Severity Of Illness-Geriatric (SOI-G):

Instrument Development

A Thesis Submitted to the College of

Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy

in the Department of Psychology

University of Saskatchewan

Saskatoon

By

Lisa Dawn Berg-Kolody

Spring 2002

UNIVERSITY OF SASKATCHEWAN

College of Graduate Studies and Research

SUMMARY OF DISSERTATION

Submitted in partial fulfillment

of the requirements for the

DEGREE OF DOCTOR OF PHILOSOPHY

by

LISA DAWN BERG-KOLODY

Psychology Department

University of Saskatchewan

Spring 2002

Examining Committee:

Dr. A. Yackulic	Dean/Associate Dean/Dean's Designate, Chair, College of Graduate Studies and Research
Dr. L. McMullen	Chair of Advisory Committee Psychology Department
Dr. D. Scott	Advisor, Psychology Department
Dr. M. Brown	Retired
Dr. L. Shepel	Psychology Department
Dr. N. Stewart	College of Nursing

External Examiner:

Professor M. Pancer
Department of Psychology
Wilfred Laurier University
Waterloo, Ontario, N2L 3C5

Severity Of Illness-Geriatric (SOI-G): Instrument Development

Controlling for the wide variability in the physical health status of geriatric populations is important as severity of illness is known to both moderate and suppress relationships examined in psychosocial research. The purpose of the present investigation was to develop a uniform, easily administered quantitative index of illness severity composed of disease-specific scales that was independent of psychosocial factors and appropriate for use with a geriatric population. As well, the aim was to collect preliminary data on the reliability and validity of the scale. The development of the Severity of Illness-Geriatric (SOI-G) scale involved the adaptation of a previously developed severity of illness instrument Severity of Renal Disease Scale (SORDS).

The present investigation involved five programmatically linked studies. Study 1 involved the determination of the items to be included on SOI-G while Study 2 defined the severity criteria for each item. In Study 3, five geriatric specialists scaled each level of each item on the same underlying threat to life scale. There was a high level of initial agreement between the raters supporting the reliability of the severity values. The final scale consisted of 32 items.

In Study 4, archival data was collected on 61 patients admitted to the geriatric unit of a rehabilitation hospital. The SOI-G was compared to the Cumulative Illness Rating Scale-Geriatric (CIRS-G) and a global severity rating. SOI-G inter-rater reliability estimates were low (likely due to rater error) but promising. SOI-G demonstrated support for content validity, face validity, and construct validity but evidence for convergent validity was not established. SOI-G scores were sensitive to differences among patients with respect to discharge outcome. The utility of SOI-G as a moderator variable in psychosocial research with the elderly could not be explored in Study 5 due to a limited sample size.

It was concluded that the present investigation demonstrated the potential usefulness of SOI-G in psychosocial research with the elderly but further research is needed before definitive conclusions can be made. The SOI-G offers researchers a tool for controlling disease variability that is not measured by psychological tests but must be accounted for in research designs.

BIOGRAPHICAL

1990	Bachelor of Arts (High Honors), Psychology University of Saskatchewan
1995	Master of Arts, Clinical Psychology University of Saskatchewan

PUBLICATIONS

Berg, L.D. (1995). Establishing a tentative typology of families with sexually abused and nonabused adolescents. Master's of Arts Thesis, University of Saskatchewan, Canada

Fischer, D.G., Stevens, G., & Berg, L. (1992). Review and Monitoring of Child Sexual Abuse Cases in Selected Sites in Saskatchewan. Prepared for Justice Canada by Peat Marwick Stevenson & Kellogg, Management Consultants

PRESENTATIONS

Litke, K., Berg, L., Torrens, D., & Crossley, M. (1998). Schizophrenia: A generalized cognitive disorder? Poster presented at the 1998 CPA in Edmonton, AB, June, 1998.

Berg, L. & Hay, D. (1997). Family functioning and social skills of adolescents sexually abused as children. Poster presented at the Biennial Meeting of the Society for Research in Child Development, Washington, D.C., April 1997

Berg, L. (1995). Family functioning and sexual abuse. Paper presented at Changing the Climate: A Multi-Disciplinary Conference for Graduate Students, Saskatoon, March 1995

Hay, D.H., Berg, L., & Safnuk, T. (1995). The social development of adolescents sexually abused as children. Poster presented at the Biennial Meeting of the Society for Research in Child Development, Indianapolis, March 1995

Fischer, D., Berg, L., & Stevens, G. (1993). Child sexual abuse: Implementation of Bill C-15. Poster presented at the 1993 CPA in Montreal, QB April, 1993

Berg, L. (1990, April). Depressogenic attribution style as a function of perceived parental rejection and control. Honours thesis paper presented at the Psychology Undergraduate Research Conference, Saskatoon, SK

Berg, L. (1989, April). Parental pressure, self-esteem and adolescent reported deviance: Bending the twig too far. Paper presented at the Psychology Undergraduate Research Conference, Saskatoon, SK

Permission to Use

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that nay copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Psychology
University of Saskatchewan
Saskatoon, Saskatchewan

Abstract

Controlling for the wide variability in the physical health status of geriatric populations is important as severity of illness is known to both moderate and suppress relationships examined in psychosocial research. The purpose of the present investigation was to develop a uniform, easily administered quantitative index of illness severity, composed of disease-specific scales, that was independent of psychosocial factors and appropriate for use with a geriatric population. As well, the aim was to collect preliminary data on the reliability and validity of the scale. The development of the Severity of Illness-Geriatric (SOI-G) scale involved the adaptation of a previously developed severity of illness instrument Severity of Renal Disease Scale (SORDS).

The present investigation involved five programmatically linked studies. Study 1 involved the determination of the items to be included on SOI-G while Study 2 defined the severity criteria for each item. In Study 3, five geriatric specialists scaled each level of each item on the same underlying threat to life scale. There was a high level of initial agreement between the raters supporting the reliability of the severity values. The final scale consisted of 32 items.

In Study 4, archival data was collected on 61 patients admitted to the geriatric unit of a rehabilitation hospital. The SOI-G was compared to the Cumulative Illness Rating Scale-Geriatric (CIRS-G) and a global severity rating.

SOI-G inter-rater reliability estimates were low (likely due to rater error) but promising. SOI-G demonstrated support for content validity, face validity, and construct validity but evidence for convergent validity was not established. SOI-G scores were sensitive to differences among patients with respect to discharge outcome. The utility of SOI-G as a moderator variable in psychosocial research with the elderly could not be explored in Study 5 due to a limited sample size.

It was concluded that the present investigation demonstrated the potential usefulness of SOI-G in psychosocial research with the elderly but further research is needed before definitive conclusions can be made. The SOI-G offers researchers a tool for controlling disease variability that is not measured by psychological tests but must be accounted for in research designs.

Acknowledgments

I would like to thank Dr. Mark Pancer from Wilfrid Laurier University for agreeing to serve as my external examiner. I enjoyed the challenge that his thoughtful questions presented and appreciated his comments regarding my research. I would also like to acknowledge Drs. Brown, Shepel, and Stewart for their role as my advisory committee. Their guidance and feedback played an important role in the preparation of this document.

I wish to express my appreciation to my supervisor, Dr. David Scott, for sharing his knowledge and expertise with me. Through his example, Dr. Scott has taught me the importance of integrity, honesty, and respect in the supervisor-student relationship.

I am deeply indebted to Dr. Darryl Rolfson of the Glenrose Rehabilitation Hospital (GRH) in Edmonton who generously gave of his time in the development of SOI-G and who was instrumental in the data collection process. I appreciated his professionalism, guidance, and remarkable dedication to research. Dr. Rolfson's support played a major role in the successful completion of my dissertation project. I would also like to express my gratitude to the administration and other professionals at the GRH that participated in this project.

I would like to thank my husband Brad for his unwavering love and support and for never letting me lose site of the forest. Thank you to Christina and Colin for being the family I always hoped I would be lucky enough to have.

Thank you to the following family and friends who fed me, gave me shelter, and consulted with me regarding the many ethical, procedural, legal, and often 'meaning of life' issues that arose during the dissertation process: Sonja and Dale, Bernie and Carol, Colleen and Andrew, Diana, Natalie, Raelyn, and Wendy. Thank you to my supervisors (Drs. Hawkins, Cooper, and Harper) at Alberta Hospital Edmonton for their role in my personal and professional development. Lastly, I wish to acknowledge the support of my research team for their many years of support and advice.

This research is dedicated to my parents, the late Bernhardt (Ben) Malfred Berg and the late Doreen Lena Berg, for their role in helping me to achieve my dreams.

Table of Contents

List of Tables

1. Introduction

Our scientific knowledge of geriatric psychology has grown substantially in recent years (Katz, 1996; Thomas, Kelman, Kennedy, Ahn, & Yang, 1992). Despite the growing body of literature, the gerontological research to date is ambiguous about the association between physical disease and psychological variables among the elderly (Stuck et al., 1999). Given that physical disease is so common among the elderly and a major contributing factor to decline in social and psychological functioning (Borchelt, Gilberg, Horgas & Geiselmann,1999), it is clear that the ability to quantitatively measure physical disease is crucial in research with older persons. The high incidence of comorbidity of physical illnesses among the elderly was documented in the Berlin Aging Study which found evidence of at least one physical illness in up to 96 percent of old persons with an estimated 88 percent suffering from at least five physical illnesses and, of those, 30 percent experienced moderate to severe illnesses (Steinhagen-Thiessen & Borchelt, 1999).

The assessment of chronic health problems independent of psychological variables is problematic as existing health status measures often combine physical and psychological variables in their ratings. For instance, the Sickness Impact Profile (SIP) is a widely used, valid and reliable measure of functional status that De-Bruin, de-Witte, Stevens, and Diederiks (1992) suggest

should stand as a gold standard against which other measures might be compared. However, the SIP consists of three physical functioning categories, four psychosocial categories, and five independent categories that are not clearly physical or psychosocial (e.g., home management). Thus, to meet the need for a better method of assessing the severity of illness in geriatric research, the primary intent of this investigation was to develop a uniform, easily administered measure of chronic health problems which is independent of psychosocial factors.

The present research was a preliminary step towards developing a valid and reliable measure of geriatric severity of illness. The overall objectives of the current research were to develop a scale that: (1) was simple and easily applied; (2) could identify subgroups of elderly persons that were homogeneous with respect to severity of illness; (3) provided a severity rating that was not a function of institutional practices or operating norms; (4) that quantified an elderly patient's severity of illness on a single numerical scale with interval properties; and, (5) assessed health status of all elderly persons independent of cognitive status (i.e., can be used to rate elderly persons with dementia).

The development of the Severity of Illness-Geriatric (SOI-G) scale involved the adaptation of a previously developed severity of illness instrument, Severity of Renal Disease Scale (SORDS), designed to evaluate the physical health of renal patients (Baltzan et al., 1987, unpublished manuscript). The SOI-G provided a quantifiable index of disease severity for a specific group of patients, namely geriatric people. This research instrument was intended to

serve as a complement to existing measures rather than a replacement for other health status measures already in existence. It was hoped that the current investigation would lead to further research designed to refine the psychometric properties of the SOI-G and ultimately to add clarity to the relationship between psychological variables and health status in the elderly.

Development of the SOI-G involved the identification of patient dimensions that reflect important differences between less sick and more sick patients. Measurement of geriatric individuals on these dimensions avoided the use of specific therapies (e.g., number of days dependent on ventilator) and was based instead on the presence and extent of physiologic derangements. Overall, the development of SOI-G involved defining and conceptualizing health and physical illness, operationalizing and measuring the health dimensions of interest, determining the standards by which this measure was compared, and the collection of preliminary data on the reliability and validity of the instrument.

The development of SOI-G was important for two reasons. First, SOI-G is a generically applicable measure composed of disease specific scales that assesses current health status in a manner that can be concisely quantified. Second, SOI-G produces a global assessment of physical disease that is specific to the conditions and organ systems most likely to be affected in a geriatric population and can be used by other researchers.

2. Review Of The Literature

Before discussing the results of the present investigation, definitions of health, functional limitations, and the elderly will be offered. This will be followed by a selective review of psychiatric research involving the elderly. The following section reviews concepts and commonly used health status scales and examines their applicability with a geriatric population. Included in this section is an overview of the scale that was adapted for the present study (i.e., SORDS).

2.1 Definition Of Health

An important first step towards the development of a health status instrument is to establish a conceptual definition of health that can be successfully operationalized (Whitelaw & Liang, 1991). While there is no one agreed upon definition of health, many researchers agree that health is a multi-dimensional concept that can include medical, emotional, social, familial, educational, economic, religious, moral, and spiritual dimensions (Feinstein, 1992). Consistent with this, the definition given by the World Health Organization (WHO) suggests that health is "a state of complete physical, mental, and social well-being, and not merely the absence of disease" (Chenier, 1993, p. 2) and Borchelt et al. (1999) consider morbidity in old age to be the consequence of a complex interaction among biological, psychological, and social influences. At the level of the individual, one may choose to define health

from a multidimensional perspective, considering all of the components mentioned in the WHO definition. Researchers, however, acknowledge that such a comprehensive approach to health is too vague or extensive for measurement purposes and thus are likely to focus on a more narrow definition (Feinstein, 1992; Hall, Epstein, & McNeil, 1989).

Dworkin and Wilson (1993) present two approaches to the definition of health status. They refer to the focus on the complex interaction of biological, psychological, and social variables as the biopsychosocial model while the focus on pathobiology is referred to as the biomedical model. Dworkin and Wilson argue that the distinction between these two models lies in differentiating between disease and illness. According to Dworkin and Wilson, disease is defined as "a biological event representing a disruption of a body structure or organ system as a result of anatomic and/or physiological change" (p. 330). In comparison, illness is considered to be a more subjective experience encompassing "physical discomfort, emotional perturbation, behavioural limitations, and psychosocial disruption of activities and relationships (p. 330)". While the merits of the biomedical model (e.g., identification of specific causative agents for certain diseases) are acknowledged by researchers, there is growing recognition among researchers of the necessity of examining the complex, interactive nature of multiple factors in the disease process. For instance, Clark, Nash, Cohen, Chase, and Niaura (1998) point out that the biomedical model has failed to identify one causal agent of cardiovascular

disease. Instead, diet, stress, smoking, depression, hostility, exercise, activity level, and alcohol abuse are all identified risk factors for this disease.

One shortcoming in geriatric health status research identified by Whitelaw and Liang (1991) is the wealth of definitions of health offered by researchers without appropriate operationalizations. These authors suggest that if the issue of health and the elderly is to evolve, research is needed that explores the theoretical and measurement properties of the individual dimensions of health included within the WHO's broad definition of health. They further maintain that this is a necessary condition before researchers can begin to understand the complex interrelationships between the dimensions.

In line with Whitelaw and Liang's (1991) suggestion that researchers focus on the individual components of the WHO's definition of health, the focus of the present research was limited to the conceptual and measurement issues associated with the physical disease dimension of the WHO definition. It is acknowledged that clinicians from various health related fields should be guided by a more encompassing conceptual framework in their everyday work with individual geriatric patients (Engel, 1997). However, the goal of the present research project was not to develop a diagnostic tool but rather to develop a research instrument that could be used in the collection of data in geriatric research. To that end, physical illness was defined as the degree to which a disease posed a threat to life. This was consistent with previous research that has accepted the definition of severity of illness as the possible threat to life and equates severity with increased risk of mortality (Aronow, 1988).

2.2 Defining Functional Limitation

Another aspect of defining health is to clarify the pathway from disease to disability. Verbrugge and Jette (1994) have developed a sociomedical conceptual scheme that describes the pathway from pathology associated with chronic and acute conditions to functional outcomes. Rather than the categorical approach of disease versus illness suggested by Dworkin and Wilson (1993), this model focused on a causal pathway. This model of disability, called the Disablement Process, outlines risk factors, interventions, and exacerbators that speed or slow disablement. Verbrugge and Jette's model is based on two main conceptual models, the International Classification of Impairments, Disabilities, and Handicaps (ICIDH) scheme (developed by WHO) and an alternative model proposed by socialist Saad Nagi.

According to the Disablement model, the main pathway begins with pathology. This involves abnormal biochemical and physiological changes that are not always directly measurable by standard medical tests. Pathology leads to impairments which involve significant dysfunction in specific body systems due to accident or injury and is often revealed by clinical examination, laboratory tests, and symptom reports. Impairments contribute to functional limitations (i.e., restrictions in basic physical and mental actions). Lastly, functional limitations lead to disability or difficulty completing activities of daily life (ADLs). In this model, disability is heavily influenced by societal factors. For example, a person confined to a wheelchair because of a spinal cord injury (i.e., impairment) is unable to walk (i.e., functionally limited) but the extent of their

disability is partly determined by the accessability of compensatory mechanisms (e.g., wheel chair accessible buildings or modified kitchens in their home). The causal sequence outlined in the disablement process has been empirically validated (Femia, Zarit, & Johansson, 2001). The relevance of this model for the development of SOI-G will be outlined later in the results/method section.

2.3 Who Are "The Elderly"?

It is important to clarify who are being referred to in "geriatric" research. A variety of synonymous terms are used interchangeably in gerontologic studies when referring to the participants. Just a few of these terms include the elderly, the old aged, older adults, and geriatric sample. The question becomes how these various interchangeable terms are operationalized. The World Assembly on Aging (assembled by the United Nations in 1982) established age 60 as the onset of old age while in Canada the age cutoff used for the collection of statistical data has been set at age 65 (Chenier, 1993).

It is important to recognize that the elderly are not a homogenous group. For example, Haynie, Berg, Johansson, Gatz, and Zarit (2001) suggest that the presence of symptoms in depression may represent a qualitatively different experience for society's oldest old relative to younger elderly persons. Acceptance of the heterogeneity of the elderly has led some researchers to further differentiate the elderly into subcategories, such as "young old" or "old old". However, the operationalization of these subcategories is not standardized and can vary between studies. For instance, Chenier (1993) refers to individuals between 65 and 75 years as the "young old", those between

75 and 85 as the "middle-aged old", and those over 85 as the "old-old". In comparison, Girzadas, Counte, Glandon, and Tancredi (1993) used the terms "young old" (60-69 years), the "old" (70-79 years), and the "old old" (80+ years). Finkel (1996) defined the "old-old" as those over 70 years of age while Haynie and colleagues defined the "oldest old" as those 80 and over.

These terms were not used in the present research. A definition of age was necessary to facilitate the development of severity criteria and to aid in the scaling of these criteria. For the purposes of the present study, data was collected from individuals 60 years and older. However, as will be explained in greater detail later in this paper, a major portion of the development of SOI-G focused on the 70 to 75 year old range.

2.4 Psychiatric Research With Elderly Persons

The primary interest of the present study was to develop a measure of medical illness that would add clarity to the literature examining the relevance of various psychological factors in elderly samples. Furthering our understanding of the importance of psychosocial variables on health outcomes of older medical patients is crucial given the present focus on medical outcome assessments and the ever increasing healthcare needs of individuals in the later years of life. The following section was not intended to be an exhaustive review of psychiatric research with elderly persons but rather a selective review designed to show examples of how psychiatric research involving the elderly would be enhanced by a more reliable and valid assessment of severity of illness.

The elderly experience the same spectrum of mental health difficulties seen in younger age groups, including depression, anxiety, alcohol and substance abuse, and schizophrenia (Knight, Santos, Teri, & Lawton, 1998). Knight et al. estimate that 15 to 25 percent of community dwelling elderly persons experience serious symptoms associated with a psychiatric disorder. Estimates for older persons in medical and institutional settings are considerable higher (70 to 90 percent of patients; Knight et al., 1998). Depression and cognitive impairment are the two most common reasons for psychiatric consultations in a medical setting (Arfken, Lichtenberg & Tancer, 1999; Borchelt et al., 1999). In addition, the relationship between reduced survival rates and certain psychiatric illness among elderly persons has been well documented (Bartels, Forester, Miles, & Joyce, 2000; Burns, Lewis, Jacoby, & Levy, 1991; Bruce & Leaf, 1989; Davidson, Dewey, & Copeland, 1988; Kay & Bergmann, 1966; Wood, Evenson, Cho, & Wagan, 1985). For example, Bruce and Leaf (1989) found that adults over the age of 55 years with a diagnosis of an affective disorder, schizophrenia, or cognitive impairment had at least 150 per cent greater chance of dying.

Many professionals recognize that the cognitive and emotional difficulties associated with physical illnesses pose a unique challenge for persons working with older persons (Arfken et al., 1999; Hall et al., 1989; Knight et al., 1998; Rossberg-Gempton & Poole, 1999) and that the boundary between 'medical' problems and 'social' problems is often not clearly demarcated (Rockwood, Hogan, & MacKnight, 2000). For instance, Arnetz (1996) and Clark et al. (1998)

have identified psychosocial determinants as both precipitating and propagating variables for a number of illnesses (e.g., cardiovascular disease). Also, the reciprocal role between depression and self-reported global health ratings has been demonstrated by both Johnson, Stallones, Garrity, and Marx (1990) and Oslin, Streim, Katz, Edell, and TenHave (2000). Other research has implicated comorbid physical illness as a contributing factor to late-onset anxiety (excluding phobias; Sadavoy & LeClair, 1997) and late-onset mania (Bartels et al., 2000).

As reported by Conwell, Forbes, Cox, and Caine (1993), scores derived from existing health status instruments have been used as both predictor variables (e.g., in longitudinal studies of affective illness in the aged) and outcome variables (e.g., in studies of the effects of social factors or health behaviour on overall physical wellness) in psychosocial research. In addition, physical health scores have served as control variables in studies assessing the effectiveness of clinical interventions (Conwell et al., 1993). For instance, accurately understanding differences in survival requires that patients assigned to different treatment interventions be equated on severity of illness prior to treatment (Pompei, Charlson, & Douglas, 1988). Further, the evaluation of treatment outcomes requires an evaluation of the contribution of severity of illness to the onset, persistence, and abatement of psychological symptoms, such as depression or anxiety (Miller et al., 1996). Researchers who fail to control for illness severity can not be sure of the extent to which differences in outcome, assumed due to the treatment intervention, are confounded with pre-existing differences in severity of illness. However, past efforts at determining

the relationship between health status and psychological variables have been frustrated by the lack of a reliable and valid measure that assesses severity of illness in the elderly (Parmelee, Thuras, Katz, & Lawton, 1995).

To illustrate the importance of adequately assessing severity of illness in psychological research with the elderly, the following section will overview three psychological adjustment variables that are commonly included as outcome variables in gerontological investigations. These areas include quality of life, dementia, and depression.

2.4.1 Quality of Life

Quality of life is an important issue among the elderly as there are indications that positive evaluations of quality of life may help the elderly to cope with the many changes and challenges that present themselves in the later years of life and thereby reduce vulnerability to stress and disease. There are a variety of measures that purport to assess health-related quality of life; however, the term is rarely defined (McDowell & Newell, 1996).

Some researchers maintain that the construct 'quality of life' has no single definition and its inherently subjective nature means it is expected to vary as a function of culture, situation, demographics, and time (Rabins, 2000). One example of a quality of life definition was that given by Crist (1999) who defined this construct as "the degree of gratification perceived from one's contextual experience, including composite satisfaction with physical, emotional, social and spiritual environmental conditions" (p. 102). Rabins suggested that measuring quality of life is useful in a variety of ways such as: (1) determining whether the

treatment benefits perceived by researchers are also perceived as benefits by patients; (2) allowing for comparisons of morbidity across different diseases; (3) comparisons of treatment outcomes; and (4) providing a method for assessing the risk/benefit ratios associated with a particular treatment.

Research focusing on quality of life for elderly persons was motivated by the increase in life expectancy among the elderly (McDowell & Newell, 1996). In Canada, average life expectancy at birth has increased from 59 years in 1920 to 78 years in 1993 (Statistics Canada, 1998a). Today 70 percent of the population are expected to reach the age of 65 years (Furner, Brody, & Jankowski, 1997). In contrast, only 25 percent of the population reached this age in 1900 (Furner et al., 1997). The proportion of Canada's population consisting of people 65 years and over has increased from 5 percent in 1921 to 10 percent in 1981 (Statistics Canada, 1998b). In 1995 there were approximately 3.6 million people 65 years and over in Canada, representing 12 percent of the total population (Statistics Canada, 1998b). It is estimated that by 2041 the proportion of the population over the age of 65 will rise to at least 23 percent of the total population (Statistics Canada, 1998b).

With the increasing number of persons who are elderly, gerontological researchers became increasingly aware that quality of life as one ages involved more than merely surviving and that optimal quality of life was largely determined by successful adaptation to the changes associated with growing older (Crist, 1999; McDowell & Newell, 1996). Older adults face many challenges and changes in the later years of life, among them the death of

family and friends, change of residence and familiar surroundings when institutionalized, deteriorating health, and physical disabilities that limit mobility. Together, these factors often act to isolate seniors placing them at risk of feelings of loneliness. Health variables that the National Advisory Council on Aging (1991) identified as risk factors for loneliness include hearing problems, physical limitations and negative evaluations of one's health.

Different types of housing is one factor thought to relate to quality of life in the elderly. Crist (1999) conducted a pilot study to examine the influence of different types of housing (i.e., personal dwellings, specialized housing, and nursing homes) on quality of life preferences. With a total sample of 87 participants, Crist determined that, contrary to expectations, there were no significant differences found between the three housing types with respect to self-rated quality of life or with respect to the importance of quality of life variables (e.g., physical well being).

With community-dwelling individuals (age 62 and older), Girzadas et al. (1993) reported that functional health status was a stronger predictor of variability in life satisfaction than physical health status. On the basis of this finding, the authors concluded that it is not the absence or presence of disease that impacts life satisfaction, but how that disease or health condition influences functional ability. However, this conclusion can be challenged on a number of grounds. First, the method used to assess presence/absence of disease involved asked participants to indicate if they had experienced any of sixteen health problems common to older adults within the past 6 months. These items

were aggregated and higher scores were taken to represent poorer health status. However, the authors did not identify the sixteen health problems in the article. Thus, it is possible that the items were not equivalent in terms of threat to life. Combining the scores by addition implies (perhaps incorrectly) that a mild illness plus a moderate illness is equivalent to two mild illnesses or is greater in severity than one severe illness. Second, the average score for the participants was 1.4 (SD of 1.58) out of a possible 8 suggesting that, on average, this particular sample presented with less than two health problems. Lastly, it is not known how severely ill the sample was as only the presence/absence was recorded. Thus, the authors' conclusion that physical disease is not related to life satisfaction may have been premature.

Perceived quality of life is an important contributing factor to the onset and maintenance of depressed mood. In a longitudinal study, Haynie et al. (2001) examined the role of well-being in depression in persons over the age of 80 years. From a sample of 549 pairs of like-sex Swedish twins (aged 80 and older), Haynie and colleagues selected a subsample of one randomly selected individual from each pair yielding a sample of 275 persons. The baseline measures of depression for these individuals was compared with two subsequent measurement waves that were two years apart. The authors concluded that lack of well-being was a bigger contributor to depression scores than was self-reported sadness and tearfulness.

Using data from the first cross-sectional wave of the Berlin Aging Study (516 participants), Geiselmann and Bauer (2000) suggested that subthreshold

levels of depression (i.e., depressive syndromes that do not meet the criteria for a diagnosis of major depression) in the elderly may reflect diminished life satisfaction and deteriorated subjective evaluation of health. Relative to major depression, they described subthreshold depression in the elderly as having fewer symptoms with less continuity (e.g., durations often less than two weeks), fewer suicidal thoughts or feelings of guilt or worthlessness, but with frequent worries about health and weariness of living. Unfortunately, physical comorbidity was not assessed and thus the extent to which physical illness may have contributed to their mood could not be determined.

One complicating factor in the assessment of quality of life in older persons involves the presence of a dementia syndrome. Dementia interferes with the person's ability to think abstractly, to remember salient events, to make comparisons across multiple domains (e.g., physical well being, personal development and fulfillment, recreation, relationships), to have insight into their condition as well as understand the impact of that condition on their well being, and to communicate their thoughts on this issue (Rabins, 2000). Using a multi-step process with different panels of experts, Rabins developed an instrument to assess quality of life in patients with dementia. The Alzheimer Disease Health-Related Quality of Life Scale (ADRQL) is a 47-item, proxy-rated, behaviourally based instrument that assesses the following five domains: (1) social interaction; (2) awareness of self; (3) response to surroundings; (4) enjoyment of activities; and, (5) feelings/mood. Rabins indicated that preliminary validity and reliability studies have demonstrated modest

correlations between the ADRQL and instruments assessing cognitive impairment, low mood, behaviour disorders, and morbidity. The author acknowledged that, although the scale does not assess severity of illness, there was a low correlation with disease severity.

Abrams, Alexopoulos, Spielman, Klausner, and Kakuma (2001) examined quality of life in elderly psychiatric patients (40 inpatients and outpatients). The authors determined that the presence of Cluster B personality disorders (i.e., antisocial, borderline, histrionic, or narcissistic personality disorders) in elderly patients was found to relate directly to declines in global functioning and perceived quality of life. Further, the presence of a personality disorder was also found to increase or worsen the impact of depressive symptoms on long-term functioning and quality of life. Although the Cluster B personality disorders occur with a low frequency in elderly patients, the authors suggested that the personality traits may exert an influence even at sub-clinical levels. The influence of these traits on quality of life was hypothesized to occur because these traits may make the person more reactive, and more likely to alienate or reject others which is expected to reduce social support. Although severity of illness, as measured by Cumulative Illness Rating Scale (CIRS-G), was not found to be related to quality of life, the level of co-morbid illness in this sample is unknown as the severity of illness scores were not reported. The failure to find a relationship between severity of illness and quality of life may also reflects two criticisms of the CIRS-G, namely the summing of ordinal data and the possible confounding of physical and psychological variables.

2.4.2 Dementia

Dementia is a common disorder among the elderly (Baldwin & Jolley, 1986; Martin, Miller, Kapoor, Arena, & Boller, 1987). Patients with a dementia syndrome account for many of the institutionalized elderly and require greater amounts of patient care (Caputo et al., 1998). It has been estimated that dementia is a major debilitating condition for more than half of the elderly living in nursing homes (Martin et al., 1987).

Individuals with dementia suffer not only cognitive changes but emotional changes as well (Bozzola, Gorelick, & Freels, 1992; Magai, Cohen, Gomberg, Malatesta, & Culver, 1996). Historically, research has focused more on the cognitive changes associated with dementia but increasingly researchers are beginning to examine the emotional aspects of this disease. This is in keeping with reports from caregivers who identify changes in emotional reactively (e.g., verbal expressions of anger) as a major factor in caregiver burden (Magai et al., 1996). With a sample of 80 patients diagnosed with dementia, Bozzola et al. (1992) reported that 61 percent experienced diminished initiative/growing apathy, 39 percent demonstrated diminished regard for the feelings of others, 36 percent coarsening of affect, 34 percent impairment in emotional responsiveness, and 19 percent impairment of emotional control.

The importance of addressing the emotional components of this dementia is reflected in the seven stages of the Global Deterioration Scale (GDS) developed by Reisberg and colleagues (1982) and used in the present research. Reisberg et al. identify anxiety as the more predominant emotion in

the early stages of the disease whereas a decrease in affectivity appears in Stage 4 and 5, overt agitation at Stage 6, and nonverbal agitation in Stage 7.

Magai et al. (1996) examined emotional expression with 198 nursing home residents diagnosed with Alzheimer's disease and found that emotional expressivity did not necessarily follow cognitive-functional declines suggested by Reisberg et al. (1982). Emotional expression was coded during interactions with family members. Patients, including those with moderate to advanced stages of Alzheimer's Disease (AD), were found to express the basic human emotions of interest, happiness, sadness, fear, and anger. The only noted decline was in the appearance of joy during the end stage of the disease. The patients were found to be more emotionally expressive during interactions with family members than during interactions with nursing home staff, likely reflecting the qualitatively different nature of the relationships. This study challenges the belief that emotion is blunted in late-stage dementia and therefore that the treatment of depression for late stage dementia patients is unnecessary.

In a review of several studies relating dementia to increased rates of mortality, Langley (1995) reported that most studies of dementia show higher rates of death among those with dementia (including both community and psychiatric samples) than among those without dementia. However, Langley also noted that there was considerable variation in the rates (from zero to four times greater) across the reviewed studies.

In a study of community dwelling elderly (85 years or older), Fichter, Meller, Schroppel, and Steinkirchner (1995) reported a high incidence of

morbidity with only 3.4 percent of those with dementia having no concomitant physical illness. Almost half of the sample had four or more co-morbid illnesses (the most frequent physical illnesses included diseases of the circulatory system, diseases of the nervous system and sense organs, musculoskeletal diseases, diseases of the digestive and genitourinary system, and endocrine, nutritional and metabolic disorders). There was also a high incidence of comorbid psychiatric disorders including depression, anxiety and paranoid states. Other research has reported an association between the presence of delusions in AD patients and worse general health (Bassiony et al., 2000).

Martin et al. (1987) compared 202 patients with dementia (AD and multi-infarct dementia) with 202 nondemented controls matched by age and sex. At a three-year follow-up, Martin and colleagues reported a significantly lower survival rate among patients with dementia than among the controls (70 percent versus 84 percent, respectively). They found no significant differences in survival rates for the two types of dementia; however, there was a trend for individuals with multi-infarct dementia to have poorer survival rates than those participants with Alzheimer type dementia.

On the basis of previous research that had suggested that the differences in mortality rates between those with dementia and nondemented controls might be explained by differences in physical illness, Martin et al. (1987) also examined the health status of their two groups (demented versus nondemented) using the Older Americans Resource Survey (OARS). Both groups of patients were divided into two groups, those with mild physical

impairment (1 to 3 on OARS) and those with severe physical impairment (4 or greater on OARS). The rate of survival was lower for patients with dementia relative to the control group for those with mild physical impairment. However, there was no significant difference between the group with dementia and the control group when the level of physical impairment was in the severe range. Thus, studies reporting no differences in mortality rates between demented and nondemented participants may have had samples with more serious illnesses than studies reporting a difference in mortality rates.

The findings of the Martin et al. (1987) study suggest that the observed variability in mortality rates reported in the literature may be accounted for by varying degrees of severity of illness within the different samples. Consistent with this interpretation, Burns et al. (1991) found the presence of physical illness to be one factor associated with reduced survival rates among elderly psychiatric patients diagnosed with Alzheimer's disease. In sum, this research highlights the need to account for level of severity of illness when investigating the relationship between dementia and mortality rates.

