6
e. Çok enerjiniz oldu?	1	2	3	4	5	6
f. Mutsuz ve kederli oldunuz?	1	2	3	4	5	6
g. Kendinizi bitkin hissettiniz?	1	2	3	4	5	6
h. Mutlu ve sevinçli oldunuz?	1	2	3	4	5	6
i. Yorgun hissettiniz?	1	2	3	4	5	6

10. Geçtiğimiz bir ay (4 hafta) içerisinde, bu sürenin ne kadarında bedensel sağlığınız va da duygusal problemleriniz, sosyal faaliyetlerinize (arkadaş, akraba ziyareti gibi) engel oldu?

(birinin etrafına daire çizin)

Her zaman1
 Çoğu zaman2
 Bazen3
 Çok ender4
 Hiçbir zaman5

11. Aşağıdaki herbir ifade sizin için ne kadar DOĞRU ya da YANLIŞ?

(her satırda bir sayının etrafına daire çizin)

	Kesinlikle Doğru	Çoğunlukla Doğru	Bilmiyorum	Çok kere Yanlış	Kesinlikle Yanlış
a. Başkalarından biraz daha kolay hastalandığımı düşünüyorum	1	2	3	4	5
b. Ben de tanıdığım herkes kadar sağlıklıyım	1	2	3	4	5
c. Sağlığımın kötü gideceğini sanıyorum	1	2	3	4	5
d. Sağlığım mükemmeldir	1	2	3	4	5

Appendix C**Symptom Checklist**

TEŞHİS:

Tarih: Eğitim:
 Cinsiyet: Meslek:
 Yaş: Medeni durum:

Bir doktor tarafından, sizde HİÇ şu hastalıklardan biri olduğu söylendi mi?

- hipertansiyon
- yakın geçmişte miyokard enfarktüsü
- kalp yetmezliği
- şeker hastalığı
- angina (koroner kalp hastalığı)
- kanser (deri kanseri hariç)
- yakın zamanda bir psikiyatrik hastalık (ör: son bir yıl içinde depresyon)
- ve başka (ör: multipl skleroz, guatr gibi).

ŞU SIRALARDA (son bir ay içinde), aşağıdaki sorunlardan biri sizde var mı?

(Bu sorunlardan herhangi biri bu son bir ay içinde bir doktora başvurmanıza yol açmış veya doktora başvurmayı düşündürecek kadar yoğun şikayet konusu ise işaretleyiniz)

- artrit (eklem hastalığı) veya romatizma
- siyatik veya sırt veya bel ağrısı gibi sorunlar
- sık tekrarlayan ve/veya şiddetli başağrıları
- kronik akciğer hastalığı
- sindirim sistemi rahatsızlığı (ör: gastrit, ülser, enterit, kolit)
- dermatit veya başka kronik deri döküntüleri
- psikolojik sorun (ör: bir psikiyatr veya psikoloğa başvurmayı düşündürecek kadar yoğun üzüntü, çöküntü; yoğun kaygı, endişe, sıkıntı; aşırı alkol alma gibi)
- sık veya sürekli kullanılan ilaç (varsa ilaç adı)
- ve başka.