2.4.3 Depression

Depression is the most common psychiatric disorder in elderly populations followed closely by dementia syndromes (Baldwin & Jolley, 1986; Burvill et al., 1991). Fredman et al (1989) and Katz (1996) explain that depression in the elderly can arise from social (e.g., bereavement, increased isolation, lower economic well-being) or biological factors (e.g., more chronic diseases, physiological effects of an illness).

Depressed elderly persons are also more likely than their younger counterparts to experience physical health difficulties and their depression is further exacerbated when mobility or independence is impaired by physical illness (Curyto, Chapleski, & Lichtenber, 1999; Melding, 1995). Research suggests that certain ethnic groups are at increased risk for depression because of higher rates of physical health problems relative to the general population. For example, Curyto and colleagues point out that Native American elderly are more likely to suffer from chronic diseases such as diabetes, liver and kidney disease, high blood pressure, emphysema and gall bladder difficulties than the general elderly population.

Geriatric depression can lead to functional and cognitive impairments, an increase in symptoms associated with medical illness, physiological decline, greater use of health care services, and increased risk of malnutrition (Katz, 1996). Research has shown poorer prognosis of depression in elderly with chronic physical health problems (Burvill, Mowry, & Hall, 1990). Depression has also been associated with higher rates of morbidity and mortality among the elderly (Schneider & Olin, 1995) and is strongly associated with completed suicide among older adults (Oslin et al., 2000). Suicide is one of the top 10 causes of death among community dwelling elderly and considered by some to be indicative of severe depression (Knight et al., 1998).

One topic in gerontological research that has been studied extensively is the association between depression in old age and an increased risk of mortality that is not accounted for by suicides (Copeland et al., 1992; O'Connor &

Vallerand, 1998; Rozzini, Frisoni, Bianchetti, Zanetti, & Trabucchi, 1991). The presence of depression, however, does not necessarily entail poorer outcome as other research has either failed to confirm the increased mortality rates in depressed elderly persons (Baldwin & Jolley, 1986; Burvill, Hall, Stampfer, & Emmerson, 1991; Fredman et al., 1989; Thomas et al., 1992) or has attributed the increase in mortality rates to physical disease (Kay & Bergmann, 1966; Murphy, Smith, Lindesay, & Slattery, 1988). Although health problems may influence factors related to mortality in geriatric research, this variable is often excluded or assessed inaccurately. Thus the discrepancies observed in the literature may, in part, be due to methodological difficulties, such as failing to adequately control for physical disease (Cohen-Cole & Kaufman, 1993; Schneider & Olin, 1995; Thomas et al., 1992).

Copeland et al. (1992) examined the association between depression and increased mortality rate in a geriatric population. Copeland and colleagues conducted a three-year longitudinal study with a community sample of 701 participants (aged 65 and over). There was a statistically significant increase in the rate of death among the depressed elderly relative to their non-depressed elderly counterparts. However, the extent to which depression directly contributed to mortality in Copeland's study can not be determined as the authors did not measure the degree of physical illness present in the sample; therefore, it is possible that depression was secondary to a medical condition that contributed to the participants' death.

In a longitudinal study involving Swedish twins (aged 80 or older), Haynie et al. (2001) reported low rates of depressive symptoms in their sample suggesting that the relationship between age and depression is not necessarily a linear one. Individuals who participated in all three waves of the study were significantly less depressed than the participants who were not included in the subsequent waves due to frailty or subsequent death. Thus, the less depressed participants were relatively healthy (based on self-reported subjective health) and higher functioning, implicating poor physical health and functional limitations in depression. The authors concluded that individuals that survive into old age may possess certain physical and psychiatric advantages that enable them to cope effectively with life challenges thus protecting them from depression. This conclusion is consistent with previous research involving late-onset anxiety disorders that suggested that good health might act as a stress buffer that inhibits panic in some individuals until such time as the buffer is jeopardized (Hassen & Pollard, 1994) and with research that has found a high rate of underlying medical illness among elderly persons diagnosed with late-onset mania (Bartels et al., 2000).

A longitudinal (four-year) study by Kay and Bergmann (1966) examined the possible relationship between mortality and mental illness with a sample of 98 elderly (age 65 or older) community participants. Kay and Bergmann found a significantly higher mortality rate among the psychiatric participants (mainly depression and anxiety) relative to the non-psychiatric participants. However, they also observed that participants with depression obtained higher ratings for

physical disability relative to the non-depressed participants. The authors concluded that the higher death rate among the psychiatric group was due to the fact that relative to the non-psychiatric group, a greater proportion of the psychiatric elderly patients had very poor physical health. However, their rating of physical disability was not based on a physical examination but rather based on what the "interviewer gleaned ... from observation (of paralyses, tremors, deformities, disorders of gait, cyanosis, breathlessness and so on)" (p. 4). Although these results should be viewed cautiously given the crudeness of their assessment of physical disease, the fact that the impact of physical disease was evident even when disease was assessed with a crude measure suggests that the variability of illness states is an extremely powerful factor.

Consistent with Kay and Bergmann (1966), Rozzini, Bianchetti, Franzoni, Zanetti, and Trabucchi (1991) found that level of depression at baseline, functional health, and somatic health status were all related to mortality with 1,201 community living elderly (70-75 years old) studied over a three-year period. Similarly, with a sample of nonclinical nursing home residents, O'Connor and Vallerand (1998) observed a relationship between depression and mortality that remained even after statistically controlling for physical health. However, their rating of health was based on a 7-point self-report scale (very poor to very good) and was not based on actual severity of illness ratings. With a sample of 124 depressed elderly patients, Murphy et al. (1988) also found a higher mortality rate for the depressed group than for a non-depressed control group (34 percent versus 14 percent, respectively). They determined that the

depressed group was significantly more ill than the non-depressed group. To determine whether or not the higher rate of physical illness explained the raised mortality of the depressed group, the authors created sub-groups consisting of depressed and non-depressed participants who had both a severe physical health event and a major chronic health difficulty. There were no significant differences between these subgroups. However, this particular method reduced their sample size to six depressed and two non-depressed participants thus calling into question the generalizability of these findings.

One hypothesis offered for the observed association between increased risk of mortality and depression is that depression influences death through behavioural factors, such as self neglect or inadequate compliance with medical treatment (Newhouse, 1996). Another explanation is that depression contributes to the higher mortality rate through biological factors, such as consecutive deterioration of immune functioning (Thomas et al., 1992). A third explanation is that illness causes depression that subsequently leads to death (Haynie et al., 2001; Thomas et al., 1992). The influence of depression on mortality may also occur via indirect pathways such as restricting performance of physical and mental actions used in daily life (Haynie et al., 2001). However, these explanations have not been adequately tested (Thomas et al., 1992) and other studies have failed to replicate the relationship between depression and increased risk of mortality among the elderly (Baldwin & Jolley, 1986; Burvill et al., 1991; Fredman et al., 1989; Thomas et al., 1992). In addition, while research has examined the role between depression and increased mortality in

community dwelling elderly, little research has examined the impact of depression for elderly persons in medical settings (Arfken et al., 1999).

With a sample of 667 patients aged 60 or older admitted to hospital, Arfken et al. (1999) determined that both depression (as determined by the Geriatric Depression Scale) and more severe cognitive impairment (assessed using the Mattis Dementia Rating Scale) independently contributed to increased risk of mortality, separate from the effects of age, medical illnesses, or disabilities. For more cognitively intact patients, moderate depression was a risk factor for death.

Burvill et al. (1991) examined the prognosis of depression in old age and the possible influence of chronic physical illness using both a self-report and other-report measure of physical illness. The authors examined the 12-month outcome of a cohort of 103 elderly persons being treated for depression (all but 5 were in-patients). In contrast to Murphy et al. (1988) and Kay and Bergmann (1966), these authors used a more reliable and valid measure of physical illness, the Older Americans Resources and Services (OARS) Physical Health Rating. They also assessed health using the patients' self-assessment of their health, and a severity measure devised by the authors. They found no significant difference in the level of chronic illness between participants who recovered from depression and those who remained depressed.

Consistent with Burvill et al. (1991), Baldwin and Jolley (1986) also found no relationship between mortality and physical health in depressed elderly. Baldwin and Jolley followed 100 depressed elderly patients over a period of

three to eight years. Physical health scores were obtained for six body systems (cardiovascular, central nervous system, musculoskeletal, respiratory, genitourinary, and gastrointestinal). Physical health for each of the six systems was scored as no problem, an inactive problem, or an active problem. Neither inactive problems nor active problems in any individual body system (taken separately) related in a significant manner to mortality. However, although the presence of illness was recorded, the severity of illness was not accounted for in the assessment of health.

Fredman et al. (1989) also reported a lack of association between depression and mortality risk (two-year follow-up) among 1,622 community adults, aged 60 years and older. However, there were a number of serious methodological problems associated with their assessment of severity of illness that threatens the validity of their conclusions. First, health status was assessed with a one-item self-report measure that assessed presence of at least one chronic disease but not the severity of illness. Second, given that many of the symptoms of depression assessed in this study overlapped with symptoms of physical illness the potential for confounding symptoms attributed to illness with those attributed to depression was heightened. For example, respondents who reported thoughts of death, depressed mood, and feelings of worthlessness received the same rating as those who reported an appetite disturbance, fatigue, and sleep disturbance.

Thomas et al. (1992) also found no relationship between depressive symptoms and mortality in a longitudinal study. Their community sample

consisted of 1,855 participants who were at least 65 years old (average age 75 years). They reported that health status was significantly related to rate of survival. However, neither baseline depression rates nor the presence of depressive symptoms for over two years increased the rate of mortality. One of the strengths of the Thomas et al. (1992) study was the assessment of severe illness and declining health. However, similar to Baldwin and Jolley (1986), they did not measure degrees of symptom severity and acknowledged that both severity and length of time with the disease might influence mortality rates.

Researchers have also examined the reciprocal role between depression and health. With a sample of 2,572 patients over the age of 60 years, Oslin et al. (2000) examined the relationship between physical illness and late life depression. Following treatment at an inpatient facility (i.e., at discharge), there was a significant reduction in depressive symptoms (with a 42 percent rate of remission). Along with improvements in mood, patients reported a decrease in self-reported physical health burden (including the domains of mental health, pain, social functioning, energy/fatigue, role emotional, role physical). As depressed mood improved there was also a significant decrease in the number of illnesses by the patients (2.84 medical conditions at admission versus 1.98 conditions at discharge). One difficulty in interpreting these results is that improvements in mood may have influenced the patients' self-reported perception of physical health. Also, understanding the relationship between improvement in depressive symptomatology and severity of illness in this study was complicated by the authors' measure of physical illness. The authors

measured physical illness using a 20-item checklist of common illnesses which may have lacked sensitivity to adequately assess severity of illness.

Research has also demonstrated gender differences with respect to depression among the elderly. Depressed men have significantly higher mortality rates than depressed women (Burvill et al., 1991; Copeland et al., 1992; Murphy et al., 1988). Schulz et al. (1994) reported that higher depression scores were predictive of poorer perceived health among both men and women but noted this effect was stronger for men than for women. However, the authors of these studies failed to control for severity of illness in their analyses. Thus, it remains uncertain whether the higher mortality rate in depressed elderly men could have been influenced by differences in severity of physical illness.

In sum, the nature of the relationship between depression in elderly populations and mortality remains unclear. The studies reviewed do not support any definitive conclusions regarding the prognostic importance of physical illness and depression in old age. It is possible that the ambiguity in this area of research is in part due to considerable variations in the method of assessing physical illness. Two large reviews of the literature (Cohen-Cole & Kaufman, 1993; Schneider & Olin, 1995) identified a failure to account for severity of illness as a major methodological flaw in research examining the comorbidity of physical illness and depression. Thus, our knowledge about the relationship between depression and medical illness has been limited by the lack of a valid geriatric severity of illness instrument. It appears that a necessary first step for researchers attempting to tease apart the complex relationship between

physical health and geriatric depression involves the development of a reliable and valid measure of illness severity.

The assessment of depression in older persons is, in part, complicated by the considerate overlap between symptoms of depression and symptoms associated with certain medical conditions. This difficulty has also been observed with late-life anxiety (Stanley, Novy, Bourland, Beck, & Averill, 2001). For instance, both anxiety responses and cardiovascular events can present with the same symptoms (e.g., heart palpitations, breathlessness, chest pain, dizziness, and sweating). This overlap of symptoms further supports the potential utility of a measure such as SOI-G which can isolate such variability due to variations in physical/medical difficulties from the somatic aspects of psychiatric conditions, such as depression or anxiety. It is important that researchers control for disease variability that is not measured by psychological tests but must be accounted for in research designs.

2.5 Measurement of Health Status

There are two major approaches to the measurement of health status among the elderly (Girzadas et al., 1993). The severity of illness model focuses on the extent to which alterations in anatomy and physiology pose a threat to life (Girzadas et al., 1993). In contrast, functional measures of illness severity rely on self-assessments of health and functional status (Girzadas et al., 1993). While the former approach emphasizes clinical criteria and physical examination, the latter approach tends to de-emphasize physical disease states and instead focuses on the participant's subjective explanation of their health.

The sections that follow describe some of the major instruments and methods whose purpose is to diagnose and assess functional status and severity of illness. The selection of instruments in the following review is not intended to be exhaustive but rather intended to highlight the scales most commonly used by researchers and clinicians. The applicability of these instruments and methods to a geriatric population will also be discussed.

2.5.1 Functional Measures of Illness Severity

Both clinicians and researchers use functional measures. The assessment of functional limitation has long been a major component in research with elderly persons (Rozzini, et al., 1993). These scales are also widely used in the public domain for purposes such as eligibility for assistance, social policy formation, and to assess disease-specific changes in health status (Johnson & Wolinsky, 1993). The 1960s witnessed a sixfold increase in measures designed to assess functional status (McHorney, 1996). Since 1970, McHorney (1996) estimates an additional fourfold increase in the number of functional status measures developed for use with an elderly population.

2.5.1.1 Sickness Impact Profile (SIP)

The Sickness Impact Profile (SIP) is a widely used, valid and reliable measure of functional status (De-Bruin et al., 1992). The SIP was developed by Bergner, Bobbitt, Carter, and Gilson (1981) and can be administered by an interviewer or self-administered. The SIP operationalizes sickness as the extent of behavioural change in the performance of one's daily life activities (De-Bruin et al., 1992). The individual items are worded so as to approximate the way in

which patients generally describe their sickness (e.g., "I am not doing any of the maintenance or repair work around the house that I usually do"). It was designed as a non-disease specific health status indicator that would be applicable across types and severities of disorders (Bergner et al., 1981). The SIP was designed for use as an outcome variable in health surveys, as a program planning and patient monitoring tool, and to inform policymakers in their decisions (De-Bruin et al., 1992).

The SIP consists of 136 items that are grouped into 12 categories (De-Bruin et al., 1992). Three categories (ambulation, mobility, body care and movement) are subsumed under a physical dimension (De-Bruin et al., 1992). A psychosocial dimension consists of four other categories (social interaction, alertness behaviour, emotional behaviour, communication) while the remaining five categories (sleep and rest, eating, work, home management, recreation and pastimes) are independent dimensions (De-Bruin et al., 1992). It has a number of scoring options: a total score, summary scores for physical and psychosocial functioning and 12 specific subscale scores (Bergner et al., 1981). Higher scores reflect greater self-perceived "sickness impact" and thus poorer health. The SIP has been applied to groups of patients with rheumatoid arthritis, angina, back pain, cancer, obesity, end stage renal disease, myocardial infarction, benign chronic pain, chronic obstructive pulmonary disease, cardiac arrest, amyotrophic lateral sclerosis, ulcerative colitis, Crohn's disease, hyperthyroidism, and physical disability (De-Bruin et al., 1992; Patrick & Deyo, 1989).

In their review of the literature, De-Bruin and colleagues (1992) concluded that the SIP represents a standard against which other measures might be usefully compared. However, they did acknowledge the following deficiencies of the SIP: (1) the clinical relevance of the measure remains uncertain; (2) the scoring procedure is in need of simplification; (3) the extent to which the SIP reflects clinical changes are not known; thus, the measure is not appropriate for longitudinal designs; and, (4) SIP scores do not reflect any theoretical underpinnings.

The SIP appears to be a useful tool with chronically or terminally ill elderly (De-Bruin et al., 1992). However, a study by Andresen, Patrick, Carter, and Malmgren (1995) concluded that the SIP was not an appropriate measure of health status with healthy, community-dwelling older adults because of observed ceiling effects. Rothman, Hedrick and Inui (1989) examined the feasibility of administering the SIP to a selected group of Veterans Administration (VA) nursing home residents (n = 186 males). Rothman and colleagues suggested several modifications of the SIP (e.g., including examples that were more relevant to a frail elderly population). Further, they suggested the inclusion of a definition of health because although the instructions on the SIP asked the respondent to rate his/her health, no definition of health is provided. Thus, the lack of a definition may increase the between subject variability. Rothman et al. also pointed out that there were categories of the SIP that were no longer relevant to nursing home residents, such as work and home management and they cautioned that the SIP is not appropriate for use with

cognitively impaired residents. Lastly, although the SIP was sensitive to small changes in impairment, the measure was somewhat long for frail participants. Consistent with this, Ware and Sherbourne (1992) point out that the SIP has four times the respondent burden of a shorter health status measure (SF-36).

2.5.1.2 Medical Outcomes Study (MOS) SF-36 Health Survey

The SF-36 Health Survey is a generic measure designed to assess the patient's self-assessment of functional status as opposed to measuring the underlying disease (Ware & Sherbourne, 1992) and was one of the measures used in the present research. The SF-36 is composed of 36 items that measures eight generic health concepts: (1) limitations in physical activities because of health problems; (2) role disability due to physical health problems; (3) bodily pain; (4) general health perceptions; (5) vitality (energy and fatigue); (6) social functioning; (7) role disability due to emotional health problems; and (8) general mental health (psychological distress and well-being; Ware & Sherbourne, 1992). Low scores reflect the least favorable health states (e.g., "feels tired and worn out all the time") while high scores represent the most favorable health state (e.g., "feels full of pep and energy all of the time").

The SF-36 was constructed to survey health status as part of a longitudinal component of the Medical Outcomes Study, a four-year observational study of variations in physician practice styles and patient outcomes in different systems of care (McHorney, 1996). The SF-36 evolved from other measures developed for the MOS, including the SF-20 Health Survey, the 149-Item Functioning and Well-Being Profile, and measures from

the Health Insurance Experiment (McHorney, 1996). Estimates suggest that the SF-36 has been used in more than 200 clinical trials (Anderson, Aaronson, & Wilkin, 1993). With a sample of 3,445 patients consisting of 24 subgroups differing in sociodemographic variables (age, gender, race, education, and poverty status), diagnosis, and disease severity, McHorney, Ware, Lu, and Sherbourne (1994) reported that the internal-consistency reliability of the SF-36 exceeded recommended standards for group comparisons (0.78 to 0.93).

McHorney (1996) reported on research that was conducted on the applicability of the SF-36 with a geriatric population. McHorney points out that the utility of the SF-36 with older adults may be limited by the fact that it does not assess variables that are applicable to elderly persons, such as mobility, sleep, and health satisfaction. Similarly, Berkman et al. (1999) reported that the SF-36 failed to assess variables that might be instrumental in a social work assessment. They reported that the SF-36 was not significantly correlated with the following psychosocial risk factors: sleep, memory, compliance with dietary restrictions, concentration, alcohol or drug abuse, hearing, sex, dizziness, managing money, vision, urinary incontinence, and use of the telephone.

Although other generic measures of health status have been observed to demonstrate ceiling effects when used with elderly persons, McHorney (1996) points out that the SF-36 exhibits differential floor and ceiling effects. That is, floor effects are more commonly reported with samples of older patients and the chronically or acutely ill, while ceiling effects are more common among younger or well general populations (Andresen et al.,1995). The observed ceiling and

floor effects may give rise to a false perception of homogeneity in health among individuals who in fact vary on the variable being measured. Ceiling effects prevent the assessment of improvement in health over time and floor effects hinder the measurement of decline in health.

McHorney and colleagues (1994) also point out that rates of data completeness were lower for elderly patients. McHorney (1996) suggests that lower rates of data completeness may be due to the readability of the questionnaire. McHorney states that the SF-36 is written at a seventh-grade reading level or higher; thus, it is estimated that at least one third of elderly persons would have difficulty completing the questions. McHorney also questions the use of an optical mark reading (computer scanning) answer format for elderly persons. Others have found high rates of missing data (21 percent) for elderly respondents and have suggested that the small typeface may be difficult for those with visual disorders which is a common difficulty experienced by elderly persons (Bjorner & Kristensen, 1999). In addition, elderly persons have less familiarity with such forms thus potentially compromising the quality and validity of responses.

Lastly, although all self-report measures are subject to the criticism of response bias, McHorney (1996) reported a specific pattern of bias, namely an age-mode interaction for the three scales that measure social and role functioning. Consistent with social role theory, older individuals appeared to be more reluctant to admit to role disability. It is expected that this pattern also generalizes to many of the self-report health status measures.

2.5.1.3 Activities of Daily Living (ADLs)

The Activities of Daily Living (ADL) is another approach to measuring general health status that emphasizes personal and social functioning as an indicator of overall health status (Parmelee et al., 1995). It is suggested that the presence of mild functional limitations may occur in preclinical or subclinical levels of disease and thus detection of these changes in functioning may identify those at high risk of developing a subsequent disability (Rozzini et al., 1993). There are two general ways in which functional disability is assessed. The first involves obtaining the person's self-reported physical functioning including the Basic Activities of Daily Living (BADL) and Instrumental Activities of Daily Living (IADL). BADL involves the most basic personal care tasks, such as feeding, grooming, toileting, transferring (moving in and out of a bed or chair), eating, dressing, bathing, and motility while IADL consists of complex activities needed for independent living, such as handling of personal finances, preparing meals, using the telephone, shopping, traveling, doing housework, and taking medications. The second approach, performance-based measures, involves having the participant perform a function that is rated by an observer (Rozzini et al., 1993).

With respect to gerontological research, both the BADL and the IADL have been criticized. McHorney (1996) was critical of BADL and IADL measures for limited sensitivity in detecting mild functional impairment (i.e., ceiling effects). For instance, it has been estimated that fewer than 5 percent of community-dwelling elderly experience difficulty using the toilet and less than 10

percent have problems with respect to bathing or dressing (McHorney, 1996). Rozzini et al. (1993) hypothesize that individuals experiencing an initial level of functional limitation might adopt compensatory (albeit more time-consuming) methods to accomplish a task and thus this type of functional decline goes undetected by ADLs. Wilms, Kanowski and Baltes (2000) discuss a second criticism, namely the difficulty in differentiating functional declines associated with psychiatric disorders (e.g., dementia) from functional limitations that are a consequence of comorbid physical health difficulties or age related effects.

2.5.1.4 The Nottingham Health Profile (NHP)

The Nottingham Health Profile (NHP) is a generic, self-report measure of health related quality of life and functional ability associated with major disabling health conditions, i.e., disease or injury (Anderson, Aaronson, & Wilkin, 1993). The NHP consists of two parts (Anderson et al., 1993). The 38 items that make up Part I address the following six domains of distress: (1) energy level; (2) pain; (3) emotional reactions; (4) sleep; (5) social isolation; and, (6) physical abilities (Anderson et al., 1993). Items that make up Part II address the impact of perceived health problems on seven areas of life: (1) work; (2) home maintenance; (3) social life; (4) home life; (5) sex life; (6) interests and hobbies; and, (7) holidays (Anderson et al., 1993).

Anderson et al. (1993) reviewed the literature on the NHP and concluded that the NHP can differentiate between the severely ill patient samples and well individuals. Further, with a sample of 1,587 elderly persons, Noro and Aro (1996) demonstrated that the NHP could reliably differentiate between the least

dependent elderly respondents living in residential care and their age peers in the general population. However, similar to other functional health measures, Anderson and colleagues reported that the NHP exhibits limited sensitivity when used with participants who demonstrate minor to mild symptoms. Anderson et al. reported that as many as 30 percent of respondents with less severe symptomatic distress obtained scores of zero, presenting the misleading picture of a problem-free state in mildly symptomatic individuals.

2.5.1.5 The OARS Multidimensional Functional Assessment Questionnaire (OMFAQ)

The Older Americans Resources and Services (OARS) Multidimensional Functional Assessment Questionnaire (OMFAQ) was developed at Duke University in the early 1970s to assess individual physical, psychological, and social functioning, and services utilization (McDowell & Newell, 1996). Unlike previously reviewed instruments that were typically developed for an adult population and then later evaluated for their applicability with the elderly, the OARS was developed specifically for use on elderly populations (Whitelaw & Liang, 1991). The OMFAQ consists of two major divisions, the Multidimensional Functional Assessment Questionnaire (part A) and the Services Assessment Questionnaire (part B). Part A is further divided into five dimensions, consisting of social and economic resources, mental and physical health, and activities of daily living (McDowell & Newell, 1996). Part B includes 24 categories of services received and needed (McDowell & Newell, 1996).

One strength of the OMFAQ is its ability to provide a comprehensive profile of personal functioning and service use (McDowall & Nowell, 1996). Whitelaw and Liang (1991) identified that OMFAQ as an appropriate choice for the assessment of physical health. However, as a severity of illness measure it would be inappropriate as much of the content of the physical health scale includes ratings that may be confounded by variables that are tangentially related to severity of illness, such as availability and access to care (e.g., physician visits, patient in hospital or nursing home) or psychological variables (e.g., self-assessment of health).

2.5.2 Strengths and Weakness of Using Functional Measures with the Elderly

One strength of functional status measures is that they are assumed to be appropriate for use with populations differing in age, race, and SES (McHorney, 1996). Secondly, Ware and Sherbourne (1992) contend that measuring the experience of health from the perspective of the patient is among the most important health care advances made during the past decade. Research suggests that the manner in which a person views his/her health is associated with ratings made by physicians, however, others (Conwell et al., 1993) report that the association between self-report measures of physical health and objective measures is poor. In addition to these contradictory findings, there are several reasons why the use of functional measures to assess severity of illness in the elderly is either inadequate or inappropriate and could potentially lead to mistaken conclusions.

One limiting factor of functional scales is the potential for combining both physical and psychological factors (e.g., social, cognitive, motivational). One cannot assume, for instance, that difficulty completing a task is synonymous with nonperformance of a task (Lohr, 1992). The latter instance may reflect a stronger influence of psychological variables (e.g., malingering, defiance, or lack of motivation) than the former, yet both instances may produce similar scores on a functional scale suggesting the presence of similar physical inability. Also, because of the extensive presence of a psychosocial component in all categories, ratings on functional measures may reflect the impact of the patient's reactions to a presenting disease state. The implication of this limitation is that the same degree of physical illness could lead to different impact scores, depending on the individual. For example, some individuals contend with chronic disease better than others do. While some individuals continue to function at a level above what would be expected for their severity of disease, others are completely incapacitated by the same level of disease (Johnson & Wolinsky, 1993). This is consistent with Gove's (1984) contention that functional impairment is considered to be a measure of sick role behaviour.

Second, functional disability is not equivalent to severity of illness. Different health problems have differential implications for health and longevity yet could theoretically lead to similar ratings on a functional index. For example, the early stage of both Alzheimer's disease (AD) and skin cancer are expected to have very little influence on functional ability yet skin cancer poses a greater threat to life (at the physiological level) than AD (Johnson & Wolinsky, 1994).

Third, many functional scales do not adequately measure mild symptom severity. It appears that the usefulness of these measures is not fully realized until there is a substantial decline in functional ability or until the development of multiple pathological conditions (i.e., comorbidity) interferes with a person's ability to perform the rated task (Rozzini et al., 1993). This suggests that the usefulness of functional measures is limited to elderly persons exhibiting more extreme levels of functional impairment. However, greater functional disability is also associated with poor performance on cognitive measures (Rozzini et al., 1993). For instance, Melding (1995) observed that individuals who are cognitively impaired have difficulties with formal operational thinking, resulting in overly simplistic self-reported concepts of pain and illness. In light of this and other research, it has been recommended that researchers exclude individuals with cognitive impairment from geriatric functional assessments to avoid compromising the self-report data (Noro and Aro, 1996; McHorney, 1996). Thus, it would appear that functional measures are most appropriate for a subgroup of elderly with considerable functional impairment but who do not exhibit cognitive deficits, thus further limiting the percentage of elderly for whom the use of self-report functional measures is considered appropriate.

Taking into account the recommended exclusion of participants with cognitive impairment, the applicability of functional measures with elderly in the upper age ranges is limited by the fact that the rate of cognitive impairment increases with advancing age (Government of Canada, 1996). Information provided by the National Advisory Council on Aging suggests that the rate of

cognitive impairment for Canadians doubles from the 75-79 to the 80-84 age ranges. The rate of increase again doubles from those 80-84 to those over the age of 85 (Government of Canada, 1996). The Canadian Study of Health and Aging Working Group (1994) estimate that the number of Canadians with dementia will rise from 252,600 in 1992 to 592,000 by 2021. In light of the growing number of elderly persons with known cognitive impairments, the need for a measure that enables researchers to assess health independent of cognitive function becomes more imperative.

Fourth, functional measures that rely on physician records to obtain data are problematic because physicians' notes frequently contain little documentation on functional impairment of elderly patients admitted to hospitals (Burns et al., 1992). In a study that examined the medical records of 2,504 patients over 65 years of age, Burns et al. (1992) reported that the degree of missing data varied by function. Findings indicated that data on bathing was missing for 20 percent of the files while dressing was unrecorded in half of the cases. Further, 10 percent of the files had no functional status information.

A fifth limitation is that many functional health status measures involve ratings based on self-report data. The use of self-report measures to assess level of severity of illness is problematic for a number of reasons. First, the participant's ability to remember and accurately describe their physical health limits the validity of the physical health ratings (Murphy et al., 1988). Second, research suggests that the underreporting of physical symptoms is a common theme among elderly persons and this underreporting could lead to unrecorded

health problems (Besdine, 1997). This pattern of underreporting may, in part, reflect a tendency of patients who have grown used to the difficulties associated with a particular illness to under-emphasize the degree of disability present (Burvill et al., 1990). Lastly, research has suggested that a person's ability to rate their physical health may be influenced in a negative direction by the presence of depressed affect (Oxman, 1996; Thompson, 1996). Thus, depressed patients may rate their health as poorer because of distorted thought patterns rather than because of an objective difference in physical health.

Finally, gender differences have been observed in both self-report functional measures of illness severity (Johnson & Wolinsly, 1993; Schulz et al., 1994) and activity limitation ratings made by others (Johnson & Wolinsky, 1994). With a sample of 5,151 elderly men and women (over the age of 70 years), Johnson and Wolinsky (1993) demonstrated that women evaluated their health more positively than men. In contrast, Schultz et al. (1994) found no gender differences in self-rated health status (n = 5,201, aged 65 years or older). However when Schultz and colleagues completed a regression analysis, the pattern of variables that were predictive of perception of health did vary as a function of gender. For example, men with lower income also reported lower perceived health while being less educated was a more important predictor of lowered perceived health for women. Exercise tolerance was another example of the gender difference in predictor variables. Exercise tolerance accounted for an additional 8 percent of the variance in perceived health for women while for men this variable accounted for only an additional 3 percent of the variance.

Johnson and Wolinsky (1994) also found evidence of a measurement bias with respect to gender. The authors tested a model of health status (previously developed by the authors) using 5,151 men and women aged 70 years or more. Activity limitations were assessed using the basic ADLs. It was determined that differences in measurement validity existed between the sexes for several elements of their model of health status. In some cases items were measured more validly for males (e.g., the need for assistance in dressing) while in other cases (e.g., toileting) the opposite was true. This means that these items will either consistently underestimate or overestimate the effect of that item dependent on the gender of the participant.

In sum, it appears that although measures of functional disability provide important information on the person's functional limitations they provide little information on the extent of organic pathology (Linn, Linn, & Guriel, 1968). One implication of the variable relationship between disease processes and functional limitation is that using functional measures in studies aimed at examining the relationship between physical illness and psychosocial variables in the elderly can confound the predictor and outcome variables. Thus, it is important to differentiate between physical illness and the functional disability it may cause (Parmelee et al., 1995).

2.5.3 Severity of Illness Measurement Systems

Severity of illness measurement systems can be further subdivided into either generic or disease specific categories. Generic measures are designed to be broadly applicable across different types and varying severity of disease,

across different medical treatments, and across cultural subgroups (Patrick & Deyo, 1989). Such measures allow for comparisons not possible with disease specific measures, such as comparing the degree of health burden associated with different diseases (Ware and Sherbourne, 1992). These measures typically yield either a single index value or a profile of interrelated scores (Bergner, et al., 1981; Stewart, Hays, & Ware, 1988). In contrast, disease specific measures are designed to assess specific diagnostic groups. Not all specific measures are disease related but instead are specific to certain conditions, such as level of consciousness with the Glasgow Coma Scale (Bastos, Sun, Wagner, Wu, & Knaus, 1993).

Specific measures may be better suited for investigations of outcome for a particular clinical intervention while generic measures may be more useful for comparisons across different diagnostic groups (Patrick & Deyo, 1989). However, both Patrick & Deyo (1989) and Lohr (1992) suggest that health status is best measured using generic measures and that disease-specific measures are most usefully applied as supplements to the generic measures. Charlson et al. (1986) indicates that much of the research with medical patients involve patients who are not critically ill for whom outcome would be best estimated by generic measures rather than by disease-specific instruments.

Parmelee et al. (1995) suggest that a measure of severity of illness that addresses the high rate of co-morbidity often found with elderly persons is needed. Parmelee et al. added that a severity of illness scale that provides a summary index of physical illness would further research with geriatric samples.

The aim of the present investigation was to merge the two sub-categories of severity of illness measurement systems and develop a generically applicable measure of severity of illness consisting of disease-specific scales. A brief and selective review of severity of illness measures will be presented with the purpose of illustrating the limitations of existing severity of illness scales. Generic measures will be reviewed first followed by disease specific scales.

2.5.4 Generic Measures of Illness Severity

The generic measures of illness severity that will be reviewed include the Cumulative Illness Rating Scale, the Cumulative Illness Rating Scale-Geriatric, the Severity of Illness Index, the Computerized Severity Index, and the Acute Physiology and Chronic Health Evaluation III.

2.5.4.1 Cumulative Illness Rating Scale (CIRS)

The Cumulative Illness Rating Scale (CIRS) was developed by Linn et al. (1968) and intended to serve as a short, comprehensive, and reliable instrument that assesses physical impairment. The scale format provides severity of impairment ratings for 12 relatively independent organ systems and a psychiatric/behavioral category. Ratings are made for each of the 13 items on a 5-point ordinal scale, ranging from none (i.e., no impairment to that organ/system) to extremely severe (i.e., impairment is life-threatening). The underlying assumption of this scale is that the more organ systems involved in a patient's condition the greater the severity of illness.

Linn and colleagues assert that, as a biologic measure of aging, the CIRS was more accurate in predicating an individual's capacity for survival than

chronological age. Conwell et al. (1993) conducted a validation study by examining how closely CIRS scores taken from retrospective accounts and medical records correlated with autopsy ratings. Subjects consisted of 72 (48 male, 24 female) completed suicide victims who ranged in age from 21 to 92 years (mean = 54.6 years, SD = 20.2 years). CIRS scores accounted for 75 percent of the variance in gross and microscopic observation of tissue obtained at autopsy. However there was a systematic bias in that the CIRS over-estimated severity of illness at low levels of tissue pathology and underrated severity of illness for moderate to severe levels of pathology.

The CIRS has also been used in research with elderly persons. Parmelee and colleagues (1995) applied the CIRS to a geriatric institutional sample of 439 residents of a multilevel care facility. The average of the 13 CIRS subscale scores were significant predictors of mortality over two years, with greater average severity associated with increased mortality. CIRS scores also predicted acute hospitalization within a year post assessment and there were significant low to moderate correlations between 11 (of 13) CIRS scales and functional ability (range r = .11 to .32).

2.5.4.2 Cumulative Illness Rating Scale-Geriatric (CIRS-G)

One problem associated with using the CIRS was that raters often requested more specific guidelines (Miller et al., 1992). In response to this concern and the need for a severity of illness measure for a geriatric population, Miller et al. (1992) modified the CIRS to reflect common problems associated with an elderly population. Their modification included a score reflecting the

number of organ-specific categories at severity level 3 and severity level 4. The modified scale was referred to as the CIRS-G and was used in the present research. However, consistent with the aim of the present research, the authors also acknowledge the potential utility of a weighting process that reflects the differences in threat to life that exist between each of the items on the CIRS-G.

Miller et al. (1992) applied the CIRS-G to sample of 141 elderly outpatient subjects. The authors reported a significant Spearman rank order correlation of 0.58 between the CIRS-G total scores and the Activities of Daily Living (ADL) score suggesting that severity of illness may influence the ability of elderly persons to perform activities of daily living. The correlation between CIRS and CIRS-G scores and functional ability observed by both Parmelee et al. (1995) and Miller et al. (1992) illustrates one limitation of the CIRS, namely that the ratings may reflect limitations in function rather than life-threatening potential per se. As pointed out by Parmelee and colleagues, it is important to distinguish severity of illness from functional disability as a functional deficit may be due to any number of specific physical and/or psychosocial problems.

One further complication of both the CIRS and CIRS-G is the inclusion of a psychiatric item in the total score. The combination of physical and psychological variables limits the usefulness of the CIRS-G as a control measure in studies aimed at examining the possible impact of health status on psychological variables (e.g., depression) within a geriatric sample. For instance, in a study of 115 elderly patients (ambulatory and without dementia), Miller et al. (1996) compared the severity of illness scores (as measured by the

CIRS-G) of recovered and non-recovered depressed patients. The authors were examining the effectiveness of combined nortriptyline (NT) and weekly interpersonal psychotherapy sessions on depression. The authors observed that the participants who did not achieve remission were neither more depressed nor more medically burdened (pre-treatment) than responders. However, we can not assume that the failure to reject the null hypothesis implies that severity of illness had no impact on the effectiveness of treatment as the severity of illness measure in this instance may not have been entirely independent of the psychological variable being assessed.

2.5.4.3 Severity of Illness Index

Horn, Sharkey, and Bertram (1983) originally designed the Severity of Illness Index in 1979-1980. It is a four-level, ordinal, generic index that is determined from seven dimensions thought to reflect the patient's "burden of illness". The seven dimensions include: (1) stage of principal diagnosis; (2) concurrent interacting conditions; (3) rate of response to therapy; (4) impairment remaining after therapy; (5) complications of the principal diagnosis; (6) patient dependency on hospital facilities and staff; and, (7) extent of non-operating room procedures. Green, Wintfeld, Sharkey, and Passman (1990) used the stage of principal diagnosis at admission (a subscale of the Severity of Illness Index) as a measure of illness severity. These authors reported that this subscale improved the accuracy of mortality predications in their study involving 13 hospitals (34,252 patients).

However, the Severity of Illness Index may not reflect the degree to which symptoms pose a threat to life as many of the variables included in the scale reflect factors other than physical illness. For example, patient dependency on hospital facilities may vary as a function of the availability of hospital resources or, in the case where patients pay directly for hospital care, "dependency" may reflect the patient's financial ability to pay for the available facilities. Also, rate of response to therapy may be influenced by psychological factors, such as motivation, presence of depression, personality traits (such as hardiness), or represent secondary gains to be attained by adopting a sick role (e.g., obtaining social support from others).

2.5.4.4 Computerized Severity Index (CSI)

Developed by Horn and colleagues (1991), the CSI categorizes patient severity of illness on the basis of physician rated clinical signs (e.g., level of consciousness, breath sounds) recorded during hospitalization and depicts a modernization of her Severity of Illness Index. The CSI was originally developed to determine the degree of improvement in the prediction of hospital length of stay and mortality that could be achieved by adjusting the existing diagnosis-related group (DRG) system for patient severity of illness (Horn et al., 1991). CSI determines severity of illness through a modification of the ICD-9-CM codes already in existence for patients' principal and secondary diagnoses.

Aronow (1988) criticizes the CSI for its lack of guidelines for timing of severity assessments. This limitation threatens the consistency of use and thus limits the CSI's reliability and makes cross study comparisons problematic.

Aronow adds that classifications based on ICD codes are limited because the ICD diagnoses are not independent of the services rendered. Thus, Aronow argues that ratings based on ICD codes may dilute the quality of the information extracted from medical records because such ratings group patients on the basis of a label (in a manner similar to Disease Staging) rather than according to illness severity. Aronow also criticizes the retrospective element of the CSI that arises when severity during hospitalization is guided by the final diagnoses and thus not necessarily how the care-giving process unfolded. One final limitation of the CSI is the need for ratings to be made by a physician.

2.5.4.5 Acute Physiology and Chronic Health Evaluation (APACHE III)

The Acute Physiology and Chronic Health Evaluation III (APACHE III) is a scale designed to categorize groups of patients on the basis of illness severity (Knaus, Zimmerman, Wagner, Draper, & Lawrence, 1981). The development of the original APACHE was driven by the need for methodology that would enable researchers to form groups of patients who were homogeneous with respect to severity of illness in order to more effectively evaluate and assess new intensive care therapies (Knaus et al., 1981). Knaus et al. (1981) maintain that the APACHE is a non-disease specific system appropriately applied to ratings groups of patients with diverse conditions or diseases.

The original APACHE method of categorizing consisted of two divisions: (1) a physiologically based score that captured the degree of acute illness; and, (2) an evaluation of health status six months before admission (Knaus et al., 1981). According to Knaus et al. (1981), the acute physiology score consisted

of 34 measurements, representing seven major physiological systems (neurological, cardiovascular, respiratory, gastrointestinal, renal, metabolic, and hematological). The pre-admission score was a measure of functional status, for example "chronic disease producing serious but not incapacitating restriction of activity" (Knaus et al., 1981).

The APACHE system has evolved since its inception. The APACHE II was developed in 1985 and consisted of 12 physiologic variables, chronic health status, and age (Knaus, Draper, Wagner, & Zimmerman, 1985). APACHE III scores reflect the extent of abnormality of 17 physiologic measures, the acute physiology score, weights for age, and weights based on seven cormorbidities that reduce immune function and influence hospital survival. Higher APACHE III scores are thought to reflect higher hospital mortality rates for patients admitted to the Intensive Care Unit (Zimmerman et al., 1996b).

Numerous studies have demonstrated the validity of the APACHE methodology (including the APACHE II and III) with respect to predicting mortality in acute care settings (e.g., ICU). For instance, Wong, Barrow, Gomez, and McGuire (1996) used the APACHE II to predict group mortality in intensive care unit (ICU) trauma victims. Prospective information was collected for 470 ICU trauma patients. Thirteen percent of patients died while 87 percent survived. Overall, APACHE II scores correctly classified greater than 90 percent of all patients. However, of those predicted to die, 25 percent were false-positives (i.e., 25 percent were predicted to die but actually survived).

In addition, the rate of death predicted by the APACHE II underestimated the actual observed death rates, particularly for high-risk patients.

Becker et al. (1995) used the acute physiology score (APS) of the APACHE III to predict hospital mortality with a sample of 2,435 patients (from six hospitals in the USA) admitted to the ICU after undergoing coronary artery by-pass surgery. The APS score explained 81.9 percent of hospital mortality. When data were combined from the six hospitals, the predicted hospital mortality (3.85 percent) matched actual hospital mortality (3.86 percent). However, there was considerable variability between hospitals. APS scores underestimated mortality at four hospitals and over-estimated mortality at the remaining two hospitals included in the study. One other limitation of this study was that data were collected after admission to the ICU, thus the APS scores used in the prediction equations reflected postoperative severity of illness.

Zimmerman and colleagues (1996a) also used the acute physiology score (APS) of the APACHE II and III to compare the outcomes for patients (from 53 U.S. hospitals) with one or more organ system failures (a total of 7,703 ICU admissions). The authors reported that severity of physiologic dysfunction (as measured by the APS) increased as the number of organ system failures increased. In addition, the APACHE III level of predicted risk was highly correlated with actual rates of hospital mortality. Lastly, APACHE III mortality predictions were more accurate predictors of mortality than were the number of organ system failures. The latter finding was replicated by Zimmerman et al.

(1996b) who demonstrated that the best predictor of death on the first day of admission to the ICU in patients with cirrhosis was the APACHE III score.

In sum, the APACHE III appears to be useful and valid as a measure of acute care illness, however there are potential limitations associated with using this scale. First, the APACHE was developed for use with patients in critical care units of hospitals and thus the usefulness of the APACHE with chronic illness is questionable (Parmelee et al., 1995). Proponents of the APACHE III (e.g., Wong et al., 1996) maintain that the pre-admission score is a measure of chronic health status; however, this claim has not been supported empirically and in fact, there is evidence to suggest that the APACHE's predictability with chronic illness is limited. For example, research with renal disease has demonstrated the superiority of the SORDS over the APACHE with respect to predicting mortality (Baltzan et al., 1987). Baltzan et al. (1987) found SORDS to be more strongly correlated to death rates than the APACHE scores. Second, the APACHE scales have been criticized for lack of parsimony relative to other, less labor-intensive indicators (Parmelee et al., 1995).

2.5.5 Disease Specific Measures of Illness Severity

The disease specific measures of illness severity that will be reviewed include Physician Diagnosed State, Disease Staging, the Glascow Coma Scale, and the Severity of Renal Disease Scale.

2.5.5.1 Physician Diagnosed State (Disease Naming)

The disease name perspective involves the identification of diagnoses and diseases. In general, disease naming is an unreliable measure of health

status because diagnostic groupings are not sensitive to the broad range of severity within a diagnosis and tend to oversimplify the complexity of most illnesses (Aronow, 1988; Conwell et al., 1993). The use of physician diagnoses per se is not being questioned but rather what is questioned is the use of physicians' diagnoses as a severity of illness measure for research. For example, Zimmerman et al. (1996a) reported varying mortality rates for patients diagnosed with two organ system failures. The mortality rate for patients with hematologic and cardiovascular failure was 20 percent while the mortality rate for patients experiencing cardiovascular and neurologic failures was 76 percent. Similarly, with a sample of 194 long-stay elderly nursing home residents, Mulrow, Gerety, Cornell, Lawrence, and Kanten (1994) observed that merely summing the numbers of disease categories accounted for an insignificant amount of the activities of daily living (ADL) variance ($r^2 = 0.03$). In comparison, when Mulrow and colleagues used summary scores that included severity ratings, these scores account for 25 percent of the variance in ADL scores.

2.5.5.2 Disease Staging

Disease Staging (DS), developed by Joseph S. Gonnella, is a severity of illness measurement method that classifies patients according to the clinical stage of their disease (Aronow, 1988). DS assigns one of four stages of severity for 400 common diseases (Aronow, 1988). One advantage of this method is that it is independent of the practices of the health care agency in which the ratings are completed thus facilitating cross study comparisons. This approach is limited, however, because it only examines the stage of each

separate disease and does not provide an overall severity score and thus does not consider the patient's total burden of illness (Horn and Horn, 1986). Also, comparisons between disease categories are not appropriate (Aronow, 1988). Lastly, DS has been criticized for its failure to consistently predict health care resource use (Aronow, 1988). The latter concern most likely reflects the large degree of variability that exists between patients at the same stage of a disease.

2.5.5.3 Glasgow Coma Scale

The Glasgow Coma Scale was developed in 1974 as a standard method for assessing neurologic status with head trauma patients (Bastos et al., 1993). Subsequent research has demonstrated an association between level of consciousness and mortality risk in critically ill patients (Bastos et al., 1993). This research led to the expanded use of the Glasgow Coma Scale with a variety of critically ill patients, both with and without head injury.

The validity of using the Glasgow Coma Scale with other medical conditions has been questioned (Bastos et al., 1993). Bastos and colleagues (1993) addressed this concern by investigating the appropriateness of using the Glasgow Coma scale with non-traumatic critically ill patients. The study population consisted of 15,973 ICU patients without trauma. A significant but nonlinear relationship was observed between the Glasgow Coma Scale score at admission and subsequent outcome. The scale lacked sensitivity in the intermediate range of scores but it demonstrated good discrimination for extreme values (i.e., those likely to die and those likely to survive). The authors reported that for those patients in the intermediate range of scores, factors such

as acute physiologic measurements, age, comorbidities, and disease etiology were better predictors of outcome than the Glasgow scale.

2.5.5.4 The Severity of Renal Disease Scale (SORDS)

SORDS is an example of a disease specific research instrument that was designed to evaluate physical health specific to renal patients (Baltzan et al., 1987; see Appendix A). It is the instrument that was adapted in the present research to develop the Severity of Illness-Geriatric (SOI-G).

SORDS quantifies the potential impact of diseases most often associated with renal failure by the use of objective symptom criteria. It provides a single numerical score, called the severity score, which reflects the disease severity. Disease severity is defined as the degree to which a person is limited in his or her ability to perform normal functions as a consequence of their disease with death identified as the most extreme limitation in functional ability. The extremes for each scale item are 0 (no limitation) to 100 (death). Theoretically, the ranges of scores for the total scale can be from zero to 1,615; however renal patients are not likely to span the entire range (Baltzan et al., 1987). Specifically, a score of zero is impossible for a renal patient because the presence of a diagnosis of chronic renal failure implies a degree of abnormality on at least some SORDS dimensions. Similarly, a score of 1,615 would reflect a total collapse of all organ systems simultaneous and patients would die before such a catastrophic outcome could occur and be measured. The total scale score yields a number on an interval scale that is indicative of severity of illness.

SORDS was originally developed as a tool for research in behavioural medicine and to meet the need to examine the psychological effects of the progression of chronic renal failure. The variation in physical health of end-stage renal patients was thought to contribute significant noise to the data obtained from such patients. This noise, if not accounted for, might act as a suppressor variable in much of the research on renal patients. SORDS was developed with this in mind and was conceived of as a possible suppressor variable measure, in that it had the potential to measure how much the severity of the illness impacted on an individual's response to treatment.

Preliminary research with SORDS demonstrated that it was face and content valid, and a reliable indicator of progressive severity of renal disease (Baltzan et al., 1987). The studies of reliability by Baltzan et al. provided encouraging indications that SORDS was a reliable and valid instrument. Using medical records, two raters individually assessed patients using the SORDS scale. The inter-rater agreement between the scores of these two raters was low but promising. The low agreement was attributed to discrepant data collection methods between the two raters (e.g., one of the raters was a nurse with personal experience of the patients involved in the study). To test the validity of the SORDS, Baltzan et al. (1987) examined the relationship between scores on SORDS and situations corresponding with severity of illness (e.g., death). An examination of patient medical files revealed a significant correlation between SORDS scores and death ($r = .59$) and the type of dialysis patients were receiving (i.e., in center versus home dialysis).

Recent empirical investigations of SORDS conducted by Alexander (2001) provided additional evidence of the validity of SORDS. Using various renal patients groups, SORDS demonstrated good convergent validity with other well-researched measures including the End Stage Renal Disease–Severity Index developed by Craven, Littlefield, Rodin, and Murray (1991) and the SF-36 which measured patients perceptions of decreased physical health and functioning. SORDS scores also differentiated between different renal patient treatment groups (i.e., pre-dialysis versus patients requiring dialysis).

Alexander (2001) also demonstrated the utility of SORDS as a moderator variable in psychological research with renal patients by reporting that type of dialysis moderated the relationship between illness severity and depression. Continuous ambulatory peritoneal dialysis (CAPD) patients were less severely ill (as indicated by lower SORDS scores) than patients receiving in-hospital haemodialysis (HD), yet the CAPD patients exhibited higher levels of depression at the same level of severity of illness. Although not formally assessed, it was suggested that the CAPD patients may have been referred for dialysis more recently than the HD group and therefore may have been at an earlier stage with respect to adjusting to the progressive nature of their illness. This latter finding suggests that SORDS scores may be useful in the assessment of adjustment to treatment and illness severity. Thus, the preliminary evidence with SORDS supports the utility and benefit of severity of illness measures in psychological research.

3. Description of Present Research

The preceding selective literature review focused on a major problem facing researchers studying psychological variables in the elderly, namely the importance of controlling for the wide variability in the physical health status of geriatric populations. In addition, the previous review outlined some of the difficulties associated with using existing severity of illness measures with elderly populations. To date, no measure of geriatric physical illness has been developed that generalizes health status across diseases while simultaneously assessing severity of illness from a physiological perspective. Such an instrument could have potential value for researchers examining morbidity, mortality and psychosocial variables in the elderly.

There are several reasons why a reliable and valid measure of physical illness would be a useful tool for researchers working with an elderly population. First, application of such a measure offers benefits in the assessment of the impact of health policies and programs developed for elderly persons (both in their present day decisions and their projections related to future demand for services). These decisions are to a large extent influenced by the hospital resource use associated with the care of patients. When these decisions are based on diagnosis-related groups the potential for error is great because of the diversity within any diagnostic group. That is, diagnostic groupings are not

sensitive to severity of illness, which can relate to both the quality and the cost of care (Aronow, 1988). In addition, comparisons of costs among different facilities are difficult because of potential differences in the standards of practice in each institution. Such comparisons can be made only if the severity of illness of the patients is considered (Horn et al., 83). Thus, methods that assess severity of illness are necessary to statistically control for the heterogeneity within the various diagnostic categories and between different facilities.

A second advantage deals with the allocation of resources within long-term care facilities. As the number of elderly persons increases so does the need for institutional care programs that meet their needs (Rothman et al., 1989). However, these programs are typically expensive and often in short supply. Measures of illness severity offer researchers a method to assess the impact of such programs on residents (Rothman et al., 1989). Also, the ability to quantify physical illness in the elderly enables researchers to better understand the level and suitability of institutional care for elderly persons thus helping to inform decisions made by front line workers in their everyday work. For instance, there may be sub-groups among the elderly requesting long-term care whose health care needs are such that they could benefit equally from less costly methods of care, such as home care (Noro & Aro, 1996).

The ability to quantify physical illness in psychological research is important in light of research that suggests a higher rate of physical diseases among psychiatric patients than other people (Burvill et al., 1990). Borchelt and colleagues (1999) underscore the reciprocal role between physical health

variables and psychological variables by suggesting that physical illness may play a role as a precipitating factor in the onset of psychiatric disorders such as depression or dementia. Conversely, physical health is also impacted by psychological and social variables. Thus, given the high incidence of physical comorbidity among the elderly, a geriatric severity of illness measure would be of potential benefit used as a control measure in psychological research with the elderly by reducing the suppressing effect of data variability attributable to differences in illness severity. The ability to accurately describe severity of illness would facilitate clearer identification of the relationship between variables of primary research interest by allowing researchers to disentangle indicators of physical health status from indicators of mental health. Although each of the advantages outlined previously are important, the necessity of quantifying severity of illness in psychiatric research with geriatric populations was the primary impetus for the present investigation.

The purpose of the present research was twofold. The first purpose was to develop a generically applicable, quantitative index of illness severity composed of disease-specific scales appropriate for use with a geriatric sample. This index was called the Severity of Illness-Geriatric (SOI-G). SOI-G is a research scale designed to yield a single numerical score with interval properties that reflects severity of illness. The present research consisted of five sequentially linked studies. The first three studies were designed to address the first purpose while two additional studies dealt with the second aim of the present research, i.e. to collect initial reliability and validity data on SOI-G.

The development of SOI-G was accomplished by employing the methodology used in the development of SORDS (Baltzan et al., 1987). The first study involved a determination of the potential pool of items and the elimination of dimensions judged less relevant to geriatric health status. Items from SORDS were used on SOI-G where appropriate. SORDS items deemed inappropriate for a geriatric sample were either modified or removed. Lastly, disease conditions not present on SORDS but determined to be important in the assessment of severity of illness in the elderly were developed for SOI-G applying the methodology used in the development of SORDS.

The use of domain-specific items was supported by research that suggests that physicians' global ratings of health are less accurate and more susceptible to subjective interpretations than are domain-specific ratings (Parmelee et al., 1995). For instance, Parmelee et al. observed that the accuracy of physician ratings increased as specificity of the items increased and as the response alternatives were better defined. Lastly, many older persons have at least one chronic condition and many have multiple conditions (e.g., Besdine, 1997; Chenier, 1993), therefore the ability of a scale to tap the most commonly experienced diseases that pose a threat to life was an important consideration in the development of SOI-G.

Study 2 involved generating generally acceptable, objective criteria (based on standard procedures) for classifying the severity of the disease into the following categorizations (where possible and appropriate): (1) mild; (2) moderate; or, (3) severe. Study 2 also involved the development of a protocol

for the inclusion of disease severity ratings for dimensions not included by Study 1. Such a protocol involved the development of an open-ended item format (i.e., an "other" category) that would allow for an adjustment to the total SOI-G score seen for a geriatric patient based on the adjudged impact of the less frequently encountered conditions. At this stage, SOI-G was capable of assessing the severity of each particular disease, however assigning numbers such as 1, 2, and 3 to the mild, moderate and severe categorizations of a disease only resulted in an ordinal scale. Thus, the SOI-G was not yet able to yield a single score on an interval scale.

Study 3 involved applying the scaling methodology developed for the SORDS to the geriatric dimensions identified in Study 1 and the severity criteria developed in Study 2. Prior to this stage, the SOI-G was capable of assessing the severity of each particular disease along an ordinal-based dysfunction/severity scale. To express the severity levels of the various diseases/syndromes on a meaningful numerical scale, it was necessary to scale each severity rating of each disease/syndrome on a common underlying dimension of illness severity. The underlying illness severity dimension was defined as the degree of dysfunction and disability caused by the disease. For the purposes of this research, the scaling dimension ranged from 0, defined as the absence of a particular disease, up to 100, defined as death. Thus, all severity levels of all items were scaled to furnish a score that reflected the extent to which the condition posed a threat to life. Intermediate and increasing

values between 0 and 100 were defined by increasing levels of dysfunction and functional decline.

The decision to scale each item in order to improve the predictive ability was based on both clinical judgment and empirical research. For example, clinical judgment suggests that a moderate laceration does not pose the same degree of a threat to life that a moderate third degree burn would. Empirical research with the SORDS established the use of a scaling process as appropriate strategy. In addition, the present scaling process has been established as an acceptable practice in cases where the scale items are relatively heterogeneous (Streiner & Norman, 1995).

A panel of five physicians with specialized training and experience in geriatric medicine were asked to rate each level (i.e., mild, moderate, severe) of each disease or syndrome along an interval-based dysfunction/severity scale ranging from 0 to 100. It was felt that by scaling the disease items from SOI-G in this manner we could obtain specific numerical values which could then be summed across diseases and give a single numerical index of severity of illness. Ratings were completed using a variant of the Delphi method (to be described in more detail later in the document) that required panelists to first complete their ratings individually and then to complete the ratings as a group. Through discussion, the panel came to an agreement as to the final scale value to be assigned to each disease level. Any initial disagreements regarding values were resolved through collegial discussion. Study 3 also entailed a preliminary

assessment of the adequacy of the scaling process by demonstrating the internal consistency of the scaling judgements.

Studies 4 and 5 involved the initial application of the SOI-G to a small sample of geriatric participants. This was accomplished by carrying out two separate studies that allowed for the collection of some preliminary reliability and validity information. Since the SOI-G was intended as a measure of illness severity it was important at this early development and revision phase to sample participants who are known to be ill. Thus, a sampling strategy was adopted that selected a small sample of geriatric participants who had been identified as physically ill (i.e., who were admitted to a geriatric unit of a rehabilitation hospital). Study 4 investigated the extent to which SOI-G assessed variability in illness severity in a consistent, reliable manner for geriatric persons suffering from a variety of physical illnesses. Study 5 assessed the extent to which SOI-G scores showed anticipated relationships with other illness severity measures and with other psychosocial variables.

The methods used in later studies were somewhat dependent on earlier studies and therefore it was not possible to present the methods section apart from the results section. For economy of presentation as well as enhancement of clarity, a combined methods/results section will be presented next. The methods/results section will offer more detailed descriptions of each individual study including a description of the methods, results, and a brief discussion of each study. A longer discussion will follow at the end of the document.

4. Method and Results

The following section outlines the development of SOI-G in terms of the developmental sequence briefly described in the previous section.

4.1 Study One: Determine potential pool of items

Dr. Darryl Rolfson (MD, FRCP(C), FACP, Specialist in Geriatric Medicine, Northern Alberta Regional Geriatric Program, Glenrose Rehabilitation Hospital, Assistant Clinical Professor, Department of Medicine, University of Alberta, Edmonton, AB) was asked to provide a list of chronic health conditions, grouped by major systems, most commonly cited for geriatric populations. Conditions were included on the list if they represented a threat to life (mortality) or were expected to cause permanent or temporary impairment (morbidity). An exhaustive list of potential diseases was obviously impractical to scale, and thus some diseases were excluded on the basis of infrequent occurrence. Items were limited to chronic conditions to reduce the length of the scale. Including acute illnesses would have at least doubled the number of illnesses on SOI-G. Furthermore, most acute illness is viewed in the context of an underlying chronic illness (D. Rolfson, personal communication, December 5, 2000). In addition, chronic illness is the most common health situation in late life and therefore of most interest in gerontological research (Verbrugge & Jette, 1999). Deaths due to accidents and adverse effects (e.g., suicide, homicide) were not included.

Dr. Rolfson's original list of possible diseases involved 72 items. Items relating to psychiatric illnesses (i.e., depression, anxiety, psychosis, alcoholism and drug abuse, mental retardation) were removed. Psychiatric items were deliberately excluded as the SOI-G was intended to assess chronic health problems independent of psychological variables. Dr. Rolfson was asked to shorten the list in order to keep the length of the scale manageable (approximately 20 to 30 items). He later removed the infectious disease category because there were no illnesses which he felt met the criteria of being common and chronic in nature. For example, "urinary tract infections are common but rarely chronic ... If an infectious disease does arise, it would be well suited to the "other" category which you have been planning" (D. Rolfson, personal communication, January 2, 2001).

Although the number and nature of the items would evolve as the next two studies unfolded, at this point SOI-G consisted of 31 diseases and syndromes involving the following 10 categories: (1) geriatric syndromes; (2) neurologic; (3) respiratory; (4) cardiovascular; (5) hematologic; (6) endocrine; (7) oncology; (8) gastrointestinal; (9) musculoskeletal/immune; and, (10) renal/urologic. The conditions selected are listed in Table 4.1. While, technically, Table 4.1 presents "results", the sequential nature of the present investigation makes it necessary to at least briefly show the structure of SOI-G to aid in understanding the subsequent discussions.

Table 4.1 SOI-G Conditions

GERIATRIC SYNDROMES	NEUROLOGIC
Dementia Syndrome [1.]	Stroke
Urinary Incontinence	Parkinson's Disease
Malnutrition	Peripheral Neuropathy
Falls	CARDIOVASCULAR
Pressure Ulcers	Ischemic Heart Disease [3.]
Constipation	Heart Failure
Special Sensory Impairment [2.]	Peripheral Vascular Disease
RESPIRATORY	HEMATOLOGIC
COPD & Asthma	Anemia
ENDOCRINE	Thromboembolic Disease
Hypothyroidism	Myelodysplastic Syndrome
Diabetes Mellitis	GASTROINTESTINAL
Osteoporosis	Gastroesophageal Reflux Disease
ONCOLOGY	Cirrhosis
Multiple Myeloma	Peptic Ulcer Disease
Chronic Lymphocytic Leukemia	MUSCULOSKELETAL/IMMUNE
Lymphomas	Rheumatoid Arthritis
Solid Tumor	Osteoarthritis
RENAL/UROLOGIC	
Renal Failure	
Benign Prostatic Hypertrophy	

Note 1. Dementia divided into Dementia Syndrome and Alzheimer's Disease (Neurologic category) in Study 2.

Note 2. SSI changed to Vision or Hearing Impairment in Study 2.

Note 3. Ischemic Heart Disease changed to Coronary Heart Disease and Angina in Study 2, and to Angina only in Study 3.

The items selected were consistent with previous research that had identified these conditions as leading causes of death among persons 65 years of age and older. For example, Furner et al. (1997) concluded that diseases of the heart, malignant neoplasms, and cerebrovascular disease were the top three causes of death, accounting for approximately 70 percent of all deaths in persons 65 and older in the United States. The same three conditions were identified as the leading causes of death in the Statistics Canada (1997) data and accounted for a similar percentage of total deaths in Canada.

A decision was made to include dementia on the SOI-G scale because of its common occurrence among elderly persons and because of the influence of dementia on a person's level of occupational or social functioning. Estimates suggest that dementia affects between 2 percent and 8 percent of those between 65 and 75, 20 percent of those over 80, and 34.5 percent of those over 85 (National Advisory Council on Aging, 1996). According to the National Advisory Council on Aging, the dementia disorders are characterized by the development of multiple cognitive deficits. The dementia disorders share a common symptom presentation but are differentiated on etiology (American Psychiatric Association, 1994). In order to meet the DSM-IV criteria for diagnosis of dementia, the cognitive deficits associated with dementia must "be sufficiently severe to cause impairment in occupational or social functioning" (p. 134) thus meeting the criteria of impairment for inclusion on the SOI-G. In addition to causing impairment, Schneck, Reisberg, and Ferris (1982) report that dementia is associated with a decline in life expectancy. In Canada, 10,000 deaths per

year can be attributed directly to dementia (National Advisory Council on Aging, 1996).

Alzheimer's Disease (AD) is the most common form of dementia accounting for over half of the all diagnosed cases of dementia and 75.2 percent of all dementias in persons over the age of 85 (National Advisory Council on Aging, 1996). Some researchers identify Alzheimer's disease as the fourth or fifth most common cause of death in the elderly in Canada and the USA (Schneck et al., 1982). However, according to the Statistics Canada (1997) data, AD accounts for less than one percent of all causes of death. The discrepancy between the Statistics Canada data and the identification of AD as a leading cause of death among the elderly by some researchers likely reflects the practice of attributing the cause of death to the immediate cause of death, such as pneumonia or heart failure (Schneck et al., 1982).

In Canada, Chenier (1993) estimates that hip fractures related to osteoporosis result in death in 12 percent to 20 percent of cases and disability in up to 75 percent of surviving patients and therefore was included on SOI-G. On the basis of data from a USA sample 65 and over, Furner et al. (1997) reported that arthritis was the most frequently reported chronic condition (affecting 48 percent of persons). Similarly, Spar and La Rue (1997) reported the prevalence of arthritis in persons aged 65-74 to be 44 percent and 55 percent for those over 75. Thus, while osteoarthritis is not expected to cause death, it can lead to varying degrees of impairment and was included on SOI-G.

Malnutrition is a serious health problem among elderly persons. The International Food Information Council Foundation (IFICF) estimated that levels of malnutrition among older persons range from 15 to 50 percent, with particularly high levels found in institutional settings. Elderly persons are at higher risk of nutritional deficiencies because of declines in taste, smell, poor appetite, mobility difficulties that interfere with purchasing or preparing food, or the presence of a feeding tube (IFICF, 1999). Malnutrition can weaken a person's body and represents a threat to life and thus was included on SOI-G.

Anemia is a common and serious problem in the geriatric population and elderly persons are more prone to anemia than non-elderly persons (MedWorks Media, 1999). Although there are different types of anemia, the outcome for each type is the same, progressive disturbances in the muscular, nervous and gastrointestinal systems. Anemia is associated with fatigue, stresses on the heart, tingling or reduced feeling in the legs, poor balance, confusion and in severe cases, congestive heart failure (MedWorks Media, 1999).

Angina (the medical term for chest pain due to coronary heart disease) accounts for less than 1% of elderly deaths but represents a sign that someone is at risk of a heart attack and thus was also included.

In sum, the consistency of SOI-G items with conditions identified in the literature and Statistics Canada information lends support to the contention that the SOI-G accounts for a notable proportion of conditions expected to occur among the elderly, thus supporting the content validity of SOI-G.

4.2 Study Two: Develop Severity Criteria

For Study 2, Dr. Rolfson was asked to provide generally acceptable, objective criteria (based on standard procedures) for classifying the severity of the disease into the following categorizations (where possible and applicable): (1) mild; (2) moderate; or, (3) severe. As was done with SORDS, Dr. Rolfson was informed that some categorizations could be collapsed or simply be stated as absent/present. The determination of the appropriate number of categorizations per item was left to his clinical judgment. Several iterations of the severity criteria were required before the final criteria were decided upon. This process involved an ongoing collaboration with Dr. Rolfson, Dr. David Scott (dissertation supervisor), and the dissertation committee members.

Dr. Rolfson stated that his decision on the best categorizations for the first round of severity criteria was based on the premise that illness severity would be defined by its functional impact. He stated the following:

> In the world of Geriatric Medicine, function is everything. ... When I say function, I am speaking about the performance of activities of daily living. Function also includes the impact on interpersonal relationships, the pursuit of activities which are enjoyable and the overall quality of life. ...without some yardstick with which to compare the various illnesses, it will be impossible to quantify their severity in a parallel fashion (D. Rolfson, personal communication, November 28, 2000).

Ratings on the SORDS were based on the hypothetical case of an individual suffering from the disease that was between 40 and 45 years of age. The assumption in the development of SOI-G was that the hypothetical person suffering from the disease is between 70 and 75 years of age. Thus, any of the

illness severity criteria on SORDS that were inappropriate for a geriatric population were adapted for the SOI-G.

Dr. Rolfson reported that his original classification of stable, unstable, and myocardial infarction for ischemic heart disease were "really a combination of two criteria in SORDS - Coronary Heart Disease and Angina Pectoris. CHD refers to the burden of damage from previous ischemia and angina refers to ongoing symptoms of inadequate coronary blood flow." (D. Rolfson, personal communication, February 22, 2001). Dr. Rolfson agreed to remove ischemic heart disease and instead use the SORDS criteria for CHD and angina in its place. He indicated that the SORDS criteria for CHD represent the Canadian Cardiovascular Society Classification which is well recognized.

Conditions on the SOI-G identified in Study 1 that overlapped with items on SORDS included the following: (1) anemia; (2) angina; (3) diabetes; (4) osteoporosis; (5) renal failure; and, (6) coronary heart disease. The only item on SOI-G that reflected gender differences in the severity criteria was anemia (as with SORDS). At this stage of development, the severity criteria for anemia and angina remained as defined on SORDS. Renal failure was defined using only one of the three defining criteria used on SORDS (i.e., creatinine clearance problems). New severity criteria were developed for diabetes and osteoporosis.

The severity criteria remained largely unchanged from the initial draft for the following items: (1) urinary incontinence; (2) pressure ulcers; (3) constipation; (4) special sensory impairment; (5) benign prostatic hypertrophy; (6) thromboembolic disease; (7) myelodysplastic syndrome;

(8) diabetes; (9) osteoporosis; (10) Parkinson's Disease; (11) rheumatoid arthritis; (12) osteroarthritis; (13) multiple myeloma; (14) chronic lymphocytic leukemia; (15) lymphomas; (16) sold tumor; (17) COPD and asthma; and, (18) gastroesophageal reflux disease. Although the severity criteria for these items was largely unchanged, the criteria for some items were made more explicit. For example, a definition of "impairment in instrumental activities of daily living" was provided to indicate what IADL involved (e.g., "complex activities needed for independent living"). Another example was the change in "special sensory impairment" to "vision or hearing impairment" and the further explication of the criteria for "vision or hearing impairment" so that "correctable with aid" became "correctable with glasses or hearing aid".

Originally, dementia was accounted for by only one item. Given the prevalence of Alzheimer's Disease (AD) in the elderly, a decision was made to add a second dementia category to the neurologic category using the Global Deterioration Scale (GDS; Reisberg, Ferris, DeLeon, & Crook, 1982). The GDS is a widely used research scale that assesses the degree of cognitive and functional decline associated with advanced stages of AD and rates cognitive impairment along a 7-point scale (Caputo et al., 1998). Some research has discovered neurological evidence of a distinct pattern of progressive cortical, extrapyramidal, and pyramidal system dysfunction in AD that is associated with the late GDS stages (e.g., Franssen, Kluger, Torossian, & Reisberg, 1993).

Dr. Rolfson suggested that we maintain the original dementia item, in addition to the AD item, as:

> The Global Deterioration Scale developed and validated by Barry Reisberg is only valid when used for Alzheimer's disease (which comprises a little more than half of all dementias.) The scale will be inaccurate and misleading when applied to any other population. I know this because we use it on a daily basis and having the instrument underestimate or overestimate the stage of the dementia is the rule rather than the exception. Furthermore, ... most physicians would not be familiar with the tool. (D. Rolfson, personal communication, February 22, 2001).

Further clarification was requested from Dr. Rolfson on the difference between "occasional falls" and "frequent falls". In his clinical judgment, he indicated that it would be important to separate the senior who has had "bad luck" from the senior who has a problem with falling. It was decided that "occasional" would be defined as no more than once/year while "frequent" would involve more than once/year.

Dr. Rolfson was also asked for more objective criteria for malnutrition. He indicated that there were few scales for nutrition and of those in existence none were well validated and very few were well known. He suggested the use of the Subjective Global Assessment which grades nutritional risk into three categories: (1) well nourished; (2) moderately malnourished; and, (3) severely malnourished. He indicated that this scale is completed by a patient in combination with their doctor, nurse or therapist based on a history and physical examination. The history takes into account weight changes, dietary intake, gastrointestinal symptoms, functional capacity, and concurrent disease. The

physical exam includes loss of subcutaneous fat, muscle wasting, ankle edema, sacral edema, ascites, mucosal lesions, cutaneous lesion and hair changes.

It was recognized that some infrequently encountered conditions that elderly persons may experience may not be captured by the SOI-G but may be relevant to the overall severity of illness score. In order to adequately address this possibility, an open-ended item was developed that allowed for an adjustment to the total SOI-G score (see Appendix B).

At this stage, the SOI-G consisted of 33 items (not including the "other" category) and was capable of assessing the severity of each particular disease along an ordinal-based dysfunction/severity scale. See Appendix C for the version that was presented to the scaling panel in Study 3.

4.3 Study Three: Scaling of new items

Stage 3 employed a variant of the Delphi method, which is a systematic method for gathering information from groups of people who have insight into a particular area of interest (Clayton, 1997; Boberg & Morris-Khoo, 1992; Fish & Bushby, 1996). In particular, the Delphi technique is appropriate in research situations requiring consensus of opinion about a particular area (Fish & Bushby, 1996). The Delphi approach originated in 1953 at the RAND Corporation where it was used to obtain general agreement from experts on defense and military matters (Boberg & Morris-Khoo, 1992; Clayton 1997; Fish & Bushby, 1996). Today this approach continues to find application in psychological, sociological, political science, and medical research arenas (Gallagher, Hares, Spencer, Bradshaw, & Webb, 1993). Although the Delphi

method is a qualitative method, its end product can be analyzed quantitatively (e.g., mean or median).

The Delphi method elicits the perceptions and judgments of the panelists. These opinions are further refined during subsequent rounds. The goal of these reviews is to achieve a blending of diverse opinions and ideas about a particular topic (Boberg & Morris-Khoo, 1992; Gowan & McNichols, 1993). Only two rounds were conducted in the present study as it has been suggested that more than two rounds tends to show little change and increases the likelihood of regression to the mean (Fish & Busby, 1996). Further, based on the experience of researchers in the development of SORDS, the method used for round two was altered from the standard Delphi method. The variation of the Delphi approach used in the present study involved convening a group meeting for the second round rather than conducting a second mailing in which panelists would be asked to reevaluate their original ratings.

4.3.1 Pilot Study

Prior to the actual scaling process, a pilot study was conducted to determine the approximate length of time to complete the task as well as to determine what information was useful to raters and which of two possible methods was preferred by the raters. A Clinical Nurse Educator and a Geriatrician at a geriatric assessment unit at a Saskatoon hospital agreed to participate as pilot raters. Unfortunately, the geriatrician failed to complete the task within the allotted time period and thus only the information provided by the nurse was used. The nurse was paid for her efforts.

The results of the pilot study lead to the inclusion of the following for the actual scaling procedure: (1) a cover letter outlining the procedure; (2) a line version of the SOI-G scaling procedure; (3) a copy of SORDS and the disability scale used in the development of SORDS; and, (4) brief background information on the development of SOI-G (see Appendix D). Dr. Rolfson also attached a cover letter to this package.

4.3.2 Scaling Process: Method

All geriatric specialists at the Northern Alberta Regional Geriatric (NARG) program at the Glenrose Rehabilitation Hospital were invited to participate. The first five volunteers were selected for the panel. Geriatric specialists were selected because of their specialized training in geriatric medicine and because of their involvement with geriatric patients on a day to day basis. The educational background of the five final panel members represented the two training pathways available to physicians specializing in geriatric medicine. Two of the panelists completed their training in Internal Medicine with a two year subspecialty in Geriatric Medicine while the remaining three completed their training in Family Medicine with a six month certification program in care of the elderly. Each panelist was provided with financial remuneration in the amount of a $300 honorarium.

Each panelist was provided with the scaling material, instructions, and a short background summary on the development of SOI-G. To aid the panelists in their judgements of the disability to be associated with the various disease levels, all were given a copy of SORDS as well as the sample descriptors of the

type and degree of disability to be associated with the scale values of 20, 40, 60, and 80 chosen from a larger group of disabilities scaled on perceived unpleasantness using hospital patients as judges (Sprecker, 1987). Sprecker indicated that the disability scale incorporated three aspects of functional ability thought to impacted by illness, namely mobility (i.e., how able is the person able to move about), social activity (i.e., how well can they interact with others), and physical activity (i.e., how well can they care for themselves).

Providing behavioural exemplars of severity of illness defined in terms of functional limitation is also in keeping with the Disablement model developed by Verbrugge and Jette (1994) described in section 2.2. Within the Disablement model, pathology leads to impairment which contributes to functional limitations in basic physical and mental actions. Dysfunction or physical restrictions as defined by the disability scale was intended to be used to illustrate how severely ill a particular individual might be at each point on the 0 to 100 scale. With increasing illness severity, or threat to life due to illness, an individual is expected to experience greater limitation with the ultimate functional limitation defined as death. The use of functional exemplars is also consistent with the process used by researchers in the development of SORDS and with the feedback provided by the panel of medical experts (outlined in the next section) regarding the validity and usefulness of incorporating functional indicators.

Members were requested to go over SOI-G and make independent ratings (i.e., without discussion with other group members) and comments prior to the scheduled panel meeting. Judges were asked to scale the level of

disability associated with the severity levels of the diseases under the following assumptions: (1) that the disease being rated was the only disease present; (2) that the person suffering from the disease was between 70 and 75 years of age; and, (3) that the disability rating to be assigned to a particular disease severity level should be that associated with the approximate middle of the anticipated range of increasing dysfunction that would result from that severity level.

Assumption one was included even though the clinical and research literature suggests that the impact of a particular disease would be influenced by the presence of other diseases. However, rating each disease level in combination with the other diseases on SOI-G would be an impossibly large scaling task. Given that the current study was a preliminary stage in the development of a geriatric severity of illness measure, the consideration of the interactive role of diseases was viewed as potentially appropriate for future testing and development of the SOI-G.

Assumption two was included to focus the scaling judgments on an age group in which confounding factors due to age-related disease impact would be hopefully minimized. It was intended to make the scaled item as broadly applicable to individuals with that condition as was possible. It was recognized that some diseases could vary in the severity of impact as a function of the age of an elderly person. However, such variability will be assumed negligible in the initial development of SOI-G.

Assumption three was necessary to allow the panelists to assign a single number that will represent the disability to be associated with a range of symptom severity such as "moderate to severe chronic heart disease".

Two weeks after initial receipt of the scaling materials, the initial panel results were returned to the primary investigator and all item scores were summarized. This summary was then distributed to each panel member in advance of the group meeting (see Appendix E). Each panel member was presented with a list of the SOI-G items listed in order of least agreement among raters to greatest agreement among raters prior to the meeting. The members then met as a group and this writer lead the discussion. Comments, criticisms, and evaluations regarding the SOI-G instrument and scaling method were also discussed. This process took approximately two hours.

4.3.3 Scaling Process: Results

Each of the five panelists completed the scaling task independent of the other raters (as described in the previous section). An analysis of inter-rater agreement between the individual panel members prior to the scaling meeting was conducted using the Coefficient alpha. The reliability estimate of the individual ratings (n = 5) was .96, indicating strong initial agreement between raters. This value is interpreted as a low bound estimate as panel members were allowed to revise their ratings at the panel meeting. Thus, it is expected that there was even greater agreement by the end of the scaling panel meeting.

Raters agreed to accept the median value 69% of the time. Although the physicians selected the median in the majority of case, panel members still

participated fully in a discussion aimed at achieving consensus for each individual item. Furthermore, their discussion often included a consideration of the ratings previously assigned. For example, a moderate level of gastroesophageal reflux disease (GERD) was deemed to be less severe than a previously assigned rating of 30 given to moderate chronic lymphocytic leukemia. It was clear from the discussion that they were not trying to equate the "moderate" rating of the two items but rather were focused on the numerical quality of a "30". Thus, GERD was rated as 27 (the median had been 26.8). In this example, it appears as if the panelists simply selected the median, however, the assigned numerical value does not reveal the full extent of the discussion and comparisons behind the decision to assign that value.

Disclosures by panel members revealed no difficulty applying the disability scale (i.e., the 0 to 100 scale) and they consistently used the behavioural descriptors as guides in the determination of severity criteria for each level of each disease. The following are some examples of how the disability scale was applied by the panel members when considering the severity of each disease/syndrome state. For example, a score of 80 was interpreted as "starting to die at this point" and therefore severe lymphomas was rated "88" as they considered this to be the middle of the anticipated range and "close to death". Another example, was the application of "can't participate fully" when considering values over 20. Even though mild vision/hearing impairment was "correctable with glasses or hearing aid", panelists agreed that there are "some things you can't do even with glasses or hearing aids". Furthermore, it

was clear from the values assigned that raters appreciated that, in some instances, there was an acceleration in disease severity from one level to the next. For example, myelodysplastic syndrome was rated as 15 for mild, 45 for moderate, and 90 for severe. Thus, it appears that the obtained SOI-G values represent an objectively measurable concept of severity of illness, namely threat to life.

Panel members experienced no difficulty achieving consensus and the scaling task was completed in less time than planned for and therefore, no items were left unscaled. It was reported that all panel members knew each other prior to the scaling meeting and had a prior working relationship with one another. Thus, the ease with which panel members were able to work together to achieve agreement likely reflected the collaborative nature of their day to day clinical work with one another. No one panel member was observed to dominate or unduly influence the decisions of other panel members. Overall, the fact that the panel achieved consensus for every item on the SOI-G seems to offer support for the face validity of the severity criteria developed in Study 2.

Anecdotal evidence that further supports the validity of SOI-G includes the observed similarity in scale values between items that appear on both the SOI-G and SORDS. The scale values assigned to each level of each disease were quite similar for many of the shared items although the values were consistently higher for SOI-G. For example, moderate angina on SORDS was rated at 35 on SORDS and 50 on SOI-G or, severe angina was 80 on SORDS and 91 on SOI-G. The observed higher values for similar items on SOI-G was

not unexpected given the fact that the SOI-G values were rated for a person 40 to 45 years of age while the hypothetical person being rated on the SOI-G was assumed to be between 70 and 75 years of age. One would expect the threat to life to be greater for a frail elderly person relative to their younger counterpart with a similar level of the same disease.

Consistently throughout the panel discussion, the comments by raters illustrated a shared body of knowledge that went beyond the explicitly stated severity descriptors. For example, even though multiple myeloma was described in stages, the discussion suggested that the raters expertise in geriatric medicine led them to expect similar clinical outcomes associated with a particular stage. For instance, the panelists agreed that severe multiple myeloma (stage 3) "restricted life because of fracture risk, fatigue". Similarly, raters agreed that mild chronic lymphocytic leukemia meant the person would be "more prone to infections". Lastly, severity ratings for mild benign prostatic hypertrophy involved the shared understanding that a person would likely be "up at night to go to the washroom and therefore losing sleep". Thus, the discussion suggested a shared level of understanding of the threat to life posed by each level of each disease/syndrome.

Where there was disagreement among raters, certain patterns emerged that suggested that the disagreement reflected: (1) greater disparity in the disease being rated; (2) the use of different internal frames of reference; (3) differences in work-related experience; or (4) difficulty assigning severity ratings to numerical laboratory findings.

The greatest disagreement seemed to reside with oncology items and least variability within the neurological items. Raters confirmed that there was more variability with cancer than neurological conditions. For example, raters reported some difficulty assigning a value for solid tumor because severity depends on the type of solid tumor and the degree of metastasis.

With respect to different internal frames of reference, two of the five geriatricians indicated that they derived their ratings based on the hypothetical case of a community dwelling elderly person with the illness while two other raters were making comparisons based on the hypothetical case of the person being admitted to a geriatric rehabilitation hospital. The community dwelling elderly person was thought to be a "healthy person with one disease". In this case, the presence of a mild illness would have a noticeable impact on their ability to perform their usual activities. In contrast, mild illness might have a less noticeable impact in the hypothetical case of a frail elderly person in a rehabilitation hospital as the mild illness may become "lost" among other difficulties the person may be experiencing. The group consensus was to view these two hypothetical situations as the end points of the anticipated range of increasing dysfunction that would result form that severity level.

In some cases, the discussion revealed that differences between raters reflected greater or lesser degrees of experience with the conditions contained on SOI-G.

One area that panelists experienced the most difficulty involved assigning severity of illness ratings to laboratory findings. There was consensus among

the raters that it was far easier to assign severity values to functional indicators than it was to laboratory results. Comments by panelists suggested that attempting to reduce the SOI-G items to numerical terms reduced the meaningfulness of the item and thus risked destroying the core nature of the disease being scaled. Another panelist made the following comment:

> For those diagnostic entities which are discrete and involve a single body system, measurement and scaling using explicit and objective criteria are easy to identify (e.g., anemia, hypothyroidism). On the under hand, as soon as you deal with a syndrome such as the geriatric syndromes, function becomes the means of rating the severity of the illness syndrome. I believe that many geriatric syndromes are not mechanistic, but are rather 'gestalt' - the whole is more that the sum of the parts. Activities of daily living represent an excellent touchstone for those syndromes which are complex and best understood in the gestalt.

The comments by panel members mirror comments made by Rockwood et al. (2000) in their discussion of the conceptualization and measurement of frailty, "there is reason to be suspicious of the reductionist approach in explaining individual states arising from the interaction of complex systems" (p. 300).

Some of the items were found to be unscalable in the form presented and thus during the panel meeting some changes were made to individual items. On the basis of the previously discussed difficulty associated with scaling laboratory results, behavioural descriptors were added to anemia and renal failure. The panel agreed that the severity criteria for Parkinson's Disease should be defined in terms of the Hoehn and Yahr Classification system (Hoehn & Yahr, 1967). Coronary heart disease was removed because it was deemed unnecessary given the inclusion of heart failure and angina. Thus, the final SOI-G scale used in Studies 4 and 5 contained 32 diseases/syndromes (see

Appendix F). Of course, to avoid potentially biasing raters, the SOI-G used in Studies 4 and 5 did not include the weighted values listed in Appendix F.

There was also some discussion about whether or not to include two dementia items. The medical experts agreed that the addition of the Alzheimer's Disease item was important because the "incipient AD" category was not captured by the other dementia item on SOI-G and because of the importance of differentiating between dementia syndromes and AD. Raters agreed that incipient AD has an impact on the person's ability to function and therefore it was important to capture this on the SOI-G. However, the panelists questioned whether GDS information would be routinely recorded in medical files. Interestingly, there was group convergence after the initial round with respect to severity ratings for the two dementia items. The medians for mild, moderate, and severe AD (44, 65, and 95, respectively) closely matched the medians for mild, moderate, and severe dementia (40, 68, and 90, respectively).

SOI-G did not include a pain item. Although past research has been criticized for not including a pain scale (e.g., SIP; De-Bruin et al., 1992), pain was considered a subjective health perception and therefore more appropriately measured by self-report health status measures. In validation studies with SORDS (Britton, 1985), the panel noted "that the degree of pain which might accompany a disease would determine the level of incapacitation". With SORDS, this issue was dealt with by assuming a constant level of pain across the various levels of each disease. However, during the development of SOI-G, panel members remarked that pain was already reflected in each disease state.

One area targeted for future development was the inclusion of a time element as part of the severity ratings. Specifically, raters noted that with some diseases (e.g., anemia, heart disease) the severity of illness is also a function of the time it takes for the disease to develop. For example, a certain hemoglobin level that develops within months might be considered mild while the same hemoglobin level developing within hours to days would be more severe.

In summary, the results of Studies 1 to 3 support the content and face validity of SOI-G. Typically, the content validity of an instrument is not formally assessed but rather the face validity or clinical credibility of the instrument is inferred from the comments of experts (D. Scott, personal communication, August 12, 2001). With respect to the development of SOI-G, a geriatric expert generated a sampling of items from a larger pool thought to reflect the chronic diseases most likely to cause morbidity or mortality in a geriatric population. The items selected were consistent with the most common causes of death among older persons listed in the literature and found in Statistics Canada information. Further, the panel of geriatricians made few changes to the pool of items presented to them suggesting the selected items and respective severity criteria reflected the aim of the instrument (i.e., to assess geriatric severity of illness). Thus, the level of agreement with previous research and Statistics Canada information along with the general acceptance of SOI-G by geriatric specialists supports the content validity of the instrument.

4.4 Studies Four and Five: Reliability and Validity of SOI-G

The overall design for the final two studies was to collect data on the reliability and validity of SOI-G by assessing the extent to which SOI-G: (1) assessed variability in illness severity in a consistent, reliable manner for geriatric persons suffering from a variety of physical illnesses; and (2) demonstrated anticipated relationships with other illness severity measures and with other psychosocial variables. Unfortunately, difficulties associated with rater error limited the results of Study 4 and the original intention of Study 5 was unrealized due to a small sample size (n = 13). The following sections will illustrate that although the results were not definitive they were encouraging.

The samples for Study 4 and 5 were medically diverse patients from Northern Alberta Regional Geriatric (NARG) program at Glenrose Rehabilitation Hospital (GRH), Edmonton, Alberta. Study 4 examined patient archival data while Study 5 involved patients currently receiving rehabilitation services at NARG. Two nurses working on the geriatric units of NARG were hired for Study 4 to collect information using the SOI-G, CIRS-G, and global severity ratings (before and after use of the other scales). Study 5 patients were administered the BDI-II, the SF-36, and a demographic questionnaire by the primary investigator while SOI-G scores were obtained from each patient's medical file by one of the nurses trained for Study 4.

Ethical approval from the University of Saskatchewan Behavioural Sciences Ethics Committee, the Health Research Ethics Board-Panel B (HRDB-B) at the University of Alberta, as well as administrative approval from the

rehabilitation hospital was obtained prior to gathering data on any patients included in Study 4 and 5.

Reliability estimates of health status measures typically rely on inter-rater reliability correlation coefficients (i.e., the degree of relationship between two sets of scores involving the same instrument). In Study 4, inter-rater reliability was measured by a Pearson product-moment correlation of scores independently obtained from the nurses who rated the patient files. The Pearson correlation was appropriate in this instance as the SOI-G variables are at an interval level of measurement (Allen & Yen, 1979). The use of Coefficient alpha, while appropriate for determining the reliability of the scaling process, was not appropriate given the intentionally heterogeneous items on the SOI-G.

A major problem facing researchers attempting to measure overall severity of illness is the lack of generally accepted valid measures of illness severity by which newly developed measures can be compared for purposes of establishing their validity (Hall et al., 1989). As discussed earlier in the review of the literature, the CIRS-G is another geriatric severity of illness instrument, however it does not represent a "gold standard" for comparison purposes with SOI-G for three reasons. First, there are important differences with respect to the level of measurement between the CIRS-G and SOI-G. CIRS-G measures severity of illness on an ordinal scale (i.e., higher numbers are given to individuals who are more severely ill). In comparison, SOI-G assesses severity of illness at the interval level of measurement by locating each level of each disease on the same underlying threat to life scale. Interestingly, the

methodology used to develop SOI-G in Study 3 was consistent with comments made by the authors of the CIRS-G, namely that CIRS-G could be improved by the assignment of weights to equate the items (Miller et al., 1992).

A second difficulty with CIRS-G involves the scoring criteria. The developers of the original CIRS (Linn et al., 1968) as well as the authors that adapted CIRS for use with a geriatric population (Miller et al., 1992) both recommend summing across the thirteen items to obtain a total pathology score. This is problematic as it is not appropriate to combine ordinal scores by summation (Siegel, 1956). While larger numbers indicate more of the property being measured, equal intervals between numbers on an ordinal scale cannot be assumed and therefore there is not a one-to-one correspondence between the properties of an ordinal scale and operations of arithmetic (Allen & Yen, 1979; Siegel, 1956). For example, the summation of the ordinal values on CIRS-G items incorrectly assumes that three mild ("current mild problem or past significant problem") diseases are equivalent to one severe ("immediate treatment required/end organ failure/severe impairment function") disease. In comparison, total SOI-G scores are derived by summing interval values and thus the total score represents the total severity of illness for a patient.

The third reason that CIRS-G does not represent a gold standard for comparison in the present study is the restriction of SOI-G items to physical illnesses. In comparison, the CIRS-G combines psychiatric illness and physical illness in the total score. The present writer was unaware of any other

physiologically based, geriatric severity of illness measures that assessed total patient severity using an interval scale.

As no gold standard for the assessment of illness severity in a geriatric population was available, convergent validity was assessed by using measures aimed at measuring variables related, but not identical to, illness severity as measured by SOI-G. In Study 4, Pearson correlations were computed between the total scores for SOI-G and the CIRS-G as well as between the SOI-G and the global severity ratings (0 to 100 scale). Comparing SOI-G scores with scores on CIRS-G and the global severity ratings assessed the extent to which SOI-G possessed convergent validity. Convergent validity would be supported if SOI-G correlated with the CIRS-G and the global severity ratings.

It is acknowledged that the analyses described above involving CIRS-G failed to conform to the necessary conditions for the use of parametric tests due to the ordinal nature of the scale. Although technically these analyses were not appropriate, they were carried out to mimic analyses used in previous research with CIRS and CIRS-G (Naughton, Saltzman, Priore, Reedy, & Mylotte, 1999; Miller et al., 1991; Miller et al., 1992; Miller et al., 1996). The primary rationale, however, for performing these analyses was to allow for a full exploration of the available data in order to glean as much information as possible about the SOI-G. Thus, these analyses were performed with the full awareness that any conclusions derived from this data should be considered with caution.

In Study 5, the intention was to compute Pearson correlations between SOI-G and the three composite scores of the Medical Outcomes Study (MOS)

SF-36 Health Survey (SF-36). The Physical Health Composite (PHC) consisted of the following subscales: (1) limitations in physical activities because of health problems; (2) role disability due to physical health problems; (3) bodily pain; and (4) general health perceptions. The Mental Health Composite (MHC) consisted of the following subscales: (1) vitality (energy and fatigue); (2) social functioning; (3) role disability due to emotional health problems; and (4) general mental health (psychological distress and well-being; Ware & Sherbourne, 1992). The General Health Composite (GHC) is a combination of the PHC and MHC. Comparing SOI-G scores with scores on the SF-36 (a less disease specific, self report health measure) was intended to assess the extent to which SOI-G possessed convergent and discriminant validity.

Support for convergent validity of SOI-G would be enhanced if SOI-G correlated with the PHC (i.e., perceived difficulties associated with physical health problems). Discriminant validity means that the instrument being studied should demonstrate a low correlation with measures that are not thought to be related to the area of interest. Discriminant validity would be supported if low correlations were observed between the SOI-G and factors assumed unrelated to the presence of a particular disease, such as perceived mental health problems as assessed by MHC. Thus, given the stronger psychosocial component of the MHC, correlations between SOI-G and the PHC were expected to be higher than the correlations between SOI-G and the MHC.

Construct validity refers to the extent to which a test measures the theoretical construct it was designed to assess (Streiner & Norman, 1995). One

way in which the construct validity of the SOI-G was assessed involved making predictions about how SOI-G scores were expected to perform in a particular situation. If these predictions were supported by the data then the process of establishing construct validity of the SOI-G would be improved (Allen & Yen, 1979; Streiner & Norman, 1995). This was explored in Study 4 by examining the relationship between SOI-G scores and discharge outcome for patients (e.g., discharged home versus to long term care). In Study 5, the intention was to explore the issue of construct validity by assessing the relationship between the SF-36, the Beck Depression Inventory-2nd Edition (BDI-II), and SOI-G. On the basis of other research that compared a similar severity of illness measure (i.e., SORDS) with the SF-36 and BDI-II (Alexander, 2001), the magnitude of the relationship between SOI-G and the SF-36 was expected to be in the low to moderate range (r = .3 to .5, respectively). Unfortunately, Study 5 only involved data from thirteen individuals, thus making any conclusions problematic.

The following sections will present analyses of the data gathered for Study 4 and Study 5 at GRH by two nurses and the primary investigator. A brief summary of the approach used to assess the reliability and validity of SOI-G will be presented separately for each study. This will be followed by description of participants, measures, and procedure.

4.4.1 Study Four: Archival Review

Two nurses who were familiar with the measures but who were blind to the study hypotheses collected data for Study 4. In Study 4, the SOI-G and the CIRS-G were retrospectively applied to medical charts from the NARG program.

Study 4 was originally intended to investigate the following hypotheses:

(1) Higher SOI-G scores would be related to greater mortality rates. Ultimately, it became clear that this particular hypothesis could not be explored as there were no recorded deaths in the files reviewed. The raters explained that death is an uncommon occurrence on the geriatric units at this particular hospital because severely ill patients are transferred elsewhere.

It was assumed that patients who required greater levels of care (e.g., long term care) at discharge were likely to be more seriously ill and therefore would have higher SOI-G scores than those patients who were healthier. Therefore, discharge outcome was used in place of mortality and the relationship between SOI-G scores and discharge outcome was investigated. The data were combined to form three groups reflecting decreasing levels of independent living and increasing levels of required care. The first group consisted of patients discharged home or home with home care (Home group). The second group (Semi-Independent) consisted of patients discharged to the community but not able to live independently (e.g., lodges). The last group (LTC) consisted of patients discharged to long-term care facilities.

The Semi-Independent group was not expected to differ significantly from either the Home group or the LTC group with respect to severity of illness. However, the Home and LTC groups were expected to differ significantly on severity of illness. More specifically, the Home group was expected to be less severely ill relative to the LTC group. Given the greater sensitivity associated with the interval nature of SOI-G relative to the ordinal nature of CIRS-G, it was

expected that SOI-G would demonstrate a stronger relationship with outcome than would CIRS-G; and,

(2) a significant relationship was expected between SOI-G and CIRS-G as well as between SOI-G and Global Severity Ratings (made before reviewing the medical chart and again after reviewing the medical chart).

All data analyses were carried out using SPSS for Windows, version 9.

4.4.1.1 Study 4 Participants

Archival data was collected on 61 patients. Description information on the study participants will be outlined in detail in the results section.

4.4.1.2 Study 4 Measures

The nurses completed the following questionnaires: (a) SOI-G (see Appendix F); (b) CIRS-G (see Appendix G), and (c) demographic questionnaire (see Appendix H). Due to recent changes in the type of patient information collected (i.e., only collecting information deemed essential or necessary to inform treatment) ethnicity information was unavailable and marital status data was missing for 12 patients. Also, data on the living situation prior to admission missing for 40 patients. The order of presentation of the SOI-G and the CIRS-G was randomized. The raters were also asked to rate each patient on a 0 to 100 scale (see Appendix I) that assessed threat to life based on a overall consideration after their initial review of the file. They were asked to provide this global severity rating (GSR) before (GSR-Bef) and after (GSR-Aft) completion of the SOI-G and CIRS-G.

4.4.1.3 Procedure

Personnel from the Clinical Records department were instructed to retrieve 100 medical files for NARG patients (not currently receiving treatment) for the time period of June 2000 to June 2001. The intention was to have raters exclude medical charts that did not include information regarding patient outcome or that contained file ambiguities or were unclear, therefore an additional 20 files were prepared in the event that some files had to be discarded. In total, 120 files were prepared for review by the nurses.

Immediately prior to the actual data collection, the primary investigator familiarized the nurses with SOI-G. For example, it was explained that there were two dementia items, one intended for AD only and the other dementia item was for all dementias other than AD. The nurses were instructed to collect data on the patients most recent admission using the discharge summary, if possible, and the written instructions on SOI-G directed raters to collect information based on the patient's most recent test results. The primary investigator was on site to answer any questions that arose during the course of data collection.

After reviewing approximately ten files, the nurses reported that the discharge summary data was insufficient for the purposes of completing the SOI-G for most files and therefore the raters were instructed to extract information on the most recent admission from the clinical chart notes. This procedure was more time consuming than originally planned for and therefore, due to time and budgetary constraints, fewer files were reviewed than originally proposed. In total, 61 files were reviewed. The files were randomly stacked in

the cubicles where they were reviewed so the actual chronological order of the files reviewed is unknown. No medical charts had to be excluded because of missing patient outcome data, file ambiguities, or discrepancies.

The two raters collected information from the same files independently (for inter-rater reliability purposes) and thus were blind to each other's findings. Both raters were interviewed together immediately following the archival review and again separately approximately one month later. The results of these interviews will be presented as part of the results section. Each nurse was provided with financial remuneration ($500 each).

In order to ensure confidentiality, names did not appear on the measures and patient files were identified by number only. Results were analyzed and reported in group form so that no individual person could be identified. All data will be stored in a locked filing cabinet for a minimum of five years.

4.4.2 Study 4 Results

The data for this study was archival and obtained though the use of charted data. Study 4 results will be presented in the following order: (1) a descriptive analysis of the basic data; (2) reliability analysis; (3) age analysis; (4) comparison of SOI-G and CIRS-G with outcome; (5) a comparison of SOI-G, CIRS-G, GSR-Bef, and GSR-Aft; and, (6) a qualitative analysis of the comments made by the two raters on the use of SOI-G. Results from the validity and reliability analyses will be described with reference to the unique hypotheses.

4.4.2.1 Descriptive analysis of the basic data

The present analyses will begin by presenting descriptive information about the participants. The descriptive data included information on the following variables: (1) age and marital status of participants; (2) gender; (3) discharge outcome; and, (4) medical conditions experienced by patients. Lastly, means and standard deviations are presented for all measures.

The data sample consisted of 61 cases (22 males, 36 females, gender was missing for 3 cases). As shown in Table 4.2, the mean age of the participants was 79.1 years (SD = 7.7 years) and the participants ranged in age from 61 years to 95 years. The majority of participants were widowed (38 percent) while 29 percent were married or living common-law, 10 percent were single or never married, and 3 percent were divorced (information on marital status was missing for 20 percent of participants).

Table 4.2 Age and marital status of Study 4 (archival review) sample

Characteristic	Males	Females	Total
Mean age in years (n [1], SD)	75.7 (22, 8.9)	81.2 (35, 6.1)	79.1 (57, 7.7)
Range	61 to 93	69 to 95	61 to 95
Marital Status			
Single/Never Married	3	3	6
Married/CL	8	10	18
Divorced	2	0	2
Widowed	5	18	23
Missing [2]	4	5	12

Note 1. Age data missing for females n = 1; both age & gender data missing for n = 3

Note 2. Of the 12 files missing marital status data, gender information was also missing for n = 3

As show in Table 4.3, approximately 46 percent of participants were discharged home or home with home care, 20 percent returned to a lodge or private care home setting, and 23 percent were transferred to long-term care (LTC) or to emergency (1 case). The remaining 12 percent of participants were discharged against physician's advice, transferred to a psychiatric hospital, or the outcome data was not recorded (7 percent).

Table 4.3 Patient outcome for Study 4 participants

Outcome	Males	Females	Total
Home	5	6	11
Home with Home Care	7	10	17
Lodge or Private Care	4	8	12
Long Term Care	6	7	13
Transferred to Emergency	0	1	1
Psychiatric Facility	0	1	1
Discharge against advice	0	2	2
Missing (gender unknown n = 3)	-	1	4

Dementia was the most frequently reported disease/syndrome by both raters (59% and 66% of all patients). Other frequently encountered items included urinary incontinence, constipation, AD, stroke, COPD and asthma, anemia, osteoporosis, osteoarthritis, and renal failure. The oncology items (multiple myeloma, chronic lymphocytic leukemia, lymphomas, and solid tumor) were the least frequently reported conditions (0% to 7%) which is not unexpected given the rehabilitative nature of GRH and the fact that oncology patients are more likely to be treated at a different facility (e.g., Cross Cancer Institute) in Edmonton. At times, the frequencies of diseases/syndromes reported by raters was variable. For example, rater 1 indicated that vision or hearing impairment was present for 12 percent of the patients while rater 2 endorsed this item for 88 percent of the patients. The issue of raters endorsing items with varying frequency will be explored more fully in later analyses.

Table 4.4 contains the means and standard deviations for each rater for each of the four severity ratings (i.e., SOI-G, CIRS-G, GSR-Bef, GSR-Aft). In addition, the total SOI-G score is presented with the "other" category and without the "other" category that was developed to include diseases not listed on SOI-G. The maximum total score possible for SOI-G was 2,512 (not including the "other" category), CIRS-G was 52, and GSR-Bef and GSR-Aft was 100.

Table 4.4 Means and Standard Deviations for all scales

	Rater 1			Rater 2		
	Mean	(SD)	Range	Mean	(SD)	Range
Total SOI-G with "other"	136.6	(78.9)	17 to 391	191.2	(81.0)	76 to 412
Total SOI-G without "other"	122.2	(79.3)	0 to 391	159.1	(82.1)	36 to 383
CIRS-G	8.0	(3.4)	2 to 17	5.5	(2.4)	2 to 13
GSR-Bef*	55.7	(8.3)	40 to 60	48.9	(10.7)	20 to 60
GSR-Aft*	57.3	(7.8)	40 to 80	48.5	(10.6)	20 to 60

* GSR-Bef = Global Severity Rating Before
GSR-Aft = Global Severity Rating After

4.4.2.2 Reliability analysis

Reliability analyses were performed for 61 patients by two independent raters for the SOI-G (SOI-G1, SOI-G2), CIRS-G (CIRS-G1, CIRS-G2), and the global severity ratings completed before (GSR-Bef1, GSR-Bef2) and after completing the other instruments (GSR-Aft1, GSR-Aft2). Inter-rater reliability of was assessed using a Pearson product-moment correlation.

The reliability estimates for SOI-G and CIRS-G are presented in Table 4.5. The correlations for the SOI-G were initially conducted for the total score including the "other" category. It was expected that reliability estimates would be high between the two raters, however, this was not the case. The correlation between the two raters (SOI-G1 and SOI-G2) was significant but low ($r = .38$, $p < .01$). An inspection of the raw data revealed that rater 2 used the "other" category more often than rater 1, recording three times as many items. It was suspected that the variability between the two raters in their use of the "other" category may have accounted for the low agreement between the two raters. Thus, the correlation between the raters was re-calculated without the "other" category as part of the total score. This led to a somewhat greater degree of agreement between raters but the correlations remained lower than anticipated ($r = .46$, $p < .01$). The correlation between rater 1 ($n = 60$) and rater 2 ($n = 61$) for CIRS-G total scores was not significant, however, the total number of categories endorsed by raters 1 and 2 on CIRS-G (CIRS-GCat1, CIRS-GCat2) was significant related ($r = .41$, $p < .01$).

Table 4.5 Inter-rater reliability correlations for SOI-G and CIRS-G

	SOI-G1 without "other"	SOI-G1 with "other"	CIRS-G1	CIRS-G Cat1
SOI-G2 without "other"	0.46**	-	-	-
SOI-G2 with 'other	-	0.38**	-	-
CIRS-G2	-	-	0.21 (NS)	-
CIRS-GCat2	-	-	-	0.41**

Note. All correlations are based on n = 61

* Correlations significant at the .05 level

** Correlations significant at the .01 level

As suggested by Miller et al. (1992), the following scores were also computed for CIRS-G: (1) Severity Score (CIRS-G total/CIRS-GCat); (2) total number of categories at the level-3 (severe) severity; and, (3) total number of categories at the level-4 (extremely severe) severity. Correlations were then computed between raters for these scores. Level-3 severity ratings were reported for 7 of 61 cases for rater 1 and 10 of 61 cases for rater 2. No level-4 values endorsed by either rater. The correlations between the raters for the Severity Score and the categories at the level-3 severity were not significant.

Lastly, the two raters had been asked to give their global impression of the patient's severity of illness using the 0 to 100 disability scale after reviewing the file but before using the SOI-G or CIRS-G (GSR-Bef1 and GSR-Bef2). The reliability estimates involving these two ratings are presented in Table 4.6. They were also asked to complete the same rating task after using the two scales

(GSR-Aft1 and GSR-Aft2). As shown in Table 4.6, these reliability estimates were also significant but lower than expected (ranging from r = .26 to r = .44).

Table 4.6 Inter-rater reliability correlations for Global Severity Ratings (Before and After)

	GSR-Bef1 (n = 61)	GSR-Aft1 (n = 60)
GSR-Bef2 (n = 61)	0.44**	
GSR-Aft2 (n = 59)		0.26*

Note.

* Correlations significant at the .05 level

** Correlations significant at the .01 level

One explanation for the lower than anticipated correlations between the raters for all severity measures (i.e., SOI-G, CIRS-G, GSR-Bef, GSR-Aft) was found to be due to differences in the data collection procedure employed by the raters. The raters' comments at the second interview revealed that rater 2 was not following the instructions. Rater 2 included diseases present across all admissions while rater 1 rated only the most recent admission. In some cases, patients had multiple admissions thus rater 2 was consistently including more diseases than rater 1 for patients with multiple admissions.

To determine the level of level of agreement when both raters agreed on the presence of a disease or syndrome, the data were recoded to exclude any diseases or syndromes not endorsed by both raters. A correlational analysis was then conducted between the new total SOI-G1 and new total SOI-G2. This correlation was high and significant ($r = .88$, $p < 0.01$). This correlation

suggests that when both raters agreed that a disease was present, there was a high level of agreement regarding the severity rating of the disease. The data for CIRS-G was also recoded in this manner and a correlation computed. The new correlation was also high and significant ($r = .737$, $p < 0.01$). The higher reliability with the recoded data is consistent with the interpretation that the low inter-rater reliability observed for both SOI-G and CIRS-G reflected differences in the application of the instructions for using the measures rather than an inability to reliably apply the severity criteria.

Correlations between rater 1 and rater 2 were also examined across the SOI-G items and the CIRS-G categories. Spearman rank order correlations were calculated for CIRS-G given the ordinal nature of the values (Allen & Yen, 1979) and Pearson correlations were computed for SOI-G. It is acknowledged that the results of this exploratory analysis should be viewed cautiously given that the alpha level was not adjusted to control for familywise error.

With respect to the SOI-G, some correlations could not be computed (n/a) because one or both of the raters did not endorse that item. As illustrated in Table 4.7, several correlations were significant, ranging from $r = .26$ (GERD) to $r = .85$ (peripheral vascular disease). Over 60% of the items (21 of 32) demonstrated moderate to high ($r = .4$ or higher) agreement between the raters.

Table 4.7 Inter-rater reliability for SOI-G items

SOI-G Item	*r*	SOI-G Item	*r*
GERIATRIC SYNDROMES		**NEUROLOGIC**	
Dementia Syndrome	.38*	Alzheimer's Disease	.54*
Urinary Incontinence	.65*	Stroke	.44*
Malnutrition	.80*	Parkinson's Disease	.58*
Falls	NS	Peripheral Neuropathy	.48*
Pressure Ulcers	.44*	**ONCOLOGY**	
Constipation	NS	Multiple Myeloma	n/a
Vision/Hearing Impairment	.68*	Chronic Lymphocytic Leukemia	n/a
CARDIOVASCULAR		Lymphomas	n/a
Angina Pectoris	.43*	Solid Tumor	NS
Heart Failure	.48*	**ENDOCRINE**	
Peripheral Vascular Disease	.85*	Hypothyroidism	NS
HEMATOLOGIC		Diabetes Mellitis	.73*
Anemia	.43*	Osteoporosis	.42*
Thromboembolic Disease	n/a	**MUSCULOSKELETALIMMUNE**	
Myelodysplastic Syndrome	n/a	Rheumatoid Arthritis	.78*
GASTROINTESTINAL		Osteroarthritis	.45*
Gastroesophageal Reflux Disease	.26*	**RENAL/UROLOGIC**	
Cirrhosis	n/a	Renal Failure	.69*
Peptic Ulcer Disease	.60*	Benign Prostatic Hypertrophy	.43*
		RESPIRATORY	
		COPD & Asthma	.55*

Note. *All significant correlations at the .01 level

A similar analysis was conducted for the CIRS-G categories (see Table 4.8). Significant correlations were revealed for 69% of the categories and ranged from r = .32 to .64. Approximately 30% of the categories demonstrated moderate to high agreement between the two raters. The pattern of correlations across categories stands in contrast to the non-significant correlation between raters for the total CIRS-G score reported in Table 4.5. Overall, these results seem to underscore the potential for misleading conclusions when using summed ordinal values on CIRS-G.

Table 4.8 Inter-rater reliability for CIRS-G categories

CIRS-G Category	*r*	CIRS-G Category	*r*
Heart	.32*	Vascular	NS
Respiratory	.54**	Eyes, ears, nose, throat, & larynx	.32*
Upper gastrointestinal tract	.62**	Lower gastrointestinal tract	.28*
Liver	NS	Renal	.54**
Genito-urinary	.64**	Musculoskeletal/integument	.34**
Neurological	NS	Endocrine/metabolic & breast	NS
Psychiatric	.34**		

Note.

* Correlations significant at the .05 level

** Correlations significant at the .01 level

Obviously, the analysis of inter-rater reliability must assess the extent to which the two raters agreed that a disease/syndrome was present, however, another facet of agreement between raters involves exploring the extent to which the raters agreed that a disease/syndrome was not present on both

SOI-G and CIRS-G (see Table 4.9). Further analyses was conducted for each item on SOI-G that examined the frequency with which the following occurred: (1) rater 1 and rater 2 agreed that a disease/syndrome was present; (2) rater 1 and rater 2 agreed that a disease was absent; (3) rater 1 indicated an item was absent but rater 2 indicated the item was present; and, (4) rater 1 agreed an item was present but rater 2 indicated the item was absent. Using dementia to illustrate, both raters agreed that dementia was present for 46% of the cases and absent for 21% of cases. Rater 1 indicated dementia was absent and rater 2 indicated dementia was present for 20% of the cases. Lastly, rater 1 thought dementia was present but rater 2 thought it was absent for the remaining 13% of cases. In total, both raters agreed on the absence and presence of dementia for 67% of the cases. Table 4.9 illustrates that when the data were recoded to determine what percentage of the time both raters agreed that an item was present and also what percentage of the time both raters agreed that an item was absent, there was a high level of agreement between raters.

Table 4.9 Agreement between raters on absence/presence of disease/syndrome for SOI-G

SOI-G Item	Agreed Present (%)	Agreed Absent (%)	Cumulative (%)
GERIATRIC SYNDROMES			
Dementia Syndrome	46	21	67
Urinary Incontinence	13	73	86
Malnutrition	3	93	96
Falls	7	61	68
Pressure Ulcers	3	92	95
Constipation	15	62	77
Vision or Hearing Impairment	10	10	20
NEUROLOGIC			
Alzheimer's Disease	8	67	75
Stroke	8	79	87
Parkinson's Disease	3	90	93
Peripheral Neuropathy	2	95	97
RESPIRATORY			
COPD & Asthma	16	74	90
CARDIOVASCULAR			
Angina Pectoris	2	92	94
Heart Failure	3	75	78
Peripheral Vascular Disease	3	89	92
HEMATOLOGIC			
Anemia	10	71	81
Thromboembolic Disease	0	93	92
Myelodysplastic Syndrome	0	100	100

Table 4.9 (Continued)

SOI-G Item	Agreed Present (%)	Agreed Absent (%)	Cumulative (%)
ENDOCRINE			
Hypothyroidism	5	79	84
Diabetes Mellitis	12	84	96
Osteoporosis	26	54	80
ONCOLOGY			
Multiple Myeloma	0	100	100
Chronic Lymphocytic Leukemia	0	100	100
Lymphomas	0	98	98
Solid Tumor	3	92	95
GASTROINTESTINAL			
Gastroesophageal Reflux Disease	7	89	96
Cirrhosis	0	100	100
Peptic Ulcer Disease	7	84	91
MUSCULOSKELETAL/IMMUNE			
Rheumatoid Arthritis	3	92	95
Osteroarthritis	38	49	87
RENAL/UROLOGIC			
Renal Failure	7	82	89
Benign Prostatic Hypertrophy	3	95	98

As indicated in Table 4.9, of the 32 diseases/syndromes on SOI-G (excluding the "other" category), the level of agreement between two raters on the presence or absence of a disease/syndrome was over 80 percent for all but six of the SOI-G items. Of the six items with less than 80 percent agreement, all but one item demonstrated a level of agreement above 67 percent. The item with the lowest level of agreement was vision or hearing impairment at 20 percent agreement. The low agreement for this item will be discussed in more detail in the future directions section of the final discussion. This analysis continues to support the argument that the observed low inter-rater reliability was attributable to differences in data collection methods rather than a lack of a reliable instrument.

The comparison of rater agreement on items both present and absent was also conducted for CIRS-G (see Table 4.10). A pattern similar to that seen with SOI-G emerged that suggested a high level of agreement between the raters. Raters displayed an 80 percent level of agreement with respect to the presence/absence for seven of the thirteen CIRS-G categories. For the remaining six categories, raters demonstrated a level of agreement between 58 percent and 78 percent.

Table 4.10 Agreement between raters on absence/presence of disease category for CIRS-G

CIRS-G Category	Agreed Present (%)	Agreed Absent (%)	Cumulative (%)
Heart	35	35	70
Vascular	31	28	59
Respiratory	18	63	81
Eyes, ears, nose, throat, & larynx	7	71	78
Upper gastrointestinal tract	15	72	87
Lower gastrointestinal tract	5	77	82
Liver	0	95	95
Renal	12	73	85
Genito-urinary	23	62	85
Musculoskeletal/integument	67	10	77
Neurological	23	35	58
Endocrine/metabolic & breast	17	45	62
Psychiatric Illness	18	62	80

The analyses described above were based on agreement across categories for CIRS-G and across items (i.e., diseases/syndromes) for SOI-G (Tables 4.9 and 4.10, respectively). Some categorical comparisons between the two scales were also worthy of note including a 58 percent level of agreement between raters on CIRS-G neurological versus a range of agreement between 75 and 97 percent on the neurologic category (AD, stroke, PD, peripheral neuropathy) of SOI-G. Further, the level of agreement on CIRS-G respiratory (81%) was comparable to SOI-G respiratory which consists of

COPD/asthma (90%). SOI-G cardiovascular (angina, heart failure, peripheral vascular disease) agreement ranged between 78 and 94 percent while CIRS-G heart was 70 percent and CIRS-G vascular was 59 percent. SOI-G endocrine (hypothyroidism, diabetes, osteoporosis) agreement was between 80 and 96 percent while CIRS-G endocrine/metabolic and breast was at a 65 percent level of agreement. SOI-G gastrointestinal (GERD, cirrhosis, peptic ulcer disease) agreement ranged between 91 and 100 percent and was comparable to CIRS-G upper gastrointestinal (87%) and lower gastrointestinal (82%). Lastly, the renal categories were comparable between the two scales (89% for renal failure on SOI-G and 85% for CIRS-G). Overall, given that previous research has established the reliability of CIRS-G, the fact that SOI-G demonstrated comparable rates of agreement across categories relative to CIRS-G in this analysis offers promise for the SOI-G to demonstrate more compelling evidence of reliability in future research.

4.4.2.3 Comparison of Age with scores on SOI-G, CIRS-G, and global severity ratings

Before the specific hypotheses were examined, the possibility that older persons might have higher scores on the severity of illness measures was investigated. Age differences were explored by performing a Pearson product-moment correlation between age and the following: (1) SOI-G total score with "other"; (2) SOI-G total score without "other"; (3) global severity ratings (before and after); and, (4) total CIRS-G score. Age was not significantly related to scores on any of the measures. The correlations ranged from $r = -.03$ to $r = .13$.

4.4.2.4 Research Hypothesis 1(Comparison of SOI-G, CIRS-G and outcome)

It was hypothesized that higher SOI-G scores would be related to discharge outcome. Scores on both SOI-G and CIRS-G were compared for patients discharged to the following three categories: (1) home or home with home care; (2) semi-independent living situations; and, (3) long term care. One patient discharged to Emergency was combined with the LTC group based on comments by the nurses that suggested that this patient was likely severely ill. Patients discharged to the psychiatric hospital (n = 1) or discharged against physicians advice (n = 2) were not recoded (and therefore excluded from the analysis) as the level of care needed by these individuals could not be determined based on the available information.

A one-way analysis of variance (ANOVA) was conducted for SOI-G total score with "other" for rater 1, rater 2, and the average of both raters (Combined). Mean SOI-G scores for the three groups were significantly different ($F = 9.66$; $df = 2, 51$; $p < .0005$) for rater 1, and the combined score ($F = 3.91$; $df = 2, 51$; $p < .026$), but not for rater 2 ($F = .374$; $df = 2, 51$; NS). A summary of ANOVA results, and subsequent pair-wise comparisons using Student Newman-Keuls procedure are shown on Table 4.11.

Table 4.11 ANOVA and pair-wise comparison results: Mean SOI-G-Rater1, SOI-G-Rater 2, SOI-G-Rater 1 and 2 combined

Variable	F-ratio [1.]	*p*	Student-Newman-Keuls: $\alpha = .05$ [2.]		
SOI-G Total-Rater 1	9.66	<.0005	Home 99.8	Semi 144.8	LTC 197.2
SOI-G Total-Rater 2	0.37	.690 NS	Home 180.5	LTC 192.9	Semi 204.3
SOI-G Total-Combined	3.91	0.026	Home 140.3	Semi 174.8	LTC 195.3

Note 1. All F-ratios have d.f. = 2, 51

Note 2. Means connected by a common underline do not differ significantly.

Post-hoc Student Newman-Keuls analyses of rater 1 data revealed that, as expected, patients discharged home or home with home care received significantly lower SOI-G ratings (i.e., were less ill) than patients discharged to long-term care (see Table 4.11). Also, there was no significant difference in SOI-G ratings between those discharged home or to semi-independent living situations. Those discharged to semi-independent situations received higher SOI-G scores than patients discharged home but the difference was not significant. However, patients discharged to semi-independent situations were significantly less ill than patients discharged to long-term care.

The analyses just described were also conducted using the SOI-G total score without "other" resulting in a similar pattern of results. Mean SOI-G scores for the three groups were significantly different ($F = 8.87$; $df = 2, 51$; p , .0005) for rater 1 and not significant for rater 2 ($F = .420$; $df = 2, 51$; NS). Also, post-hoc Student Newman-Keuls analyses revealed the same pattern of significant differences between the groups that was demonstrated in the previous analysis.

One possible criticism of this analysis is that each rater knew the outcome before completing the SOI-G ratings and, therefore, their SOI-G ratings were influenced by knowledge of the patients outcome. However, if that were true, one would expect to have seen a similar pattern of scores emerge for rater 2 but that was not the case. As discussed in the earlier reliability section, rater 1 completed the SOI-G as instructed, i.e., rating only the most recent admission, while rater 2 rated diseases present for all admissions. Severity of illness at discharge (and its influence on function) was expected to be one of the major contributing factors with respect to discharge placement decisions made by the NARG clinical team. Overall, the empirical evidence is more consistent with the interpretation that the significant differences observed for rater 1 but not for rater 2 were due to more accurate recording by rater 1 relative to rater 2 than with the interpretation of rater bias.

If rater bias accounted for the significant pattern of differences observed between the three outcome groups for rater 1, a similar pattern of differences should have emerged using CIRS-G. A one-way analysis of variance (ANOVA)

indicated there were no significant differences for either rater 1 (F = .437; df = 2, 51; NS) or rater 2 (F = .209; df = 2, 51; NS). The failure to find significant differences between outcome groups for CIRS-G scores is inconsistent with the possibility that the severity of illness ratings made by rater 1 on SOI-G were biased by her knowledge of patient outcome.

The SOI-G was expected to perform better than the CIRS-G for two reasons: (1) greater specificity of domain-specific ratings present on SOI-G relative to the CIRS-G; and, (2) the interval properties of SOI-G equates all items on SOI-G on the same underlying scale while CIRS-G items assessed on an ordinal scale and therefore ratings are not equivalent. The results outlined above offers support for this hypothesis.

4.4.2.5 Research Hypothesis 2 (Comparison of SOI-G, CIRS-G, GSR-Bef, GSR-Aft)

Comparing SOI-G scores with scores on the Cumulative Illness Rating Scale-Geriatric (CIRS-G) and the Global Severity Ratings-Before and After (GSR-Bef and GSR-Aft, respectively) assessed the extent to which SOI-G possessed convergent validity. It was hypothesized that SOI-G scores would correlate highly with scores on the CIRS-G as well as the GSR-Bef and GSR-Aft. Correlations were performed using a Pearson product-moment correlation.

As illustrated in Table 4.12, significant correlations were observed between both the SOI-G with "other" and the SOI-G without "other" total scores and the other severity measures (i.e., CIRS-G, GSR-Bef, GSR-Aft) for rater 2. In contrast, neither SOI-G score (i.e., with "other" or without "other") for rater 1

correlated significantly with the CIRS-G or GSR-Aft score although a significant correlation was observed between the SOI-G scores (both with and without "other") and the GSR-Bef.

Table 4.12 Correlations between SOI-G and other scales for each rater

	Rater 1		Rater 2	
	SOI-G with "other"	SOI-G without "other"	SOI-G with "other"	SOI-G without "other"
CIRS-G	0.05	0.06	0.35**	0.36**
GSR-Bef	0.30*	0.31**	0.29*	0.35**
GSR-Aft	0.07	0.04	0.26*	0.32**

Note.

* Correlations significant at the .05 level

** Correlations significant at the .01 level

The relationship between SOI-G and GSR-Bef was significant but low for raters 1 and 2. Further, the relationship between SOI-G and GSR-Aft was significant but low for rater 2 and non-significant for rater 1. The low correlations and non-significant relationship may have been attenuated due to restriction of range. Both raters did not vary more than 40 points on either their before or after ratings. Rater 1 varied only 20 points on her before ratings and 40 points on her after ratings. Similarly, rater 2 demonstrated a range of 40 points for her before and after ratings. With a restricted range of scores the correlation is expected to be smaller than a similar correlation based on an unrestricted range of scores (Allen & Yen, 1979). Another explanation for the lack of relationship between the global ratings and SOI-G may lie in the obvious

differences in specificity between the two scales. As indicated previously, Parmelee et al. (1995) suggest that global ratings of health are less accurate than ratings with greater specificity and well defined response alternatives, such as that found on the SOI-G.

4.4.2.6 Qualitative comments on use of SOI-G

Rater 2 commented that some of the patient files contained diseases listed on the SOI-G but the patient was "asymptomatic" at the time of admission and therefore did not meet the criteria as defined on SOI-G. For example, the patient had heart failure but did not meet the mild classification "symptoms with moderate exercise" so was rated as "absent" on SOI-G. Some of the lack of agreement between raters and between measures may have been reflected in how this type of situation was interpreted. It is possible that the other rater decided to code this as "mild" on SOI-G or absent on SOI-G but present on CIRS-G. This type of difficulty poses a potential floor effect on SOI-G for diseases that are present but less than mild as these conditions might not be rated. Future development of SOI-G should include a training and instructional manual for users that directs raters to use the "other" category in cases where the disease is present but less than "mild" or greater than "severe".

The nurses indicated that severity criteria (i.e., mild, moderate, severe) for Parkinson's Disease (PD) as indicated in the medical chart did not always match the same severity criteria listed on SOI-G. They stated this was also true for other conditions such as anemia and renal disease. The potential variability among physicians or institutions with respect to what it means to have a mild,

moderate or severe form of an illness and the values listed on SOI-G can be dealt with by developing a detailed instructional manual. Instruction to raters should be clear that raters are to be guided by the severity criteria as defined on SOI-G unless this information can not be determined from file information. The issue of subjectivity between individuals/institutions in the use of descriptors such as "mild, moderate, and severe" highlights one of the advantages of SOI-G, namely standardization of severity criteria for each disease/syndrome. Offering descriptions of each level of each disease/syndrome (as was done with SOI-G) provides researchers with a tool that controls for variability in how severity criteria are operationalized. This, in turn, is expected to facilitate empirical comparisons of geriatric severity of illness across physicians and across institutions.

The nurses also reported that staging information listed on SOI-G for Alzheimer's Disease (AD) and Parkinson's Disease (PD) was not readily available from file information and therefore had to be extrapolated from nursing notes or therapies information (e.g., physical therapy reports). They added that the dementia item was easier to score than the AD item as it required less approximation on their part. One difficulty with extrapolating from other file information is the potential for a possible confound between diseases. For example, a patient may suffer from mild PD and severe AD and be wheelchair bound as a consequence of the AD but not the PD. In this example, raters unfamiliar with the client may misinterpret information in file (i.e., confined to wheelchair) as indicative of severe PD. Information in an instruction manual is

needed to caution raters to be sure that symptoms associated with one disease are not mistakenly used to assign severity ratings to a co-morbid disease.

With respect to the use of laboratory findings, the nurses indicated that at times lab results in the form of severity criteria made the extraction of information simple and straightforward. Other times, however, the laboratory information was unavailable in the file. The emphasizes the importance of including behavioural indicators with laboratory findings as suggested by geriatricians in Study 3.

4.4.2.7 Study 4 Conclusions

In summary, the results of Study 4 were encouraging but not definitive. It appears that with properly trained raters, SOI-G has the potential to be a reliable and valid severity of illness measure. However, even with the present limitations of the instrument, stable relationships emerged between SOI-G and outcome. Interestingly, these relationships did not emerge with an instrument known to be reliable and valid, the CIRS-G, and outcome. Typically it is assumed that without reliability there can not be validity, however, the fact that SOI-G demonstrated validity in Study 4 indirectly implies potential reliability.

4.4.3 Study Five: Patient Interviews

One of the two nurses trained for the archival file review in Study 4 collected SOI-G information from patient files in Study 5. The primary investigator (PI) conducted the individual interviews with patients for Study 5. In Study 5, 13 rehabilitation patients were interviewed using the SF-36 and the BDI-II while a nurse completed the SOI-G using patient file information.

Unfortunately, the original intent of Study 5 was unrealized due to limited sample size (n =13). Problems with data collection included almost half of the patients being excluded due to dementia (or other exclusionary criteria) and lack of follow-through from one physician asked to recruit their patients for the study. Thus, there was insufficient data to allow for any substantive analyses. The following hypotheses had been generated for Study 5:

1. Since higher SF-36 composite scores represent better overall health status and well being, there should be an inverse relationship between severity of illness and quality of life. As such, increasing severity of illness as reflected by high SOI-G scores were expected to be associated with diminished functional status and diminished general well being as reflected by low SF-36 scores (i.e., SOI-G scores were expected to be negatively correlated with SF-36 composite T-scores). It was expected that the correlation between SOI-G and the Physical Health composite score would be higher than the correlation between the SOI-G and the Mental Health composite score.

2. The strength of the relationship between SF-36 composite scores and SOI-G (Hypothesis #1) was expected only to be low to moderate as the SF-36 assesses the patient's reaction to their health status, i.e., the psychosocial component of the SF-36 was expected to add variability to the SF-36 scores which was relatively independent of the physiologically based illness severity measure.

3. Given previous research that has suggested a high incidence of depression in medically ill older adults, it was expected that higher BDI-II scores would be associated with greater severity of illness as measured by SOI-G.

4. Diminished functional status and well being as measured by self-report on the SF-36 was expected to be associated with higher depression scores on the BDI-II, i.e., both SF-36 Physical Health composite score and Mental Health composite scores should correlate with scores on BDI-II.

5. Given the psychosocial component of the SF-36, correlations between the BDI-II and the SF-36 composite scores (Hypothesis #4) were expected to be greater than the correlations between the SOI-G and the BDI-II (Hypothesis #3).

All data analyses were carried out using SPSS for Windows, version 9.

4.4.3.1 Study 5 Participants

Recruitment of participants for Study 5 occurred at NARG (Northern Alberta Regional Geriatric) program inpatient unit. Approximately 70 participants (aged 65 years or older) were invited to participate. Exclusion criteria for Study 5 included dementia, mental retardation, head injury, or thought disorder that would interfere with the participant's ability to complete the questionnaires or to give consent to participate in the study. Barriers to communication such as non-English speakers or deafness were also reasons for exclusion. Medical personnel working on the geriatric units where the study was conducted indicated to the primary investigator that approximately 80 to 90 percent of the patients on these units suffer from some form of dementia.

Of the 70 patients available during the data collection time period, 13 agreed to participate. The remaining 57 (26 male, 31 females) patients did not participate for the following reasons: 11 (7 male, 4 female) patients were excluded because the one physician did not return the patient list; 8 (3 male, 5 female) patients declined the invitation to participate; 34 (13 male, 21 female) patients met the exclusion criteria listed above; 3 (2 male, 1 female) patients were discharged at the time of invitation; and, 1 male patient was unavailable because he was in isolation.

Participants were volunteers and were not paid for their participation. Participants were informed that their participation was appreciated but their decision to participate in no way affected their ability to receive services. Participants were advised that their continued participation was voluntary and that they could terminate their participation in the study at any point.

4.4.3.2 Measures

Applicants were asked to complete the SF-36 (see Appendix J), BDI-II (Appendix K), and the demographic questionnaire (see Appendix L).

The SF-36 Health Survey is a generic measure designed to assess the patient's self-assessment of functional status as opposed to measuring the underlying disease (Hays, 1998; Ware & Sherbourne, 1992). Previous research has demonstrated that the SF-36 effectively measures factors related to physical and mental health (Berkman et al., 1999). The present study focused on the composite scores. Low scores on the Global Health Composite suggest that the individual's perception of his/her health problems are impeding life

functioning while low scores on the Physical Health Composite and Mental Health Composite suggest that perceived physical health problems and perceived mental health problems, respectively, are imposing limitations in functioning.

The Beck Depression Inventory-Second Edition (BDI-II) is a widely used, 21-item self-report scale that evaluates severity of depression in adults (Beck et al., 1996). Each item on the BDI-II is scored on a scale from 0 to 3, with a range of scores from 0 to 63. Increasing scores reflect increasing severity of depression. The use of the BDI-II has been empirically validated for use with elderly psychiatric inpatients (Steer, Rissmiller, & Beck, 2000), clinically depressed outpatients (Steer, Ball, Ranieri, & Beck, 1999) as well as with primary care medical patients (Arnau, Meagher, Norris, & Bramson, 2001). Two scores were calculated for the BDI-II. The first calculation was completed with all 21 items intact. The second calculation removed items believed to overlap between physical disease and depression (e.g., decreased appetite, insomnia, decreased energy, fatigue) and resulted in a new total score. The BDI-II *Cognitive* factor (BDI-CS) served as a measure of depression without the confounding problem of somatic symptoms.

4.4.3.3 Procedures

Study 5 required a person with health education (e.g., a nurse) to rate the patient using the SOI-G and an interviewer to administer the SF-36, the BDI-II, and the demographic questionnaire. The nurse was provided with financial remuneration ($65). In keeping with University of Alberta's ethical guidelines,

participants were not recruited by the primary investigator but rather each patient's attending physician recruited all study participants during regular rounds. Six physicians were asked to approach patients. One physician was away on vacation and the physician covering her patients was asked to approach those patients. All physicians agreed to participate but one physician did not return the patient list to the primary investigator. Participants who agreed to take part in the study were approached by the primary investigator who then discussed the study and consent form with them. The participant was given a copy of an information letter and the consent form (see Appendix M). The information letter and consent form were read aloud to each participant.

Once written consent was obtained from a participant, they were interviewed individually. A copy of the consent form was given to each participant. The SF-36 and the BDI-II were administered in random order. The instructions and questions for each of the questionnaires was read aloud by the interviewer and the interviewer recorded the participant's responses in the questionnaire booklet. The answer selections for each questionnaire were printed in large print on a laminated card that was placed in front of each participant during the interview. This procedure was followed for all participants to standardize the procedure and to assist participants who may have difficulty completing the questionnaires on their own (e.g., who have difficulty holding a pencil because of arthritis). Further, this approach was used in deal with the issue of data incompleteness associated with reading and completion difficulties on the SF-36 identified by Bjorner and Kristensen (1999) and McHorney et al.

(1994) and discussed earlier in the literature review. The majority of interviews required approximately 30 minutes.

The primary investigator was alert for possible comprehension, language or hearing difficulties that may have been evident in comments or expressed non-verbally (e.g., obvious hesitation when answering items). Participants were informed that if they experienced fatigue at any point in the study they could take a break or terminate their participation if they wished. No participants requested a break and therefore all measures were completed in one interview.

In order to ensure confidentiality, names did not appear on any of the questionnaires. Instead, numbers only identified all questionnaires. The completed consent forms and patient questionnaires will be filed separately and stored in a locked cabinet for a minimum of five years. The information was used only for research purposes. Results were analyzed and reported in group form, therefore no individual person can be identified.

4.4.4 Study 5 Results

The results of Study 5 were hampered by the meager sample size of 13 participants and thus the hypotheses could not be adequately tested. Study 5 results will be discussed beginning with a description of the demographic data followed by a discussion of an age analysis and a brief discussion related to the hypotheses outlined earlier.

4.4.4.1 Descriptive analysis of the basic data

The present analyses will begin by describing the demographic information provided by patients during the interviews. The descriptive data

includes information on age, gender, marital status, usual occupation, highest level of school completed, living situation, and length of hospitalization at the Glenrose. In addition to information about the patients, means and standard deviations are presented for all measures.

The sample consisted of 13 Caucasian participants (5 males, 8 females). The average age of participants was 79 years (SD = 10 years) and the participants ranged in age from 64 years to 94 years (see Table 4.13). As with Study 4, the majority of participants was widowed (69%) while 23 percent were married or common law, and 8 percent were divorced. None of the participants were single or never married. All of the women were widowed while the majority of the men were married or common law.

Table 4.13 Age, marital status, and education of Study 5 sample

Characteristic	Males	Females	Total
Mean age in years (n, SD)	75.0 (5, 10.6)	82.0 (8, 9.2)	79.3 (13, 10.0)
Range	69 to 94	64 to 93	64 to 94
Marital Status			
Married/CL	3	0	3
Divorced	1	0	1
Widowed	1	8	9
Education			
< High School	2	2	4
Some High School	2	3	5
Technical/Business	1	1	2
University	0	2	2

Table 4.13 also presents the highest grade level attained by participants. The majority of participants had some high school or less than high school (69%) while the remaining 31 percent had either technical or business school training or a university degree. The lowest grade level attained was Grade 7 and one participant reported no formal academic training due to the fact that he had polio as a child. Participants worked at a variety of occupations including service industry worker (e.g., hospital dietary), plumber, meter reader, nursing aide, clerical, truck driver, cab driver, teacher, homemaker, construction worker, janitor, and machinist.

Table 4.14 presents the length of stay at the GRH and the patients living situation. The majority of patients lived in their own home (54%) with approximately 28 percent of those having some form of outside assistance, such as home care or meals on wheels. Thirty-one percent lived in seniors's housing and two of the patients (15%) were in the process of being assessed for placement. The average hospital stay was 27 days although some of the participants had been transferred to the Glenrose from other hospitals and therefore had been hospitalized for longer than what is stated in Table 4.14.

Table 4.14 Patient living situation and length of hospital stay

Variable	Males	Females	Total
Mean length of hospitalization in days (SD)	24.6 (23.5)	29.0 (23.2)	27.3 (22.4)
Range	5 to 63	2 to 70	2 to 70
Living Situation			
Own home/apartment	3	2	5
Own home/ apartment with outside assistance (e.g., home care)	0	2	2
Senior's housing	2	2	4
Awaiting placement	0	2	2

Table 4.15 contains the means and standard deviations for each measure. The maximum total score possible for SOI-G (not including "other") was 2,512, the BDI-II is 62, and the BDI-CS is 24. Composite T scores were calculated for the SF-36 based on age-based norms. The SF-36 composite scores are transformed from raw scores to standardized scores with a mean is 50 with a standard deviation of 10. According to Hays (1998), low composite scores indicated that perceived health problems are impeding life functioning.

Table 4.15 Means and Standard Deviations for all Measures

Measure	Mean (SD)	Range
SOI-G Total with "other"	199.2 (96.6)	73 to 351
SOI-G Total without "other"	146.9 (73.4)	53 to 241
SF-36: Global Health Composite (GHC)	34.8 (10.2)	25 to 55
SF-36: Mental Health Composite (MHC)	35.5 (12.9)	22 to 55
SF-36: Physical Health Composite (PHC)	34.6 (7.3)	24 to 54
BDI-II	13.9 (11.9)	0 to 41
BDI-CS	4.9 (5.5)	0 to 16

As illustrated in Table 4.15, the average composite score for GHC, MHC, and PHC were approximately 1.5 standard deviations below the mean when compared to same-aged peers. Based on comparisons with the age-based standardization sample (65 years or older) provided by Hays (1998), it was determined that 12 of the 13 participants (92%) received a PHC that was below 81 percent of their peers (i.e., at the 19th percentile) with two of the participants scoring below 99 percent of same aged peers. Eight participants (62%) received a MHC score that was below the 91 percent of their peers and ten participants (77%) received a GHC score that was lower than 80 percent of their peers. Overall, these data suggest that most of the patients perceived their health difficulties to be imposing limitations in their daily functioning. Although decisive conclusions can not be derived from such a small sample of participants, it was interesting to note that this sample of people who perceived themselves to be in poor health obtained an average SOI-G score (199.2)

comparable to that obtained by participants discharged to long-term care (197.2) in Study 4 (see earlier Table 4.9).

With respect to the BDI-II, 5 participants (39%) obtained scores in the clinically depressed range while the remaining 8 subjects (61%) obtained scores indicative of no depression. Thus, the majority of patients were not depressed. However, for those scoring in the depressed range, four obtained scores suggestive of moderate depression and one participant scored in the severe depression range (41 out of a total of 63).

4.4.4.2 Comparison of Age with Scores on SOI-G and BDI-II

An analysis of possible age differences on the SOI-G and BDI-II was conducted. As the SF-36 composite scores were age-based, the SF-36 was not included in this analysis. Age differences were explored by performing a Pearson product-moment correlation between age and both the SOI-G and the BDI-II. The correlational analysis revealed that age was not significantly related to scores on either SOI-G or BDI-II. The correlations ranged between $r = -14$ and $r = .02$.

4.4.4.3 Research Hypotheses 1 to 5

The first hypothesis predicted that greater severity of illness (as indicated by high SOI-G scores) would be associated with diminished functional status and diminished general well being (as indicated by low SF-36 composite scores) while hypothesis 2 suggested that the magnitude of this relationship would be low to moderate. Pearson product-moment correlations were conducted between the SOI-G and the SF-36 but no significant relationships were found.

In light of previous research suggesting a relationship between depression and physical illness in the elderly, the third hypothesis predicted that higher BDI-II scores would be associated with greater severity of illness as measured by SOI-G. A Pearson product-moment correlation revealed no significant relationship. This results was not unexpected given the small sample size and the even smaller number of depressed participants (n = 5).

Previous research (Alexander, 2001) had demonstrated a relationship between the BDI-II and the SF-36 for different renal patient groups. Thus, hypothesis 4 predicted that diminished functional status and diminished well being (as indicated by low PHC and MHC scores) would be associated with higher depression scores on the BDI-II. Given the psychosocial component of the SF-36, correlations between the BDI-II and the PHC and MHC scores were expected to be greater than the correlations between SOI-G and BDI-II explored in hypothesis 3. Pearson product-moment correlations revealed a significant correlation between the BDI-II and the MHC ($r = -.62$, $p < .05$) but no significant relationship was observed between the BDI-II and the PHC.

Lastly, there was some limited anecdotal support for the validity of SOI-G. First, each patient was asked by the interviewer "what physical condition brought you to GRH". In all but one case (where the person reported "dizziness and confusion" that SOI-G revealed as PD and dementia), SOI-G ratings matched the participant's self-report. Second, in light of limited agreement between raters in Study 4 on the "vision and hearing" item, the primary investigator documented whether or not the person wore glasses and compared

this to SOI-G ratings. In all but two instances, the SOI-G ratings on "vision and hearing difficulties" matched ratings on SOI-G.

4.4.4.4 Comments on Study 5

In summary, severity of illness as measured by SOI-G was not related to self-reported health functioning or to self-reported levels of depression. Difficulties obtaining participants for Study 5 resulted in meager numbers of participants that would allow an adequate test of the proposed hypotheses. There was no evidence to support the convergent validity of SOI-G as there was insufficient data for anything but non-significant results. Further, given that only low to moderate relationships had been expected, detection of such relationships at the .05 level of significance would require a minimum of 40 participants (D. Scott, personal communication, January 2001).

A additional complication with Study 5 involved the use of the SF-36. Participants were asked to evaluate the impact of their physical and mental health difficulties on their ususal activities "in the past 4 weeks". For patients hospitalized for more than 4 weeks, their response to this question varied. Some indicated it had no effect as they were able to "get up for breakfast, go to therapy". Others attempted to guess how their health problems "might" influence their usual activities which was not the internal frame of reference used by the standardization sample against which participants were compared. In addition, the extent to which their health interfered with their social activities depended on the mobility of their friends and family to come to visit rather than on the participants ability to socialize (which was the intent of the SF-36).

5. Discussion

The present investigation involved a series of five programmatically linked studies aimed at developing a valid and reliable measure of severity of illness composed of disease-specific scales appropriate for use with a geriatric population suffering from a variety of physical illnesses. This index was called the Severity of Illness-Geriatric (SOI-G) and was intended to quantify severity of illness on a single numerical scale with interval properties independent of psychosocial variables.

The initial three studies focused specifically on the construction of SOI-G while the final two studies attempted to examine the reliability and validity of SOI-G. The evidence of reliability will be discussed first followed by evidence in support of validity. The evidence from the present study failed to conclusively support the reliability of SOI-G. However, as will be revealed in the following discussion, the conclusion that SOI-G was unreliable can also be challenged.

The reliability estimates of the initial severity ratings obtained from the Study 3 scaling panel yielded evidence for the reliability of the severity values. The strong initial agreement (Coefficient alpha = .96) on the values obtained from five independent raters suggests that the method by which severity values were obtained was reliable. The initial agreement between raters was viewed as a low bound estimate as the nature of the panel meeting (i.e., achieving

consensus on each level of each item) suggests that even greater agreement was achieved by the end of the second round of the scaling task.

Study 4 revealed statistically significant but lower than anticipated inter-rater reliability. While the lack of favorable results from the reliability study could be interpreted as evidence that SOI-G was unreliable, there was evidence to support alternative interpretations of this data. It was revealed during subsequent interviews with the raters that one rater did not follow the instructions and therefore the two raters were not applying the instrument in a parallel manner. Rater 2 revealed that she included data from every admission to the rehabilitation hospital while the other rater correctly included data only from the most recent admission. This difference could account for the lack of agreement between raters as some patients had multiple admissions. Unfortunately, the number of admissions was not recorded and therefore this interpretation could not be tested empirically.

If the low inter-rater reliability was primarily due to the raters documenting different chart histories and not due to SOI-G being unreliable, one would expect to see a high level of agreement when the raters agreed on the presence of a disease. In fact, this is what subsequent exploratory analyses revealed. When the analyses were performed only for items that both raters agreed were present, the resulting reliability was high and significant ($r = .879$). Similarly, correlations between raters for each SOI-G item revealed moderate to high levels of agreement across most of the items ($r = .4$ or higher). Lastly, when the level of agreement on diseases not present was evaluated, a high level of

agreement emerged again across diseases/syndromes with raters agreeing that a disease was absent or present for over 80% of the cases for most items.

If it were true that the raters had applied SOI-G in a consistent fashion and thus the results were due to SOI-G being an unreliable instrument, one would expect to see a high level of agreement between raters for measures with known reliability such as the CIRS-G. However, this was not the case and, in fact, the reliability between the raters for CIRS-G was even lower ($r = .23$) than SOI-G and was not significant. In addition, similar to SOI-G, the reliability improved when only the items that were agreed upon by both raters were considered. Thus, the similar pattern of reliability results for both SOI-G and CIRS-G is consistent with the interpretation that the low reliability observed with both measures was due to differences in the data collection procedure.

It was surprising that the present investigation failed to replicate previous research that demonstrated the reliability and validity of CIRS-G. In addition to the previously acknowledged difficulties, the inconsistency with previous research may have been due to inherent difficulties with the suggested scoring method (i.e., summing ordinal data to obtain a total severity score). Each level of each disease/syndrome on SOI-G fit an interval scale, i.e., the severity criteria values are assumed to be equidistant in terms of the underlying threat to life scale. Furthermore, all items on SOI-G contribute equally to the overall scale (e.g., a "20" represents the same amount of the underlying construct for every item on the scale). Thus, SOI-G meets the implicit assumptions associated with a simple summation approach while CIRS-G does not. The

advantages of an interval scale combined with the greater specificity of domain-specific ratings present on SOI-G suggests that total SOI-G scores may be more meaningful than CIRS-G total scores in research requiring an overall assessment of geriatric severity of illness.

Together the pattern of results described for the present research suggested that the failure to demonstrate strong reliability with SOI-G was associated with rater error rather than an unreliable instrument. Although these results were encouraging, they must be viewed with caution given their exploratory nature. If the apparent difficulty that raters had with respect to the application of SOI-G can be remedied, SOI-G promises to be a reliable measure of illness severity; however, it remains essential that the reliability of SOI-G be empirically demonstrated in future research.

Assessing the validity of an instrument is, of course, important in any study. The validity of a measure indicates the degree to which it measures what it claims to measure. Traditionally, three basic types of validity are considered when assessing the validity of an instrument, namely construct, content, and criterion related validity (Allen & Yen, 1979). The present study focused primarily on only two of these forms of validity: content and construct validity. The validity of SOI-G was supported by the following results: (1) consistency of SOI-G items with previous research and government statistics; (2) qualitative comments made by scaling panel; and, (3) demonstration of expected differences in SOI-G scores for patients with different discharge outcomes. These results will be discussed in more detail in the following section.

The content validity of SOI-G was evaluated in different ways. A geriatric specialist was asked to select the items and severity criteria that he felt would reflect important differences between less sick and more sick elderly persons. The final list of items derived in Study 1 replicated the substantial body of literature which has identified many of the SOI-G diseases/syndromes as those most commonly experienced by persons in the later years of life. The items selected were also consistent with the leading causes of death for individuals over the age of 65 years identified by Statistic Canada data.

Further refinement of the scale came about through the development of severity criteria and by applying the Delphi technique (Studies 2 and 3). The panel of five geriatric specialists made some revisions to items and criteria but largely accepted the items as they were at the end of Study 2. This acceptance supports the face and content validity of SOI-G suggesting that the SOI-G items accurately reflect the medical community's view of the impact of individual diseases/syndromes on a person's life. Lastly, there was anecdotal evidence of validity observed in the similarity in scale values between SORDS and SOI-G.

Construct validity was more difficult to assess in the present investigation because of the lack of other generally accepted valid measures of illness severity measures with interval properties. As there was no gold standard for comparison, Study 4 examined convergent validity with a similar health status measure, the CIRS-G. There appeared to be weak evidence of convergent validity evidence for rater 2 but given that it was revealed that rater 2 did not

follow the instructions, it is difficult to know how to interpret the significant albeit low correlations observed across the SOI-G, CIRS-G and GSR for rater 2.

Given that SOI-G and CIRS-G scores assess severity of illness, it was expected that patients needing greater levels of care would receive higher scores on these measures. This prediction was supported with SOI-G but not with CIRS-G. Patients discharged to long-term care had significantly higher SOI-G scores than patients discharged home (including patients discharged home with home care) or discharged to the community but not able to live independently (e.g., lodges). Thus, support for this prediction provides encouraging support for the construct validity of SOI-G.

It appears that SOI-G scores were sensitive to differences among patients with respect to discharge outcome. As such, SOI-G is potentially sensitive to differences in placement need (e.g., patients with greater independence and lower levels of care versus those who are less independent and requiring greater levels of care).

Validity could not be adequately explored in Study 5 because of a limited sample size. It is recognized that the small number of subjects in Study 5 was a significant limitation of the present research as it did not allow for an examination of the potential usefulness of SOI-G in psychosocial research. Future research with SOI-G should focus on examining the relationship between SOI-G and other health status measures. Such research, if successful, would provide strong evidence for SOI-G convergent validity. Given the difficulties with patients being excluded because of dementia in Study 5, it is suggested that

future research consider the use of health status measures that can be completed by others (such as family members or nurses). In addition, for patients who have been hospitalized for extended periods of time, some of the SF-36 questions may not be appropriate. The SF-36 may be more appropriate in research with recently admitted acute care patients or with community dwelling patients seeking treatment at a clinic.

One last issue to consider in the evaluation of SOI-G is that of validity generalization. In short, this refers to the assumption that relationships observed between instruments in one setting should generalize to other similar settings (D. Scott, personal communication, June 12, 2001). Previous research with SORDS (Alexander, 2001), a severity of illness measure with properties that are similar to SOI-G, established relationships between SORDS and the SF-36 with renal patients. Given the pattern of relationships observed in Alexander's research, one has reason to expect that SOI-G may perform in a similar fashion once the difficulties identified in the present investigation are successfully remedied.

Continuing research is needed to definitively establish the reliability and validity of SOI-G. Future research with SOI-G should consider the following: (1) an examination of the relationship between SOI-G scores and mortality; (2) the development of training procedure and instruction manual; (3) the inclusion of time element ; (4) future development of "other" category; (5) changes to specific items; and, (6) clarifications regarding user of SOI-G . Each of these recommendations for future research will be addressed next.

The present research had planned to examine the ability of SOI-G to predict mortality in Study 4 but was unable to do so because there were no deaths indicated in the reviewed files. Subsequent to the data collection, it was revealed that death was an uncommon occurrence for patients at the hospital where the data was collected. Clients severely ill enough to be near death would be outside the scope of rehabilitation services and thus the majority of patients included in the present study were likely in the low to middle range of severity of illness. Support for this supposition was the observation that many of the values reported for the CIRS-G categories were either mild or moderate (89% of cases for rater 1; 84% of cases for rater 2). Future studies could examine the relationship between SOI-G and mortality by examining one or both of the following: (1) an archival review of randomly selected acute care patient files; or, (2) a longitudinal study using randomly selected patients admitted to hospital. The latter alternative could be conducted by examining SOI-G scores at admission, discharge, and at follow-up. This type of study would also yield useful information regarding treatment outcome. For example, for individuals discharged home, one would expect to see a decrease in SOI-G scores thus reflecting successful treatment interventions.

In order to further improve upon the existing reliability of SOI-G, it is recommended that an instruction manual be developed and a more intense training session be considered. The inconsistent use of SOI-G by the raters may have been due to a lack of clarity in the written instructions or the training procedure. Although there was a statement at the top of SOI-G directing the

raters to "record the patient's latest test results" and each rater was directed by the primary investigator to only rate the most recent admission, the difficulties experienced in Study 4 suggest that it may be necessary to develop a training manual with explicit rating instructions to reduce procedural differences in data collection method. Further, it may be necessary to include a longer training session than was used in this study (i.e., 30 minutes prior to data collection and ongoing instruction as difficulties emerged during data collection) to ensure that the raters are perfectly clear on the task. Training might be enhanced by the inclusion of several case examples to illustrate to potential raters the intended use of the SOI-G. While these suggestions may help to improve consistency between raters in the future, it was important in the present study not to bias raters by providing too much information. By keeping the instructions simple at this point in the development of the instrument, it was possible to identify potential problem areas without unduly influencing the raters.

The geriatricians in Study 3 stated that, for some diseases, the severity of illness is also a function of the time it takes for the disease to develop. Thus, further development of SOI-G may wish to consider incorporating a time element as part of the severity criteria. This would require the SOI-G to be re-scaled for some conditions where specification of a time frame is judged to have a potential impact upon the severity level rating. For such conditions, the original scaling procedure using a panel of geriatricians would most likely need to be replicated. While the incorporation of a time element may be an important future consideration, it is sufficiently premature to say definitively whether or not

such a change would significantly enhance the effectiveness of SOI-G to assess geriatric severity of illness. The degree of precision in the SOI-G is related to the intended measurement purpose. Should future research reveal that total scores are not sufficiently precise to be useful as a control variable in psychosocial research, the time element may need to be revisited.

Another purpose of the present study was the development of an "other" category to capture items not on SOI-G. One rater used the "other" category more often than the other rater. However, this same rater was rating all admissions for each patient while the other rater only rated the most recent admission. This, in all likelihood, accounted for the lower reliability between raters when the "other" category was included in the total score. That being said, the need for more explicit instructions on how to use the "other" category would likely improve the reliability of this item. Items that were recorded frequently included atrial fibrillation, ulcerative colitis, goiter, hypertension, hip fractures, dysphagia, and gout. While it may be necessary to incorporate these items into SOI-G in the future, additional research is needed to determine whether the frequency of the "other" items identified with medically ill rehabilitation patients would generalize to other geriatric populations. For example, hip fractures are not expected to occur frequently among community dwelling elderly.

The one item that demonstrated the greatest disparity between raters was that of vision or hearing difficulties. The discrepancy between the raters for this item is difficult to interpret. We know that one rater was rating all

admissions while the other rated the most recent admission. Therefore, it may have been the case that eye difficulties were present on previous admissions but not on the most recent admission. It is also possible that the raters were unclear about what type of vision or hearing difficulty to include. It may be important to clarify in the instructional manual that mild vision and hearing difficulties does not necessary have to be due to a disease process (e.g., cataracts) but rather should include any visual or hearing impairment (e.g., nearsightedness). Lastly, the instructional manual should clarify for raters that they are not to consider prognosis in rating decisions. In Study 5, the SOI-G rater rated "vision and hearing" as severe (i.e., irreversible blindness) for a patient with macular degeneration. While blindness may be the expected outcome for this patient, the interviewer observed that this patient was able to read. This inconsistency, however, may have resulted because the rater was unable to determine the severity of the condition from the medical chart alone.

Further investigations of reliability need to assess whether SOI-G can be used by any medical professional or whether only physicians can reliably use the scale. In addition, the difficulty just described with macular degeneration leads one to question whether or not familiarity with the patient being rated is necessary in order for a rater to use SOI-G in a reliable manner. For instance, it is important that raters be able to differentiate between the symptoms of co-morbid diseases when using SOI-G. It is not yet clear whether this can be done from file information alone. Future study could explore these issues by comparing SOI-G ratings for the following groups: (1) physicians who know

patient; (2) physicians who don't know patient; (3) nurses who know patient; and, (4) nurses who don't know patient.

The failure to account for the possible interactive effect of multiple conditions and a possible exponential increase in disease severity when multiple conditions are present is a limitation of SOI-G that may need to be addressed in future research. Two common problems associated with multiple pathology include disease-disease interactions and disease-treatment interactions (Besdine, 1997). Disease-disease interactions involve the interaction of two diseases that together work to the detriment of the patient (Besdine, 1997). Disease-treatment interactions involve the iatrogenic harm associated with the interaction between an unidentified illness and treatment undertaken to manage a diagnosed problem (Besdine, 1997). However, as with the issue of including a time element, it may be true that this level of precision is not needed for SOI-G to be a useful tool in geriatric psychosocial research.

One last area to be discussed involves the scale use, training, and necessary qualifications a person using the SOI-G should possess. SOI-G has potential for use as an outcome measure in the assessment of the impact of health policies and programs developed for elderly persons. For example, incorporating the SOI-G as part of the routine discharge summary information in hospital medical files would provide easy access to severity of illness information that might be used to assess the effectiveness of treatment interventions. Study 4 suggests the potential usefulness of SOI-G in the allocation of resources within long-term care.

With respect to training and user qualifications, raters should at a minimum have some sort of formal medical training in order to administer the instrument reliably. It may also be of benefit to train raters to criterion; i.e., have raters rate several hypothetical medical scenarios until a certain level of reliability is achieved. It remains to be established in future research whether physicians alone are qualified to complete the scale or whether nurses have sufficient training to complete the instrument. Given the multi-disciplinary nature of most health care facilities and treatment teams, the most effective solution may involve having multiple raters from the same research team complete the SOI-G. For example, physicians could rate the items requiring more in depth medical knowledge and differential diagnoses (e.g., myelodysplastic syndrome) while nursing staff could complete the items that they have more direct experience with based on their day to day interactions with patients (e.g., rating behaviours associated with the Dementia Syndrome item).

Despite the previously identified cautions about the interpretation of the results and the inherent difficulties of any scale attempting to quantify complex medical diseases and syndromes, the SOI-G in its current state of development demonstrated promising utility for the quantification of severity of illness in geriatric patients. In terms of methodological and theoretical contributions, the present research demonstrated the potential usefulness of using a measure of chronic health problems in the elderly which is independent of psychosocial variables. The SOI-G was simple and easy to apply. It was able to identify subgroups of elderly participants in Study 4 that were homogenous with respect

to severity of illness. It provided a single, numerical score with interval qualities that was not a function of hospital practices and could be used on patients with varying levels of cognitive impairment. Lastly, in light of evidence to suggest that the SOI-G items reflect the medical community's view of the impact of each level of each item on a person's life, this measure appears appropriate for use in clinical settings.

In conclusion, while the evidence failed to conclusively support the reliability of SOI-G, the pattern of results did not disconfirm reliability. Further research is needed to address the issue of SOI-G reliability before definitive conclusions can be made. The results of the present investigation supported the reliability of the severity values obtained in Study 3 as well as the content and face validity of SOI-G. It was sensitive to differences among discharge outcome groups supporting the construct validity of SOI-G. In light of the fact that reliability is a precondition for validity, this would only happen if SOI-G was reliable. Convergent validity of SOI-G was not established in Study 4 perhaps due to difficulties associated with the comparison instrument (i.e., summing ordinal values on CIRS-G). An attempt to assess convergent and divergent validity with geriatric patients in Study 5 was not realized due to many potential patients being excluded because of dementia and the non-participation of one physician who had originally agreed to recruit patients. Thus, while the results of the present investigation were encouraging, further research is needed to assess the utility of SOI-G in psychosocial research with elderly persons.

6. References

Abrams, R.C., Alexopoulos, G.S., Spielman, L.A., Klausner, E., & Kakuma, T. (2001). Personality disorder symptoms predict declines in global functioning and quality of life in elderly depressed patients. American Journal of Geriatric Psychiatry, 9, 67-71.

Alexander, D. (2001). An investigation of the validity and reliability of the Severity of Renal Disease Scale (SORDS). Unpublished doctoral dissertation, University of Saskatchewan, Saskatoon, Saskatchewan.

Allen, M.J., & Yen, W.M. (1979). Introduction to measurement theory, Bemont, CA: Wadsworth Inc.

American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.

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 An International Journal of Quality of Life Aspects of Treatment, 2, 369-395.

Andresen, E.M., Patrick, D.L., Carter, W.B., & Malmgren, J.A. (1995). Comparing the performance of health status measures for healthy older adults. Journal of the American Geriatrics Society, 43, 1030-1034.

Arfken, C.L., Lichtenberg, P.A., & Tancer, M.E. (1999). Cognitive impairment and depression predict mortality in medically ill older adults. Journal of Gerontology, 54A, M152-M156.

Arnau, R.C., Meagher, M.W., Norris, M.P., & Bramson, R. (2001). Psychometric evaluation of the Beck Depression Inventory-II with primary care medical patients. Health Psychology, 20, 112-119.

Arnetz, B.B. (1996). Causes of change in the health of populations: A biopsychosocial viewpoint. Social Science and Medicine, 43, 605-608.

Aronow, D.R. (1988). Severity-of-illness measurement: Applications in quality assurance and utilization review. Medical Care Review, 45, 339-366.

Baldwin, R.C., & Jolley, D.J. (1986). The prognosis of depression in old age. British Journal of Psychiatry, 149, 574-583.

Baltzan, M., Faulkner, Firor, W., Keegan, D., Scott, D., Shepel, L., Plypchuk, G., & Wall, R. (1987). [Development of Severity of Renal Disease Scale-SORDS]. Unpublished manuscript.

Bartels, S.J., Forester, B., Miles, K.M., & Joyce, T. (2000). Mental health service use by elderly patients with bipolar disorder and unipolar major depression. American Journal of Geriatric Psychiatry, 8, 160-166.

Bassiony, M.M., Steinberg, M.S., Warren, A., Rosenblatt, A., Baker, A.S., & Lyketsos, C.G. (2000). Delusions and hallucinations in Alzheimer's Disease: Prevalence and clinical correlates. International Journal of Geriatric Psychiatry, 15, 99-107.

Bastos, P.G., Sun, X., Wagner, D.P., Wu, A.W., & Knaus, W.A. (1993). Glasgow Coma Scale score in the evaluation of outcome in the intensive care unit: Finding from the Acute Physiology and Chronic Health Evaluation III study. Critical Care Medicine, 21, 1459-1465.

Beck, A.T., Steer, R.A., & Brown, G.K. (1996). Beck Depression Inventory-2nd Edition (BDI-II) Manual. San Antonio: The Psychological Corporation.

Becker, R.B., Zimmerman, J.E., Knaus, W.A., Wagner, D.P., Seneff, M.G., Draper, E.A., Higgens, T.L., Estafanous, F.G., & Loop, F.D. (1995). The use of APACHE III to evaluate ICU length of stay, resource use, and mortality after coronary artery by-pass surgery. Journal of Cardiovascular Surgery, 36, 1-11.

Bergner, M., Bobbitt, R., Carter, W., & Gilson, B.S. (1981). The Sickness Impact Profile: Development and final revision of a health status measure. Medical Care, XIX, 787-805.

Berkman, B., Chauncey, S., Holmes, W., Daniels, A., Bonander, E., Sampson, S., & Robinson, M. (1999). Standardized screening of elderly patients' needs for social work assessment in primary care: Use of the SF-36. Health & Social Work, 24, 9-16.

Besdine, R.W. (1997). Clinical approach: An overview. In C. Cassel, H. Cohen, E. Larson, D. Meier, D. Resnick, L. Rubenstein, & L. Sorenson (Eds.), Geriatric Medicine (3rd ed., pp. 155-168). New York, NY: Springer.

Bjorner, J.B. & Kristensen, T.A. (1999). Multi-item scales for measuring global self-rated health. Research on Aging, 21, 417-439.

Boberg, A.L., & Morris-Khoo, S. (1992). The Delphi method: A review of methodology and an application in the evaluation of a higher education program. Canadian Journal of Program Evaluation, 7, 27-39.

Borschelt, M., Gilberg, R., Horgas, & Geiselmann, B. (1999). On the significance of morbidity and disability in old age. In P.B. Baltes & K.U. Mayer (1999). The Berlin Aging Study: Aging from 70 to 100 (pp. 403-429). New York, NY: Cambridge University Press.

Bozzola, F.G., Gorelick, P.B., & Freels, S. (1992). Personality changes in Alzheimer's Disease. Archives of Neurology, 49, 297-300.

Britton, P.H. (1985). Reliability, validity and SORDS: The Severity of Renal Disease Scale. Unpublished honours thesis, University of Saskatchewan, Saskatoon, SK, Canada.

Bruce, M.L., & Leaf, P.J. (1989). Psychiatric disorders and 15-month mortality in a community sample of older adults. American Journal of Public Health, 79, 727-730.

Burns, A., Lewis, G., Jacoby, R., & Levy, R. (1991). Factors affecting survival in Alzheimer's disease. Psychological Medicine, 21, 363-370.

Burns, R.B., Moskowitz, M.A., Ash, A., Kane, R.L., Finch, M.D., & Bak, S.M. (1992). Self-report versus medical record functional status. Medical Care, 30, MS85-MS95.

Burvill, P.W., Hall, W.D., Stampfer, H.G., & Emmerson, J.P. (1991). The prognosis of depression in old age. British Journal of Psychiatry, 158, 64-71.

Burvill, P.W., Mowry, B., & Hall, W.D. (1990). Quantification of physical illness in psychiatric research in the elderly. International Journal of Geriatric Psychiatry, 155, 673-679.

Canadian Study of Health and Aging Working Group (1994). Canadian Study of Health and Aging: Study methods and prevalence of dementia. Canadian Medical Association Journal, 150, 899-913. Retrieved September, 1999, http://www.uottawa.ca/academic/med/epid/core1.htm

Caputo, L., Boller, F., Verny, M. Hugonot-Diener, L., Saillon, A., Traykov, L., Jahchan, T., Forette, F., De Ladoucette, O. and the Greco Study Group on Severe Demencia (1998). Severe dementia in the elderly: How to evaluate it? Archives of Gerontological Geriatrics, suppl. 6, 57-64.

Charlson, M.E., Sax, F.L., Mackenzie, C.R., Fields, S.D., Braham, R.L., & Douglas, R.G. (1986). Assessing illness severity: Does clinical judgment work?. Journal of Chronic Disease, 39, 439-452.

Chenier, N.M. (1993). The health of the Canadian elderly, (November, 1993), Ottawa: Library of Parliament.

Clark, M.W., Nash, J.M., Cohen, R., Chase, C., & Niaura, R. (1998). Psychological factors in cardiac disease. In Kleespies, P.M. (Ed.), Emergencies in Mental Health Practice: Evaluation and management (pp. 355-376). New York, NY: The Guildford Press.

Clayton, M.J. (1997). Delphi: A technique to harness expert opinion for critical decision-making tasks for education. Educational Psychologist, 17, 373-386.

Cohen-Cole, S.A. & Kaufman, K.G. (1993). Major depression in physical illness: Diagnosis, prevalence, and antidepressant treatment (A ten-year review: 1982-1992). Depression, 1, 181-204.

Conwell, Y., Forbes, N.T., Cox, C., & Caine, E.D. (1993). Validation of a measure of physical illness burden at autopsy: The Cumulative Illness Rating Scale. Journal of the American Geriatrics Society, 41, 38-41.

Copeland, J.R.M., Davidson, I.A., Dewey, M.E., Gilmore, C., Larkin, B.A., McWilliam, C., Saunders, P.A., Scott, A., Sharma, V., & Sullivan, C. (1992). Alzheimer's disease, other dementias, depression, and pseudo-dementia: Prevalence, incidence, and three-year outcome in Liverpool. British Journal of Psychiatry, 161, 230-239.

Craven, J., Littlefield, C., Rodin, G., & Murray, M. (1991). The End Stage Renal Disease–Severity Index (ESRD-SI). Psychological Medicine, 21, 237-243.

Crist, P.A. (1999). Does quality of life vary with different types of housing among older persons? A pilot study. Aging in place: Designing, adapting, and enhancing the home environment, 101-116.

Curyto, K.J., Chapleski, E.E., & Lichtenberg, P.A. (1999). Prediction of the presence and stability of depression in the Great Lakes Native American elderly. Journal of Mental Health and Aging, 5, 323-340.

Davidson, I.A., Dewey, M.E., & Copeland, J.R.M. (1988). The relationship between mortality and mental disorder: Evidence from the Liverpool longitudinal study. International Journal of Geriatric Psychiatry, 3, 95-98.

De-Bruin, A.F., de-Witte, L.P., Stevens, F., & Diederiks, J.P. (1992). Sickness Impact Profile: The state of the art of a generic functional status measure. Social Science and Medicine, 35, 1003-1014.

Dworkin, S.F., & Wilson, L. (1993). Measurement of illness behavior: Review of concepts and common measures. In P.M. Conn (Ed.), Paradigms for the study of behavior. Methods in Neurosciences. (Vol. 14, pp. 329-342). San Diego, CA: Academic Press.

Engel, G.L. (1997). From biomedical to biopsychosocial: Being scientific in the human domain. Psychotherapy and Psychosomatics, 66, 57-62.

Feinstein, A. (1992). Benefits and obstacles for development of health status assessment measures in clinical settings. Medical Care, 30, MS50-MS56.

Femia, E.E., Zarit, S.H., & Johansson, B. (2001). The Disablement Process in very late life: A study of the oldest-old in Sweden. Journal of Gerontology, 56B, P12-P23.

Fichter, M.M., Meller, I., Schroppel, H., & Steinkirchner, R. (1995). Dementia and cognitive impairment in the oldest old in the community: Prevalence and comorbidity. British Journal of Psychiatry, 166, 621-629.

Finkel, S.I. (1996). Efficacy and tolerability of antidepressant therapy in the old-old. Journal of Clinical Psychiatry, 57, 23-28.

Fish, L.S., & Bushby, D.M. (1996). The Delphi method. In D.H. Sprenkle, & S.M. Moon (Eds.), Research methods in family therapy, (pp. 469-482). New York: The Guilford Press.

Franssen, E.H., Kluger, A., Torossian, C.O., & Reisberg, B. (1993). The neurologic syndrome of severe Alzheimer's Disease: Relationship to functional decline. Archives of Neurology, 50, 1029-1039.

Fredman, L., Schoenbach, V.J., Kaplan, B.H., Blazer, D.G., James, S.A., Kleinbaum, D.G., & Yankaskas, B. (1989). The association between depressive symptoms and mortality among older participants in the epidemiologic catchment area-Piedman Health Survey. Journal of Gerontology, 44 S149-156.

Furner, S.E., Brody, J.A., & Jankowski, L.M. (1997). Epidemiology and aging. In C.K. Cassel, H.J. Cohen, E.B. Larson, Meier, D.E., Resnick, N.M., Rubenstein, L.Z., Sorenson, L.B. (Eds.), Geriatric Medicine (3rd Ed., pp. 37-44). New York, NY: Springer.

Gallagher, M., Hares, T., Spencer, J., Bradshaw, C., & Webb, I. (1993). The nominal group technique: A research tool for general practice? Family Practice, 10, 76-81.

Geiselmann, B., & Bauer, M. (2000). Subthreshold depression in the elderly: Qualitative or quantitative distinction?. Comprehensive Psychiatry, 41, 32-38.

Girzadas, P.M., Counte, M.A., Glandon, G.L., & Tancredi, D. (1993). An analysis of elderly health and life satisfaction. Behavior, Health, and Aging, 3, 103-117.

Gove, W. (1984). Gender differences in mental and physical illness: The effects of fixed roles and nurturant roles. Social Science Medicine, 19, 77-91.

Government of Canada (1996). Age and disability, Info-Age, 16. Retrieved September, 1999, from http://www.hc-sc.gc.ca/seniors-aines/issues/english/pubs/info_16e.htm

Gowan, J.A., & McNichols, C.W. (1993). The effects of alternate forms of knowledge representation on decision-making consensus. International Journal of Man-Machine Studies, 38, 489-507.

Green, J., Wintfeld, N., Sharkey, P., & Passman, L.J. (1990). The importance of severity of illness in assessing hospital mortality. Journal of the American Medical Association, 263, 241-246.

Hall, J.A., Epstein, A.M., & McNeil, B.J. (1989). Multidimensionality of health status in an elderly population. Medical Care, 27, S168-S177.

Hassan, R., & Pollard, A. (1994). Late-Life-Onset Panic Disorder: Clinical and demographic characteristics of a patient sample. Journal of Geriatric Psychiatry and Neurology, 7 (Apr.-June), 86-90.

Haynie, D.A., Berg, S., Johansson, B., Gatz, M., Zarit, S.H. (2001). Symptoms of depression in the oldest old: A longitudinal study. Journal of Gerontology, 56B, P111-P118.

Hays, R.D. (1998). RAND-36 Health Status Inventory, San Antonio: The Psychological Corporation.

Hoehn, M.M., & Yahr, M.D. (1967). Parkinsonism: Onset, progression, and mortality. Neurology, 17, 427-442.

Horn, S.D., & Horn, R.A. (1986). Reliability and validity of the severity of illness index. Medical Care, 24, 159-178.

Horn, S.D., Sharkey, P.D., & Bertram, D., (1983). Measuring severity of illness: Homogeneous case mix groups. Medical Care, XXI, 14-30.

Horn, S. D., Sharkey, P. D., Buckle, J. M., Backofen, J. E., Averill, R.F., & Horn, R. A. (1991). The relationship between severity of illness and hospital length of stay and mortality. Medical Care, 29(4), 305-317.

International Food Information Council Foundation (1999). Better eating for better aging. Retrieved from http://www.ific.org/proactive/newsroom/release.vtml?id=18320&PROACTIVE_ID=cecfcfcbcdcccec9cac5cecfcfcfc5cecfcac6ccc6cfcec9c9c5cf

Johnson, T.P., Stallones, L., Garrity, T.F., & Marx, M. (1990). Components of self-rated health among adults: Analysis of multiple data sources. International Quarterly of Community Education, 11, 29-41.

Johnson, R.J., & Wolinsky, F.D. (1993). The structure of health status among older adults: Disease, disability, functional limitation, and perceived health. Journal of Health and Social Behaviour, 34, 105-121.

Johnson, R.J., & Wolinsky, F.D. (1994). Gender, race, & health: The structure of health status among older adults. Gerontologist, 34, 24-35.

Katz, I.R. (1996). On the inseparability of mental and physical health in aged persons: Lessons from depression and medical comorbidity. The American Journal of Geriatric Psychiatry, 4, 1-16.

Kay, D.W.K., & Bergmann, K. (1966). Physical disability and mental health in old age: A follow-up of a random sample of elderly people seen at home. Journal of Psychosomatic Research, 10, 3-12.

Knaus, W.A., Draper, E.A., Wagner, D.P., & Zimmerman, J.E. (1985). APACHE II: A severity of disease classification. Critical Care Medicine, 13, 818-829.

Knaus, W.A., Zimmerman, J.E., Wagner, D.P., Draper, E.A., & Lawrence, D.E. (1981). APACHE—Acute Physiology and Chronic Health Evaluation: A physiologically based classification system. Critical Care Medicine, 9, 591-597.

Knight, B.G., Santos, J., Teri, L., & Lawton, M.P. (1998). Introduction: The development of training in clinical geropsychology. In B.G. Knight, L. Teri, P. Wohlford, & J. Santos (Eds.), Mental health services for older adults. (pp. 1-8). Washington, DC: American Psychological Association.

Langley, A.M. (1995). The mortality of mental illness in older age. Reviews in Clinical Gerontology, 5, 103-112.

Linn, B.S., Linn, M.W., & Guriel, L. (1968). Cumulative Illness Rating Scale. Journal Americian Geriatric Society, 16, 622-626.

Lohr, K.N. (1992). Applications of health status assessment measures in clinical practice. Medical Care, 30, MS1-MS14.

Magai, C., Cohen, C., Gomberg, D., Malatesta, C., & Culver, C. (1996). Emotional expression during mid- to late-stage dementia. International Psychogeriatrics, 8, 383-395.

Martin, D.C., Miller, J.K., Kapoor, W., Arena, V.C., & Boller, F. (1987). A controlled study of survival with dementia. Archives of Neurology, 44, 1122-1126.

McDowell, I., & Newell, N. (1996). Measuring health: A guide to rating scales and questionnaires, New York: Oxford University Press.

McHorney, C.A., (1996). Measuring and monitoring general health status in elderly persons: Practical and methodological issues in using SF-36 health survey. Gerontologist, 36, 571-583.

McHorney, C.A., Ware, J.E., Lu, J.F., & Sherbourne, C.D., (1994). The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Medical Care, 32, 40-66.

MedWorks Media (1999). Retrieved from http://www.cme-reviews.com.

Melding, P.S. (1995). How do older people respond to chronic pain? A review of coping with pain and illness in elders. Pain Reviews, 2, 65-75.

Miller, M.D., Pollock, B.G., Rifai, A.H., Paradis, C.F., Perel, J.M., George, C., Stack, J.A., & Reynolds III, C.F. (1991). Longitudinal analysis of nortriptyline side effects in elderly depressed patients. Journal of Geriatric Psychiatry and Neurology, 4, 226-230.

Miller, M.D., Paradis, C.F., Houck, P.R., Rifai, A.H., Mazumdar, S., Pollock, B., Perel, J.M., Frank, E., Reynolds III, C.F. (1996). Chronic medical illness in patients with recurrent major depression. Americian Journal of Geriatric Psychiatry, 4, 281-290.

Miller, M.D., Paradis, C.F., Houck, P.R., Mazumdar, S., Stack, J.A., Rifai, H., Mulsant, B., & Reynolds III, C.F. (1992). Rating chronic medical illness burden in geropsychiatric practice and research: Application of the Cumulative Illness Rating Scale. Psychiatry Research, 41, 237-248.

Mulrow, C.D., Gerety, M.B., Cornell, J.E., Lawrence. V.A. , & Kanten, D.N. (1994). The relationship between disease and function and perceived health in very frail elders. Journal of the American Geriatrics Society, 42, 374-380.

Murphy, E., Smith, R., Lindesay, J., & Slattery, J. (1988). Increased mortality rates in late-life depression. British Journal of Psychiatry, 152, 347-353.

National Advisory Council on Aging (1996). A quick portrait of dementia in Canada. Aging Vignettes, #34. Retrieved September, 1999, from http://www.hc-sc.gc.ca/english/azindex_def.htm#D

National Advisory Council on Aging (1991). Mental Health and Aging (Cat. No. H71-2/1-11-1991E). Ottawa, ON: Author.

Naughton, B.J., Saltzman, S., Priore, R., Reedy, & Mylotte, J. (1999). Using admission characteristics to predict return to the community from a post-acute geriatric evaluation and management unit. Journal of the American Geriatric Society, 47, 1100-1104.

Newhouse, P.A. (1996). Use of serotonin selective reuptake inhibitors in geriatric depression. Journal of Clinical Psychiatry, 57 (Suppl. 5), 12-22.

Noro, A., & Aro, S. (1996). Health-related quality of life among the least dependent institutional elderly compared with the non-institutional elderly population. Quality of life research: An international journal of quality of life aspects of treatment, 5, 355-366.

O'Connor, B.P., & Vallerand, R.J. (1998). Psychological adjustment variables as predictors of mortality among nursing home residents. Psychology and Aging, 13, 368-374.

Oslin, D.W., Streim, J., Katz, I.R., Edell, W.S., & TenHave, T. (2000). Change in disability follows inpatient treatment for late life depression. Journal of the American Geriatric Society, 48, 357-362.

Oxman, T.E. (1996). Antidepressants and cognitive impairment in the elderly. Journal of Clinical Psychiatry, 57 (Suppl. 5), 38-44.

Parmelee, P.A., Thuras, P.D., Katz, I., & Lawton, M.P. (1995). Validation of the Cumulative Illness Rating Scale in a geriatric population. The American Geriatrics Society, 43, 130-137.

Patrick, D.L., & Deyo, R.A. (1989). Generic and disease-specific measures in assessing health status and quality of life. Medical Care, 27, S217-S232.

Pompei, P., Charlson, M.E., Douglas, R.G., Jr. (1988). Clinical assessments as predictors of one-year survival after hospitalizations: Implications for prognostic stratification. Journal of Clinical Epidemiology, 41, 275-284.

Rabins, P.V. (2000). Measuring quality of life in persons with dementia. International Psychogeriatrics, 12 (Suppl. 1), 47-49.

Reisberg, B., Ferris, S.H., DeLeon, M.J., & Crook, T. (1982). The Global Deterioration Scale for assessment of primary degenerative dementia. American Journal of Psychiatry, 139, 1136-1139.

Rockwood, K., Hogan, D.B., & MacKnight, C. (2000). Conceptualisation and measurement of frailty in elderly people. Drugs and Aging, 17(4), 295-302.

Rossberg-Gempton, I.E., & Poole, G.D. (1999). An intergenerational creative dance program for children and frail older adult. Gerontology and Geriatrics Education, 20(2), 49-68.

Rothman, M.L., Hedrick, S., & Inui, T. (1989). The Sickness Impact Profile as a measure of the health status of noncognitively impaired nursing home residents. Medical Care, 27, S157-S167.

Rozzini, R. Bianchetti, A., Franzoni, S., Zanetti, O., & Trabucchi, M. (1991). Social, functional and health status influences on mortality: Consideration of a multidimensional inquiry in a large elderly population. Journal of Cross-Cultural Gerontology, 6, 83-90.

Rozzini, R., Frisoni, G., Bianchetti, A., Zanetti, O., & Trabucchi, M. (1993). Physical performance test and activities of daily living scales in the assessment of health status in elderly people. Journal of the American Geriatrics Society, 41, 1109-1113.

Sadavoy, J., & LeClair, J.K. (1997). Treatment of anxiety disorders in late life. Canadian Journal of Psychiatry, 42 (Suppl 1), 28S-34S.

Schneider, L.S., & Olin, J.T. (1995). Efficacy of acute treatment for geriatric depression. International Psychogeriatrics, 7, 7-25.

Schneck. M.K., Reisberg, B., & Ferris, S. (1982). An overview of current concepts of Alzheimer's disease. American Journal of Psychiatry, 139, 165-173.

Schulz, R. Mittelmark, M., Kronmal, R., Polak, J., Hirsch, C. German, P., & Bookwala, J. (1994). Predictors of perceived health status in elderly men and women. Journal of Aging and Health, 6, 419-447.

Siegel, S. (1956). Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill Book Company, Inc.

Spar, J.E., & LaRue, A. (1997). Concise guide to geriatric psychiatry. City: Publisher.

Sprecker, S. (1987). Illness Index's Dysfunction Dimension Rating Comparisons Between Hospitalized and Non-Hospitalized (University Student) Populations. Unpublished Honours Thesis, University of Saskatchewan, Saskatoon, SK.

Stanley, M.A., Novy, D.M., Bourland, S.L., Beck, J.G., & Averill, P.M. (2001). Assessing older adults with generalized anxiety: a replication and extension. Behaviour Research and Therapy, 39, 221-235.

Statistics Canada (1997). Mortality-Summary list of causes, 1995. (Catalogue No. 84-209-XPB). Ottawa, ON: Minister of Industry.

Statistics Canada (1998a). "Life expectancy at birth, Canada, selected years". (Catalogue No. 82-221-XDE). Retrieved September, 1999, from http://www.statcan.ca/english/kits/life/life1.htm

Statistics Canada (1998b). "Selected highlights from a Portrait of Seniors in Canada: Second Edition". (Catalogue No. 89-519-XPE). Retrieved September, 1999, from http://www.statcan.ca/english/ads/89-519-XPE/link.htm

Steinhagen-Thiessen, E. & Borchelt, M. (1999). Morbidity, medication and functional limitations in very old age. In P.B. Baltes & K.U. Mayer (1999). The Berlin Aging Study: Aging from 70 to 100 (pp. 131-166). New York, NY: Cambridge University Press.

Steer, R.A., Ball, R., Ranieri, W.F., & Beck, A.T. (1999). Dimensions of the Beck Depression Inventory-II in clinically depressed outpatients. Journal of Clinical Psychology, 55, 117-128.

Steer, R.A., Rissmiller, D.J., & Beck, A.T. (2000). Use of the Beck Depression Inventory-II with depressed geriatric inpatients. Behaviour Research and Therapy, 38, 311-318.

Stewart, A.L., Hays, R., & Ware, Jr., J.E. (1988). The MOS short-form general health survey. Medical Care, 26, 724-735.

Streiner, D.L., & Norman, G.R. (1995). Health measurement scales: A practical guide to their development and use (2nd ed.) Oxford: Oxford University Press.

Stuck, A.E., Walthert, J.M., Nikolaus, T., Bula, C.J., Hohmann, C., and Beck, J.C. (1999). Risk factors for functional status decline in community-living elderly people: A systematic literature review. Social Science and Medicine, 48, 445-469.

Thomas, C., Kelman, H.R., Kennedy, G.J., Ahn, C., & Yang, C. (1992). Depressive symptoms and mortality in elderly persons. Journal of Gerontology, 47, S80-87.

Thompson, L.W. (1996). Cognitive-behavioral therapy and treatment for late-life depression. Journal Clinical Psychiatry, 57 (Suppl. 5), 29-37.

Verbrugge, L.M., & Jette, A.M. (1994). The disablement process. Social Science and Medicine, 38, 1-14.

Ware, J.E., & Sherbourne, C.D., (1992). The MOS 36-item short-form health survey (SF-36). Medical Care, 30, 473-483.

Whitelaw, N., & Liang, J. (1991). The structure of the OARS physical health measures. Medical Care, 29, 332-347.

Wilms, H.U., Kanowski, S., & Baltes, M.M. (2000). Limitations in activities of daily living: Towards a better understanding of subthreshold mental disorders in old age. Comprehensive Psychiatry, 41, 19-25.

Wong, D.T., Barrow, P.M., Gomez, M., & McGuire, G.P. (1996). A comparison of the Acute Physiology and Chronic Health Evaluation (APACHE) II score and the Trauma-Injury Severity Score (TRISS) for outcome assessment in intensive care unit trauma patients. Critical Care Medicine, 24, 1642-1648.

Wood, J.B., Evenson, R.C., Cho, D.W., & Hagan, B.J. (1985). Mortality variations among public mental health patients. Acta Psychiatrica Scandinavica, 72, 218-229.

Zimmerman, J.E., Knaus, W.A., Wagner, D.P., Sun, X., Hakim, R.B., & Nystrom, P.O. (1996a). A comparison of risks and outcomes for patients with organ system failure: 1982-1990. Critical Care Medicine, 24, 1633-1641.

Zimmerman, J.E., Wagner, D.P., Seneff, M.G., Becker, R.B., Sun, X., & Knaus, W.A. (1996b). Intensive care unit admissions with cirrhosis: Risk-stratifying patient groups and predicting individual survival. Hepatology, 23, 1393-1401.

APPENDIX A
Severity Of Renal Disease Scale (SORDS)

Anemia

	Men	Women
0 ◯ absent	133 gm/ml or more	117 gm/ml or more
15 ◯ mild	90 to 132 gm/ml or more	90 to 116 gm/ml or more
40 ◯ moderate	60 to 89 gm/ml or more	60 to 89 gm/ml or more
70 ◯ severe	59 gm/ml or less	59 gm/ml or less

Pneumonia
- 0 ◯ absent
- 30 ◯ mild . . . . . patchy airspace consolidation
- 60 ◯ moderate segmental/lobar consolidation
- 80 ◯ severe . . . multiple lobe involvement

Pleuritis
- 0 ◯ absent
- 20 ◯ mild . . . . . pleurisy without effusion
- 35 ◯ moderate pleurisy with small effusion(s)
- 60 ◯ severe . . . pleurisy with large effusion(s)

Pulmonary embolism
- 0 ◯ absent . . insignificant X-ray episode
- 20 ◯ minor . . . X-ray evidence only
- 80 ◯ major . . . X-ray evidence plus clinical symptoms

Pulmonary edema
- 0 ◯ absent
- 30 ◯ mild . . . . . upper lobe vessel changes
- 60 ◯ moderate interstitial edema
- 85 ◯ severe . . . airspace edema

Coronary heart disease (based on evidence from any of the following: (a) EKG - recent or old myocardial damage; (b) positive thalium scan; and/or (c) positive coronary angiogram
- 0 ◯ absent
- 20 ◯ minimal CHD
- 65 ◯ moderate to severe CHD

Angina pectoris
- 0 ◯ absent . . no angina
- 35 ◯ moderate angina present, but only with exertion
- 80 ◯ severe . . . angina present at rest

Hypertension

Age:	45 to 40	41 to 60	61 or older
◯ absent . . .	less than 140/90 . .	less than 160/95 . .	less than 165/100
◯ mild	140/90 to 159/104 .	160/95 to 169/104	165/100 to 174/109
◯ moderate	160/105 to 199/129	170/105 to 199/129	175/110 to 199/129
◯ severe . .	200/130 or higher .	200/130 or higher . .	200/130 or higher

Peripheral ischemia
- 0 ◯ absent . . no evidence of claudication
- 45 ◯ moderate present, but only with exertion
- 70 ◯ severe . . . present at rest

Pericarditis
- 0 ◯ absent
- 60 ◯ present

Pruritis
- 0 ◯ absent
- 5 ◯ mild . . . . . symptoms of itch
- 15 ◯ moderate evidence of scratching or rubbing
- 20 ◯ severe . . . evidence of excoriation

Diabetes
- 0 ◯ absent
- 10 ◯ present but not requiring insulin
- 15 ◯ present and only controllable through insulin injections

Hyperparathyroidism
- 0 ◯ absent
- 10 ◯ present, calcium levels of 11 to 13 mg/100 ml
- 40 ◯ present, calcium levels of more than 13 mg/100 ml

Non-specific nausea and vomiting
- 0 ◯ absent
- 15 ◯ mild . . . . nausea at least once a day
- 30 ◯ moderate vomiting or vomiting and nausea at least once a day

Ascites
- 0 ◯ absent
- 40 ◯ present

Esophagitis
- 0 ◯ absent
- 20 ◯ present

Hepatitis
- 0 ◯ absent
- 50 ◯ present . . any evidence of same

Peptic ulcer
- 0 ◯ absent
- 5 ◯ mild . . . . X-ray evidence only, no symptoms
- 30 ◯ moderate X-ray evidence plus clinical symptoms; slow blood loss
- 80 ◯ severe . . . X-ray evidence plus clinical symptoms; hemorrhage plus perforations

Rapid weight loss (over past six weeks, all dry weights)
- 0 ◯ absent . . no weight loss
- 5 ◯ mild . . . . 5% to 10% of body weight lost
- 40 ◯ moderate 11% to 20% of body weight lost
- 60 ◯ severe . . . more than 20% of body weight lost

Osteo dystrophy
- 0 ◯ absent
- 15 ◯ moderate present, but no fractures
- 65 ◯ severe . . . present with fractures

Aseptic necrosis
- 0 ◯ absent .
- 5 ◯ mild . . . . no symptoms, early X-ray evidence
- 25 ◯ moderate mild intermittent groin pains; X-ray evidence clear
- 70 ◯ severe . . . continuous pain; crutches necessary; arthritis on X-ray

Osteporosis
- 0 ◯ absent
- 15 ◯ moderate present, but no fractures
- 65 ◯ severe . . . present with fractures

Cerebrovascular accident severity (based on Glasgow Coma Scale)
- 0 ◯ absent
- 30 ◯ mild . . . . 1 to 6 points
- 60 ◯ moderate 7 to 12 points
- 85 ◯ severe . . . 13 points or more

Peripheral nerve disease: Scaling instructions - E.M.G. velocities slowest of either the peroneal or posterior tibial nerves.
- 0 ◯ absent . . E.M.G. velocity of 40 m/sec or greater
- 10 ◯ mild . . . . E.M.G. velocity of 35 m/sec to 39 m/sec
- 25 ◯ moderate E.M.G. velocity of 30 m/sec to 34 m/sec
- 55 ◯ severe . . . E.M.G. velocity of 29 m/sec or less

Creatinine clearance problems
- 0 ◯ none . . . . 70 ml per minute or more
- 10 ◯ mild . . . . 21 to 69 ml per minute
- 45 ◯ moderate 5 to 20 ml per minute
- 80 ◯ severe . . . 0 to less than 5 ml per sec - requires dialysis

Glomerular filtration rate problems
- 0 ◯ none . . . . 70 ml per minute or more
- 10 ◯ mild . . . . 21 to 69 ml per minute
- 45 ◯ moderate 5 to 20 ml per minute
- 80 ◯ severe . . . 0 to less than 5 ml per sec - requires dialysis

Urine volume problems
- 0 ◯ none . . . . 1000 ml per day or more
- 10 ◯ mild . . . . 500 ml per day to 999 ml per day
- 45 ◯ moderate 100 ml per day to 499 ml day
- 75 ◯ severe . . . 0 ml per day to 99 ml per day

APPENDIX B
Open-Ended "Other" Category For SOI-G

Please (A) list each additional item and (B) rate the seriousness of disability of each item on a scale of 0 to 100 using the following scale:

0: Absence of the scaled disease

20: Can do usual work but is unable to participate fully in other normal activities
Can travel about community freely. Can walk WITHOUT limitations

40: Cannot work, cannot play fully, but is able to dress, bathe and feed self
Can travel about community freely. Can walk WITHOUT limitations
OR
Can work and can play but for each is limited in amount and kind
Can go outside alone but requires help to get about community.
Can walk WITH limitations
OR
Can do usual work but is unable to participate fully in other normal activities
Can go outside alone but requires help to get about community.
Can walk WITHOUT limitations

60: Cannot work, cannot play fully, may require human help to dress, bathe or feed
Confined to house or requires human assistance to go outside.
May have limitations in ability to walk
OR
Cannot work, cannot play fully, may require human help to dress, bathe and feed self
Confined to house or requires human assistance to go outside.
Can walk WITHOUT limitations
OR
Cannot work, cannot play fully, may require human help to dress, bathe and feed self
Confined to house or requires human assistance to go outside.
Cannot walk but can propel self in wheelchair

80: Cannot work, cannot play fully, may require human help to dress, bathe or feed
Confined to special unit such as intensive care, special treatment or isolation ward
Confined to bed or chair for most or all of the day
OR
Cannot work, cannot play fully, may require human help to dress, bathe or feed
Confined to hospital, nursing home or similar institution.
Confined to bed or chair for most or all of the day

100: Death

Other #1 **Seriousness rating =**
Other #2 **Seriousness rating =**
Other #3 **Seriousness rating =**
Other #4 **Seriousness rating =**
Other #5 **Seriousness rating =**

APPENDIX C
SOI-G Severity Criteria (Study 2)

GERIATRIC SYNDROMES

Dementia Syndrome (not Alzheimer's Disease–see Neurological disorders)

mild	Impairment in Instrumental Activities of Daily Living (e.g., complex activities needed for independent living, including handling of personal finances, preparing meals, using the telephone, shopping, traveling, doing housework, and taking medications).
moderate	Impairment in Basic Activities of Daily Living [e.g., the most basic personal care tasks, including feeding, grooming, toileting, transferring (moving in and out of a bed or chair), eating, dressing, bathing, and motility]
severe	Total Dependence for all Activities of Daily Living

Urinary Incontinence (not associated with Alzheimer's Disease or other dementia)

mild	Minor incontinence only (i.e., no medications or aids required)
moderate	Incontinence requiring medications or other aids
severe	Incontinence not controlled with medications or other aids or requiring catheter.

Malnutrition (based on Subjective Global Assessment)

mild	Well nourished (minimal restriction of food intake and/or absorption with minimal change in function and body weight)
moderate	Moderately malnourished (clear evidence of food restriction with functional changes but little evidence of any changes in body mass)
severe	Severely malnourished (both changes in intake and body mass with poor function).

Falls

mild	Occasional falls (no more than once/year) but normal activity
moderate	Frequent falls (more than once/year) or excessive decline in activity
severe	Frequent falls with complications such as fractures or head injury

Pressure Ulcers

mild	Stage 1 or 2 (through epidermis or dermis)
moderate	Stage 3 (subcutaneous tissue)
severe	Stage 4 (full thickness involving muscle, bone or supporting structures)

Constipation

mild	Constipation with no social or functional impact
moderate	Constipation with social or functional impact
severe	Constipation with fecal impaction or overflow

Vision or Hearing Impairment

mild	Correctable with glasses or hearing aid
moderate	Correctable only with surgery
severe	Irreversible deafness or blindness

NEUROLOGIC

Alzheimer's Disease (based on Global Deterioration Scale)

Stage 2 (Normal Aging)	Cognition:	Subjective deficit, e.g., in name and word recall
	Functioning:	Subjective deficit, e.g., recalling location of objects
Stage 3 (Incipient AD)	Cognition:	Subtle but manifest deficits in cognition
	Functioning:	Decreased performance in complex occupational and social tasks
Stage 4 (Mild AD)	Cognition:	Clearly manifest cognitive deficits
	Functioning:	Decreased capacity in complex activities of daily life (e.g., handling finances, marketing, meal preparation)
Stage 5 (Moderate AD)	Cognition:	Deficits of sufficient magnitude to preclude independent survival
	Functioning:	Decreased capacity to choose proper clothing for the season and the occasion
Stage 6 (Moderate-Severe AD)	Cognition:	Deficits of sufficient magnitude to interfere with the capacity to handle basic activities of daily life
	Functioning:	Progressive impairment in: (a) putting on clothing properly, (b) handling mechanics of bathing, (c) handling mechanics of toileting, (d) urinary continence, (e) fecal continence
Stage 7 (Severe AD)	Cognition:	Deficits of sufficient magnitude to require continuous assistance in managing basic activities of daily life
	Functioning:	Progressive impairment as follows: (a) speech ability limited to approximately a half dozen words in the course of an intensive contact, (b) speech ability limited to a single word in the course of an intensive contact, (c) ambulatory ability lost, (d) ability to sit up lost, (e) ability to smile lost, (f) ability to hold up head lost.

Stroke

mild	Impairment in Instrumental Activities of Daily Living (e.g., complex activities needed for independent living, including handling of personal finances, preparing meals, using the telephone, shopping, traveling, doing housework, and taking medications).
moderate	Impairment in Basic Activities of Daily Living [e.g., the most basic personal care tasks, including feeding, toileting, transferring (moving in and out of a bed or chair), dressing, bathing, and motility]
severe	Total Dependence for all Activities of Daily Living

Parkinson's Disease [1.]

mild	Unilateral symptoms
moderate	Symptoms with functional impact
severe	Wheelchair bound

Peripheral Neuropathy

mild	Mild sensory loss
moderate	Sensorimotor impairment with functional impact

RESPIRATORY

COPD & Asthma

mild	Occasional symptoms
moderate	Daily symptoms
severe	Daily symptoms with functional limitation or ongoing oxygen requirement

Note 1. Severity criteria for PD changed to Hoehn & Yahr Classification system in Study 3

CARDIOVASCULAR

Coronary Heart Disease [2] (based on evidence from any of the following: (a) EKG - recent or old myocardial damage; (b) positive thallium scan; and/or (c) positive coronary angiogram)

minimal CHD
moderate to severe CHD

Angina Pectoris

moderate	Angina present, but only with exertion
severe	Angina present at rest

Heart Failure

mild	Symptoms with moderate exercise
moderate	Symptoms with minimal exercise
severe	Symptoms at rest

Peripheral Vascular Disease

mild	Symptoms with exercise
moderate	Symptoms at rest
severe	Gangrene or amputation

HEMATOLOGIC

Anemia [3]	**Men**	**Women**
mild	90 to 132 gm/ml or more...	90 to 116 gm/ml or more
moderate	60 to 89 gm/ml or more.....	60 to 89 gm/ml or more
severe	59 gm/ml or less.....	59 gm/ml or less

Thomboembolic Disease

mild	Hypercoagulability with infrequent thrombosis
moderate	Frequent or chronic thrombosis causing chronic pain and edema
severe	Chronic marked edema with ulcerations and disability

Myelodysplastic Syndrome

mild	Refractory anemia with or without ringed sideroblasts
moderate	Refractory anemia with excess blasts
severe	Acute myelogenous leukemia

ENDOCRINE

Hypothyroidism

mild	Subclinical: TSH>6 with normal thyroid hormones (T3 & T4)
moderate	Clinical: TSH>6 with below normal thyroid hormones
severe	Myxedema

Diabetes Mellitis

mild	Symptomatic with adequate glycemic control
moderate	End-organ damage without functional impairment
severe	End-organ damage with functional impairment

Osteoporosis

mild	Osteopenia only
moderate	Osteopenia with fractures
severe	Osteopenia with immobility or other complications

Note 2. Coronary Heart Disease removed in Study 3

Note 3. Behavioural criteria added in Study 3

ONCOLOGY

Multiple Myeloma

mild	Stage 1 (normal hemoglobin, calcium, X Rays, and M component)
moderate	Stage 2 - intermediate between Stage 1 and 3
severe	Stage 3 (anemia, hypocalcemia, lytic bone lesions, and high M component)

Chronic Lymphocytic Leukemia

mild	Adenopathy
moderate	Splenomagaly
severe	Anemia or Thrombocytopenia

Lymphomas

mild	Stage 1 or 2a
moderate	Stage 2b or 3a
severe	Stage 3b or 4

Solid Tumor

mild	Localized disease
moderate	Regional disease
severe	Metastatic disease

GASTROINTESTINAL

Gastroesophageal Reflux Disease

mild	Occasional symptoms
moderate	Daily symptoms
severe	Weight loss

Cirrhosis

mild	Hepatomegaly only
moderate	Malnutrition, jaundice, ascites, or edema
severe	Severe portal hypertension (encephalopathy, bleeding abnormalities)

Peptic Ulcer Disease

mild	Occasional symptoms
moderate	Daily symptoms
severe	Obstruction or other complications

MUSCULOSKELETAL/IMMUNE

Rheumatoid Arthritis

mild	Minimal symptoms without deformity or functional impact
moderate	Synovitis or deformity with functional impact
severe	Extraarticular involvement with the potential to threaten life

Osteroarthritis

mild	Minimal symptoms without deformity or functional impact
moderate	Synovitis or deformity with functional impact
severe	Immobility

RENAL/UROLOGIC

Renal Failure (Creatinine clearance problems) [4]

mild	21 to 69 ml per minute
moderate	5 to 20 ml per minute
severe	0 to less than 5 ml per sec - requires dialysis

Benign Prostatic Hypertrophy

mild	Minimal symptoms
moderate	Compensated with medications or other aids
severe	Obstructive renal failure with hydronephrosis

Note 4. Behavioural indicators added in Study 3

APPENDIX D
Background Information, Panel Instructions and Scaling Material (Study 3)

Background on the development of Severity of Illness-Geriatric (SOI-G)

The development of Severity of Illness-Geriatric (SOI-G) involves the adaptation of a previously developed severity of illness instrument, Severity of Renal Disease Scale (SORDS), designed to evaluate the physical health of renal patients (Baltzan et al, 1987, unpublished manuscript). A copy of SORDS is attached for your information.

While completing my predoctoral internship at the Alberta Hospital Edmonton and Glenrose Rehabilitation Hospital Internship Consortium in Edmonton, AB, I began collaborating with Dr. Darryl Rolfson. Dr. Rolfson's expert opinion was solicited with respect to the identification of the chronic diseases that are (1) commonly experienced by elderly persons and (2) that pose a threat to life. Thirty-three diseases and syndromes were selected for the SOI-G. We recognize that we may have omitted some relevant conditions however we intend to include an "other" category to capture any missed conditions. Dr. Rolfson also provided his expert opinion regarding the generally acceptable, objective criteria (based on standard procedures) for classifying the severity of the each disease item into categorizations such as "absent, mild, moderate, or severe". Some of the criteria for SOI-G (e.g., anemia) were taken directly from the original renal scale (i.e., SORDS).

At present, the SOI-G is capable of assessing the severity of each particular disease, but only on an ordinal scale. To express the severity levels of the various diseases/syndromes on a meaningful numerical scale, it is necessary to scale each severity rating of each disease on a common underlying dimension of illness severity. Each level of each disease will be scaled using a modified Delphi technique. This will involve contacting physicians with training in geriatric medicine and asking for judgements of the disability to be associated with the various disease levels.

The scale values assigned by each medical expert for the diseases on SOI-G will be combined with those assigned by the other participants in the scaling process. The responses from the initial step will be summarized and distributed to the participants who will then meet as a group. The medical experts will be asked to discuss their responses and come to an agreement as to the final scale value to be assigned to each disease level.

Severity of Renal Disease Scale (SORDS)

The development of the SOI-G, following the procedures used to develop SORDS, requires that numerically based severity ratings (i.e., weights) be determined for each of the objective diagnostic criteria (i.e., mild, moderate, severe). You have been provided with sample descriptors of the type and degree of disability that was associated with the scale values of 20, 30, 60, and 80 used in the development of SORDS. These values were chosen from a larger group of disabilities scaled on perceived unpleasantness using hospital patients as judges. These sample descriptors illustrate functional limitations associated with different levels of disease.

Please keep in mind that the weights that were assigned for each level of each disease/symptom on SORDS were developed under the assumption that the hypothetical person suffering from the disease was between 40 and 45 years of age. Thus, given the assumption in the present study that the hypothetical person suffering from the disease is between 70 and 75 years of age, it may be appropriate to assign different weights for the SOI-G than for similar items on SORDS.

As shown in the examples, the dysfunction/limitation descriptors associated with each anchor point are intended to be used as general guides; not as strict criteria. If a person actually had the specific cluster of disabilities in the areas of work/play, physical limitations and mobility limitations, the severity rating would be 20, 40 60 etc. It is understood that the severity levels for the various diseases may not be easily converted to a specific disability cluster however, and thus the clusters are presented to give an idea of the types of limitations that would generate such a severity rating.

SEVERITY OF ILLNESS-GERIATRIC (SOI-G)

SCALING INSTRUCTIONS

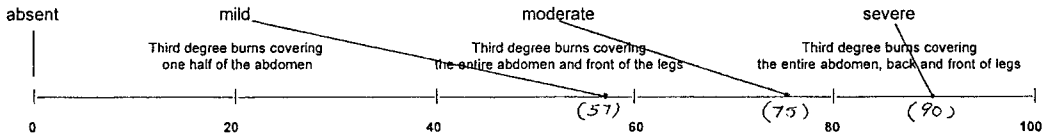
You have been given a copy of SOI-G along with a disability scale. The disability scale has 6 anchor points along a scale ranging from **0 (Absence of the Scaled Disease/Syndrome) to 100 (Death)**. You are asked to scale each level of each disease/syndrome item from SOI-G along this 0 to 100 scale in terms of how dysfunctional the disease is, using the anchor points as guides to level of severity. Judges will be asked to scale the level of disability associated with the severity levels of the diseases under the following assumptions:

(1) that the disease being rated was the only disease present
(2) that the person suffering from the disease was between 70 and 75 years of age
(3) that the disability rating to be assigned to a particular disease severity level should be that associated with the approximate middle of the anticipated range of increasing dysfunction that would result from that severity level.

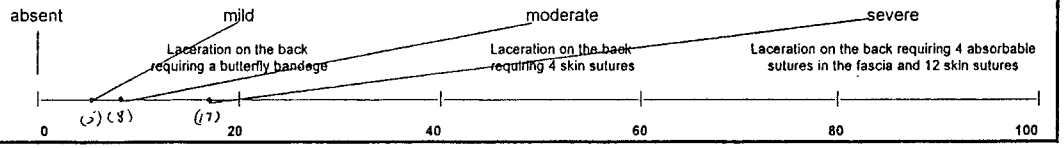
Please make sure that your judgment is based ONLY on the hypothetical patient's current condition. Assume that the disease being scaled is the only disease present.
In addition, do not base your judgement upon a patient's prognosis, but instead upon your perception of the potential level of disability that would be caused by the particular disease.

Using the rating scale below for each disease or syndrome,
Please rate the seriousness of each level (i.e., mild, moderate, severe)
For each of the diseases/syndromes listed by drawing a line from EACH LEVEL of EACH DISEASE/SYNDROME
Onto the 0 (Absence of the disease) to 100 (Death) scale (*see Examples*).
You may use any part of the line and not just the parts physically under the description of mild, moderate, severe.

EXAMPLE #1:
Third Degrees Burns



EXAMPLE #2:
Laceration



SEVERITY OF ILLNESS-GERIATRIC (SOI-G)

GERIATRIC SYNDROMES

	absent	mild	moderate	severe
Dementia Syndrome (NOT Alzheimer's Disease, see Neurological disorders)		Impairment in Instrumental Activities of Daily Living (e.g., complex activities needed for independent living, including handling of personal finances, preparing meals, using the telephone, shopping, traveling, doing housework, and taking medications)	Impairment in Basic Activities of Daily Living [e.g., the most basic personal care tasks, including feeding, grooming, toileting, transferring (moving in and out of a bed or chair), eating, dressing, bathing, motility]	Total dependence for all Activities of Daily Living
	0	20	40 60	80 100
Urinary Incontinence (not associated with Alzheimer's Disease)	absent	mild	moderate	severe
		Minor incontinence only (i.e., no medications or aids required)	Urinary incontinence compensated with medications or other aids	Urinary incontinence not controlled with medications or other aids or requiring catheter
	0	20	40 60	80 100
Malnutrition (based on Subjective Global Assessment)	absent	mild	moderate	severe
		Well nourished (minimal restriction of food intake and/or absorption)	Moderately nourished (clear evidence of food restriction with functional changes but little evidence of any changes in body mass)	Severely malnourished (both changes in intake and body mass with poor function)
	0	20	40 60	80 100
Falls	absent	mild	moderate	severe
		Occasional falls (no more than once/year) but normal activity	Frequent falls (more than once/year) or excessive decline in activity	Frequent falls with complications such as fractures or head injury
	0	20	40 60	80 100

COMMENTS

Example of Scaling Task (Study 3)

Disability Scale used in the development of SORDS

0: Absence of the scaled disease

20: Can do usual work but is unable to participate fully in other normal activities
Can travel about community freely
Can walk WITHOUT limitations

40: Cannot work, cannot play fully, but is able to dress, bathe and feed self
Can travel about community freely
Can walk WITHOUT limitations

OR

Can work and can play but for each is limited in amount and kind
Can go outside alone but requires help to get about community
Can walk WITH limitations

OR

Can do usual work but is unable to participate fully in other normal activities
Can go outside alone but requires help to get about community
Can walk WITHOUT limitations

60: Cannot work, cannot play fully, and may require human help to dress, bathe or feed
Confined to house or requires human assistance to go outside
May have limitations in ability to walk

OR

Cannot work, cannot play fully, may require human help to dress, bathe and feed self
Confined to house or requires human assistance to go outside
Can walk WITHOUT limitations

OR

Cannot work, cannot play fully, may require human help to dress, bathe and feed self
Confined to house or requires human assistance to go outside
Cannot walk but can propel self in wheelchair

80: Cannot work, cannot play fully, and may require human help to dress, bathe or feed
Confined to special unit such as intensive care, special treatment or isolation ward
Confined to bed or chair for most or all of the day

OR

Cannot work, cannot play fully, may require human help to dress, bathe or feed
Confined to hospital, nursing home or similar institution
Confined to bed or chair for most or all of the day

100: Death

APPENDIX E
Individual Rating Summary

Disease or Syndrome	Rater #	Mild	Moderate	Severe
GERIATRIC SYNDROMES				
Dementia				
	1	40	68	90
	2	40	70	90
	3	44	68	94
	4	40	60	80
	5	45	70	95
MEAN		**41.8**	**67.2**	**89.8**
MEDIAN		**40**	**68**	**90**
Urinary Incontinence				
	1	2	7	30
	2	25	45	70
	3	27	47	70
	4	9	49	62
	5	20	38	62
MEAN		**16.6**	**37.2**	**58.8**
MEDIAN		**20**	**45**	**62**
Malnutrition				
	1	2	10	40
	2	10	40	70
	3	8	40	62
	4	20	40	88
	5	7	25	65
MEAN		**9.4**	**31**	**65**
MEDIAN		**8**	**40**	**65**
Falls				
	1	3	14	60
	2	20	50	95
	3	10	50	80
	4	3	50	97
	5	10	53	72
MEAN		**9.2**	**43.4**	**80.8**
MEDIAN		**10**	**50**	**80**

GERIATRIC SYNDROMES (continued)

Disease or Syndrome		Mild	Moderate	Severe
Pressure Ulcers				
	1	10	30	50
	2	10	30	60
	3	20	44	82
	4	20	62	94
	5	7	21	45
MEAN		**13.4**	**37.4**	**66.2**
MEDIAN		**10**	**30**	**60**
Constipation				
	1	2	30	40
	2	10	40	68
	3	5	37	50
	4	4	20	47
	5	3	23	50
MEAN		**4.8**	**30**	**51**
MEDIAN		**4**	**30**	**50**
Vision or Hearing Impairment				
	1	2	20	50
	2	10	50	70
	3	19	30	65
	4	1	40	72
	5	2	22	62
MEAN		**6.8**	**32.4**	**63.8**
MEDIAN		**2**	**30**	**65**

NEUROLOGIC

Alzheimer's Disease							
	1	5	25	44	75	85	95
	2	8	22	44	65	75	91
	3	30	42	54	70	82	95
	4	1	5	20	44	74	99
	5	10	22	45	57	71	90
MEAN		**10.8**	**23.2**	**41.4**	**62.2**	**77.4**	**94**
MEDIAN		**8**	**22**	**44**	**65**	**75**	**95**

NEUROLOGIC (continued)				
Disease or Syndrome		**Mild**	**Moderate**	**Severe**
Stroke				
	1	35	60	94
	2	43	66	92
	3	44	75	93
	4	31	60	91
	5	43	75	90
MEAN		**39.2**	**67.2**	**92**
MEDIAN		**43**	**66**	**92**
Parkinson's Disease				
	1	29	65	89
	2	25	60	85
	3	30	65	94
	4	2	40	80
	5	22	51	65
MEAN		**21.6**	**56.2**	**82.6**
MEDIAN		**25**	**60**	**85**
Peripheral Neuropathy				
	1	20	53	
	2	21	51	
	3	21	44	
	4	2	40	
	5	7	31	
MEAN		**14.2**	**43.8**	
MEDIAN		**20**	**44**	
RESPIRATORY				
COPD & Asthma				
	1	7	50	71
	2	27	45	80
	3	37	58	70
	4	7	48	87
	5	13	35	52
MEAN		**18.2**	**47.2**	**72**
MEDIAN		**13**	**48**	**71**

CARDIOVASCULAR

Disease or Syndrome		Mild	Moderate	Severe
Angina Pectoris				
	1		30	71
	2		55	90
	3		35	79
	4		40	98
	5		22	47
MEAN			**36.4**	**77**
MEDIAN			**35**	**79**
Heart Failure				
	1	30	69	88
	2	40	60	90
	3	35	60	85
	4	20	61	95
	5	17	30	56
MEAN		**28.4**	**56**	**82.8**
MEDIAN		**30**	**60**	**88**
Peripheral Vascular Disease				
	1	24	56	90
	2	29	62	84
	3	25	60	85
	4	20	60	97
	5	20	38	62
MEAN		**23.6**	**55.2**	**83.6**
MEDIAN		**24**	**60**	**85**

HEMATOLOGIC

Disease or Syndrome		Mild	Moderate	Severe
Anemia				
	1	11	44	94
	2	18	57	80
	3	28	50	74
	4	8	54	87
	5	3	21	38
MEAN		**13.6**	**45.2**	**74.6**
MEDIAN		**11**	**50**	**80**
Myelodysplastic Syndrome				
	1	5	69	93
	2	23	51	85
	3	28	50	74
	4	11	46	98
	5	5	23	50
MEAN		**14.4**	**47.8**	**80**
MEDIAN		**11**	**50**	**85**
Thromboembolic Disease				
	1	20	60	75
	2	36	60	75
	3	30	50	70
	4	5	40	80
	5	30	52	73
MEAN		**24.2**	**52.4**	**74.6**
MEDIAN		**30**	**52**	**75**

ENDOCRINE

Hypothyroidism				
	1	3	30	88
	2	16	32	61
	3	5	28	50
	4	1	29	80
	5	4	28	47
MEAN		**5.8**	**29.4**	**65.2**
MEDIAN		**4**	**29**	**61**
Diabetes Mellitus				
	1	8	67	85
	2	27	40	65
	3	24	30	74
	4	5	51	78
	5	12	37	53
MEAN		**15.2**	**45**	**71**
MEDIAN		**12**	**40**	**74**

ENDOCRINE (continued)				
Disease or Syndrome		**Mild**	**Moderate**	**Severe**
Osteoporosis				
	1	4	71	85
	2	18	65	85
	3	10	50	72
	4	3	42	80
	5	6	55	66
MEAN		**8.2**	**56.6**	**77.6**
MEDIAN		**6**	**55**	**80**
ONCOLOGY				
Multiple Myeloma				
	1	29	68	89
	2	23	41	80
	3	43	67	95
	4	2	40	80
	5	10	33	45
MEAN		**21.4**	**49.8**	**77.8**
MEDIAN		**23**	**41**	**80**
Chronic Lymphocytic Leukemia				
	1	9	31	80
	2	24	52	80
	3	38	52	89
	4	2	30	60
	5	5	12	23
MEAN		**15.6**	**35.4**	**66.4**
MEDIAN		**9**	**31**	**80**
Lymphomas				
	1	30	69	89
	2	25	50	80
	3	30	50	88
	4	7	40	79
	5	5	20	38
MEAN		**19.4**	**45.8**	**74.8**
MEDIAN		**25**	**50**	**80**

ONCOLOGY (continued)

Disease or Syndrome		Mild	Moderate	Severe
Solid Tumor				
	1	20	60	93
	2	25	51	87
	3	42	64	94
	4	8	53	89
	5	17	31	42
MEAN		**22.4**	**51.8**	**81**
MEDIAN		**20**	**53**	**89**

GASTROINTESTINAL

Gastroesophageal Reflux Disease				
	1	3	9	50
	2	13	35	60
	3	21	40	63
	4	2	33	80
	5	5	17	21
MEAN		**8.8**	**26.8**	**54.8**
MEDIAN		**5**	**33**	**60**
Cirrhosis				
	1	28	66	95
	2	15	50	93
	3	22	53	76
	4	1	52	99
	5	10	42	71
MEAN		**15.2**	**52.6**	**86.8**
MEDIAN		**15**	**52**	**93**
Peptic Ulcer Disease				
	1	5	40	87
	2	15	51	79
	3	17	68	94
	4	7	40	88
	5	5	21	38
MEAN		**9.8**	**44**	**77.2**
MEDIAN		**7**	**40**	**87**

MUSCULOSKELETAL/IMMUNE

Disease or Syndrome		Mild	Moderate	Severe
Rheumatoid Arthritis				
	1	15	47	87
	2	21	51	89
	3	28	52	94
	4	1	31	96
	5	15	31	58
MEAN		**16**	**42.4**	**84.8**
MEDIAN		**15**	**47**	**89**
Osteoarthritis				
	1	10	52	85
	2	20	41	81
	3	30	50	90
	4	1	29	80
	5	10	45	66
MEAN		**14.2**	**43.4**	**80.4**
MEDIAN		**10**	**45**	**81**

RENAL/UROLOGIC

Renal Failure				
	1	8	20	89
	2	18	57	80
	3	10	53	97
	4	3	53	91
	5	8	28	45
MEAN		**9.4**	**42.2**	**80.4**
MEDIAN		**8**	**53**	**89**
Benign Prostatic Hypertrophy				
	1	8	17	70
	2	20	35	88
	3	23	44	73
	4	5	34	88
	5	12	21	41
MEAN		**13.6**	**30.2**	**72**
MEDIAN		**12**	**34**	**73**

APPENDIX F
Severity Of Illness-Geriatric (SOI-G) with assigned weights

For each of the following diseases/syndromes, record the patient's latest test results by placing a check (✓) in the appropriate category.

GERIATRIC SYNDROMES

Dementia Syndrome (not Alzheimer's Disease–see Neurological disorders)

40	mild	Impairment in Instrumental Activities of Daily Living (e.g., complex activities needed for independent living, including handling of personal finances, preparing meals, using the telephone, shopping, traveling, doing housework, and taking medications).
68	moderate	Impairment in Basic Activities of Daily Living [e.g., the most basic personal care tasks, including feeding, grooming, toileting, transferring (moving in and out of a bed or chair), eating, dressing, bathing, and motility]
90	severe	Total Dependence for all Activities of Daily Living

Urinary Incontinence (not associated with Alzheimer's Disease or other dementia)

20	mild	Minor incontinence only (i.e., no medications or aids required)
45	moderate	Incontinence requiring medications or other aids
62	severe	Incontinence not controlled with medications or other aids or requiring catheter.

Malnutrition (based on Subjective Global Assessment)

8	mild	Well nourished (minimal restriction of food intake and/or absorption with minimal change in function and body weight)
40	moderate	Moderately malnourished (clear evidence of food restriction with functional changes but little evidence of any changes in body mass)
65	severe	Severely malnourished (both changes in intake and body mass with poor function).

Falls

10	mild	Occasional falls (no more than once/year) but normal activity
50	moderate	Frequent falls (more than once/year) or excessive decline in activity
85	severe	Frequent falls with complications such as fractures or head injury

Pressure Ulcers

10	mild	Stage 1 or 2 (through epidermis or dermis)
30	moderate	Stage 3 (subcutaneous tissue)
60	severe	Stage 4 (full thickness involving muscle, bone or supporting structures)

Constipation

4	mild	Constipation with no social or functional impact
30	moderate	Constipation with social or functional impact
60	severe	Constipation with fecal impaction or overflow

Vision or Hearing Impairment

10	mild	Correctable with glasses or hearing aid
30	moderate	Correctable only with surgery
70	severe	Irreversible deafness or blindness

NEUROLOGIC

Alzheimer's Disease (based on Global Deterioration Scale)

0	absent	Stage 1	
5	Stage 2 (Normal Aging)	Cognition:	Subjective deficit, e.g., in name and word recall
		Functioning:	Subjective deficit, e.g., in recalling the location of objects
22	Stage 3 (Incipient AD)	Cognition:	Subtle but manifest deficits in cognition
		Functioning:	Decreased performance in complex occupational and social tasks
44	Stage 4 (Mild AD)	Cognition:	Clearly manifest cognitive deficits
		Functioning:	Decreased capacity in complex activities of daily life (e.g., handling finances, marketing, meal preparation)
65	Stage 5 (Moderate AD)	Cognition:	Deficits of sufficient magnitude to preclude independent survival
		Functioning:	Decreased capacity to choose proper clothing for the season and the occasion
75	Stage 6 (Moderate-Severe AD)	Cognition:	Deficits of sufficient magnitude to interfere with the capacity to handle basic activities of daily life
		Functioning:	Progressive impairment in: (a) putting on clothing properly, (b) handling mechanics of bathing, (c) handling mechanics of toileting, (d) urinary continence, (e) fecal continence
95	Stage 7 (Severe AD)	Cognition:	Deficits of sufficient magnitude to require continuous assistance in managing basic activities of daily life
		Functioning:	Progressive impairment as follows: (a) speech ability limited to approximately a half dozen words in the course of an intensive contact, (b) speech ability limited to a single word in the course of an intensive contact, (c) ambulatory ability lost, (d) ability to sit up lost, (e) ability to smile lost, (f) ability to hold up head lost.

Stroke

43	mild	Impairment in Instrumental Activities of Daily Living (e.g., complex activities needed for independent living, including handling of personal finances, preparing meals, using the telephone, shopping, traveling, doing housework, and taking medications).
66	moderate	Impairment in Basic Activities of Daily Living [e.g., the most basic personal care tasks, including feeding, toileting, transferring (moving in and out of a bed or chair), dressing, bathing, and motility]
92	severe	Total Dependence for all Activities of Daily Living

Parkinson's Disease (based on Hoehn & Yahr Classification)

20	Stage 1	Unilateral involvement
40	Stage 2	Bilateral symptoms with bilateral impairment
60	Stage 3	Bilateral symptoms with some postural instabilities
75	Stage 4	Severe disability but able to walk or stand unassisted
85	Stage 5	Wheelchair bound or bedridden unless aided

Peripheral Neuropathy

20	mild	Mild sensory loss
44	moderate	Sensorimotor impairment with functional impact

RESPIRATORY

COPD & Asthma

15	mild	Occasional symptoms
48	moderate	Daily symptoms
75	severe	Daily symptoms with functional limitation or ongoing oxygen requirement

CARDIOVASCULAR

Angina Pectoris

50	moderate	Angina present, but only with exertion
91	severe	Angina present at rest

Heart Failure

40	mild	Symptoms with moderate exercise
60	moderate	Symptoms with minimal exercise
88	severe	Symptoms at rest

Peripheral Vascular Disease

24	mild	Symptoms with exercise
60	moderate	Symptoms at rest
85	severe	Gangrene or amputation

HEMATOLOGIC

Anemia

		Men	Women
11	mild	90 to 132 gm/ml or more	90 to 116 gm/ml or more Chronic mild anemia without symptoms
50	moderate	60 to 89 gm/ml or more	60 to 89 gm/ml or more Chronic mild anemia causing symptoms (dyspnea, fatigue, dizziness, angina)
80	severe	59 gm/ml or less	59 gm/ml or less Congestive heart failure

Thromboembolic Disease

30	mild	Hypercoagulability with infrequent thrombosis
52	moderate	Frequent or chronic thrombosis causing chronic pain and edema
75	severe	Chronic marked edema with ulcerations and disability

Myelodysplastic Syndrome

15	mild	Refractory anemia with or without ringed sideroblasts
45	moderate	Refractory anemia with excess blasts
90	severe	Acute myelogenous leukemia

ENDOCRINE

Hypothyroidism

10	mild	Subclinical: TSH>6 with normal thyroid hormones (T3 & T4)
29	moderate	Clinical: TSH>6 with below normal thyroid hormones
61	severe	Myxedema

Diabetes Mellitis

12	mild	Asymptomatic with adequate glycemic control
40	moderate	End-organ damage without functional impairment
74	severe	End-organ damage with functional impairment

Osteoporosis

6	mild	Osteopenia only
55	moderate	Osteopenia with fractures
80	severe	Osteopenia with immobility or other complications

ONCOLOGY

Multiple Myeloma

10	mild	Stage 1 (normal hemoglobin, calcium, X Rays, and M component)
40	moderate	Stage 2 - intermediate between Stage 1 and 3
75	severe	Stage 3 (anemia, hypercalcemia, lytic bone lesions, and high M component)

Chronic Lymphocytic Leukemia

15	mild	Adenopathy
30	moderate	Splenomagaly
80	severe	Anemia or Thrombocytopenia

Lymphomas

27	mild	Stage 1 or 2a
50	moderate	Stage 2b or 3a
88	severe	Stage 3b or 4

Solid Tumor

20	mild	Localized disease
53	moderate	Regional disease
90	severe	Metastatic disease

GASTROINTESTINAL

Gastroesophageal Reflux Disease

5	mild	Occasional symptoms
27	moderate	Daily symptoms
60	severe	Weight loss

Cirrhosis

10	mild	Hepatomegaly only
52	moderate	Malnutrition, jaundice, ascites, or edema
93	severe	Severe portal hypertension (encephalopathy, bleeding abnormalities)

Peptic Ulcer Disease

7	mild	Occasional symptoms
40	moderate	Daily symptoms
87	severe	Obstruction or other complications

MUSCULOSKELETAL/IMMUNE

Rheumatoid Arthritis

15	mild	Minimal symptoms without deformity or functional impact
47	moderate	Synovitis or deformity with functional impact
89	severe	Extraarticular involvement with the potential to threaten life

Osteroarthritis

12	mild	Minimal symptoms without deformity or functional impact
45	moderate	Synovitis or deformity with functional impact
81	severe	Immobility

RENAL/UROLOGIC

Renal Failure (Creatinine clearance problems)

8	mild	21 to 69 ml per minute Chronic renal insufficiency without symptoms
53	moderate	5 to 20 ml per minute/Uremia
89	severe	0 to less than 5 ml per min - requires dialysis Obstructive renal failure with hydronephrosis

Benign Prostatic Hypertrophy

18	mild	Minimal symptoms
34	moderate	Requiring medications or other aids
73	severe	Obstructive renal failure with hydronephrosis

APPENDIX G
Cumulative Illness Rating Scale-Geriatric (CIRS-G)

Instructions: **Please write brief descriptions of the medical problem(s) that justified the endorsed score on the line following each item. (Use the reverse side for more writing space).**

Rating Strategy

0 **No problem**
1 **Current mild problem or past significant problem**
2 **Moderate disability or morbidity/requires "first line" therapy**
3 **Severe/constant significant disability/ "uncontrollable" chronic problems**
4 **Extremely severe/immediate treatment required/end organ failure/severe impairment function**

	Score
Heart	_____
Vascular	_____
Respiratory	_____
Eyes, ears, nose, throat, and larynx	_____
Upper gastrointestinal tract	_____
Lower gastrointestinal tract	_____
Liver	_____
Renal	_____
Genito-urinary	_____
Musculoskeletal/integument	_____
Neurological	_____
Endocrine/metabolic and breast	_____
Psychiatric Illness	_____

APPENDIX H
Demographic Questionnaire
Study 4

1. Date of Birth _____

2. Gender (circle): Male Female

3. Marital Status (circle):
    1. Single/Never Married
    2. Married
    3. Divorced
    4. Common-Law
    5. Widowed

4. Living situation (circle):
    a. In own home/apartment
    b. In own home/apartment with outside assistance (e.g., HomeCare)
    c. Senior's Housing (please name) _____
    d. Nursing Home (please name) _____

5. Ethnicity _____

APPENDIX I
Global Severity Ratings (Before and After)

Please give your overall consideration of how severely ill the patient was using the following 0 to 100 scale.

Seriousness rating after initial review of file but BEFORE completion of other instruments _____

0: Absence of the scaled disease

20: Can do usual work but is unable to participate fully in other normal activities
Can travel about community freely.
Can walk WITHOUT limitations

40: Cannot work, cannot play fully, but is able to dress, bathe and feed self
Can travel about community freely.
Can walk WITHOUT limitations

60: Cannot work, cannot play fully, may require human help to dress, bathe or feed
Confined to house or requires human assistance to go outside.
May have limitations in ability to walk

80: Cannot work, cannot play fully, may require human help to dress, bathe or feed
Confined to special unit such as intensive care, special treatment or isolation ward
Confined to bed or chair for most or all of the day

100: Death

Please give your overall consideration of how severely ill the patient was using the following 0 to 100 scale.

Seriousness rating
AFTER completion of other instruments_____

0: Absence of the scaled disease

20: Can do usual work but is unable to participate fully in other normal activities
Can travel about community freely.
Can walk WITHOUT limitations

40: Cannot work, cannot play fully, but is able to dress, bathe and feed self
Can travel about community freely.
Can walk WITHOUT limitations

60: Cannot work, cannot play fully, may require human help to dress, bathe or feed
Confined to house or requires human assistance to go outside.
May have limitations in ability to walk

80: Cannot work, cannot play fully, may require human help to dress, bathe or feed
Confined to special unit such as intensive care, special treatment or isolation ward
Confined to bed or chair for most or all of the day

100: Death

APPENDIX J
Medical Outcomes Study (MOS) SF-36 Health Survey (SF-36)

NB: The SF-36 is copyright protected; this copy has been prepared solely as part of this submission. Copies of the SF-36 used in the present research were purchased from The Medical Outcomes Trust Inc.

RAND–36 Health Status Inventory Question/Answer Sheet

THE PSYCHOLOGICAL CORPORATION
Harcourt Brace & Company
SAN ANTONIO
Orlando • Boston • New York • Chicago • San Francisco • Atlanta • Dallas
San Diego • Philadelphia • Austin • Fort Worth • Toronto • London • Sydney

Name	Date of Birth	Age	Sex ☐F ☐M	Date of Testing

Do you currently have a physical disability/condition? ☐ No ☐ Yes If yes, please specify:

The following questions ask about your health as it relates to how you have felt and gone about your daily activities in the past 4 weeks. Circle one number for each item. Please be sure to answer *all* of the questions.

1. In general, would you say your health is:	Excellent	1
	Very good	2
	Good	3
	Fair	4
	Poor	5
2. Compared to 1 year ago, how would you rate your health in general now?	Much better now than 1 year ago	1
	Somewhat better now than 1 year ago	2
	About the same as 1 year ago	3
	Somewhat worse now than 1 year ago	4
	Much worse now than 1 year ago	5

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?	YES, LIMITED A LOT	YES, LIMITED A LITTLE	NO, NOT LIMITED AT ALL
3. Vigorous activities, such as running, lifting heavy objects, participating in strenuous sports	1	2	3
4. Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf	1	2	3
5. Lifting or carrying groceries	1	2	3
6. Climbing several flights of stairs	1	2	3
7. Climbing one flight of stairs	1	2	3
8. Bending, kneeling, or stooping	1	2	3
9. Walking more than a mile	1	2	3
10. Walking several blocks	1	2	3
11. Walking one block	1	2	3
12. Bathing or dressing yourself	1	2	3

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

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)?	YES	NO
17. Cut down the amount of time you spent on work or other activities	1	2
18. Accomplished less than you would like	1	2
19. Didn't do work or other activities as carefully as usual	1	2

Distributed by The Psychological Corporation. 1 2 3 4 5 6 7 8 9 10 11 12 A B C D E 0158132157

SF-36 (Page 2)

20. 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?	Not at all	1
	Slightly	2
	Moderately	3
	Quite a bit	4
	Extremely	5
21. How much **bodily** pain have you had during the **past 4 weeks**?	None	1
	Very mild	2
	Mild	3
	Moderate	4
	Severe	5
	Very severe	6
22. During the **past 4 weeks**, how much did pain interfere with your normal work (including both work outside the home and housework)?	Not at all	1
	A little bit	2
	Moderately	3
	Quite a bit	4
	Extremely	5

The following 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 time during the **past 4 weeks:**	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
23. Did you feel full of pep?	1	2	3	4	5	6
24. Have you been a very nervous person?	1	2	3	4	5	6
25. Have you felt so down in the dumps that nothing could cheer you up?	1	2	3	4	5	6
26. Have you felt calm and peaceful?	1	2	3	4	5	6
27. Did you have a lot of energy?	1	2	3	4	5	6
28. Have you felt downhearted or blue?	1	2	3	4	5	6
29. Did you feel worn out?	1	2	3	4	5	6
30. Have you been a happy person?	1	2	3	4	5	6
31. Did you feel tired?	1	2	3	4	5	6

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

How true or false is **each** of the following statements for you?	DEFINITELY TRUE	MOSTLY TRUE	DON'T KNOW	MOSTLY FALSE	DEFINITELY FALSE
33. I seem to get sick a little easier than other people.	1	2	3	4	5
34. I am as healthy as anybody I know.	1	2	3	4	5
35. I expect my health to get worse.	1	2	3	4	5
36. My health is excellent.	1	2	3	4	5

APPENDIX K
Beck Depression Inventory-2nd Edition (BDI-II)

NB: The BDI-II is copyright protected; this copy has been prepared solely as part of this submission. Copies of the BDI-II used in the present research were purchased from The Psychological Corporation

BDI-II

Date:

Name: _____ Marital Status: _____ Age: _____ Sex: _____

Occupation: _____ Education: _____

Instructions: This questionnaire consists of 21 groups of statements. Please read each group of statements carefully, and then pick out the **one statement** in each group that best describes the way you have been feeling during the **past two weeks, including today**. Circle the number beside the statement you have picked. If several statements in the group seem to apply equally well, circle the highest number for that group. Be sure that you do not choose more than one statement for any group, including Item 16 (Changes in Sleeping Pattern) or Item 18 (Changes in Appetite).

1. Sadness
- 0 I do not feel sad.
- 1 I feel sad much of the time.
- 2 I am sad all the time.
- 3 I am so sad or unhappy that I can't stand it.

2. Pessimism
- 0 I am not discouraged about my future.
- 1 I feel more discouraged about my future than I used to be.
- 2 I do not expect things to work out for me.
- 3 I feel my future is hopeless and will only get worse.

3. Past Failure
- 0 I do not feel like a failure.
- 1 I have failed more than I should have.
- 2 As I look back, I see a lot of failures.
- 3 I feel I am a total failure as a person.

4. Loss of Pleasure
- 0 I get as much pleasure as I ever did from the things I enjoy.
- 1 I don't enjoy things as much as I used to.
- 2 I get very little pleasure from the things I used to enjoy.
- 3 I can't get any pleasure from the things I used to enjoy.

5. Guilty Feelings
- 0 I don't feel particularly guilty.
- 1 I feel guilty over many things I have done or should have done.
- 2 I feel quite guilty most of the time.
- 3 I feel guilty all of the time.

6. Punishment Feelings
- 0 I don't feel I am being punished.
- 1 I feel I may be punished.
- 2 I expect to be punished.
- 3 I feel I am being punished.

7. Self-Dislike
- 0 I feel the same about myself as ever.
- 1 I have lost confidence in myself.
- 2 I am disappointed in myself.
- 3 I dislike myself.

8. Self-Criticalness
- 0 I don't criticize or blame myself more than usual.
- 1 I am more critical of myself than I used to be.
- 2 I criticize myself for all of my faults.
- 3 I blame myself for everything bad that happens.

9. Suicidal Thoughts or Wishes
- 0 I don't have any thoughts of killing myself.
- 1 I have thoughts of killing myself, but I would not carry them out.
- 2 I would like to kill myself.
- 3 I would kill myself if I had the chance.

10. Crying
- 0 I don't cry anymore than I used to.
- 1 I cry more than I used to.
- 2 I cry over every little thing.
- 3 I feel like crying, but I can't.

_____ Subtotal Page 1

Continued on Back

THE PSYCHOLOGICAL CORPORATION®
Harcourt Brace & Company
SAN ANTONIO
Orlando • Boston • New York • Chicago • San Francisco • Atlanta • Dallas
San Diego • Philadelphia • Austin • Fort Worth • Toronto • London • Sydney

0154018392

BDI-II (Page 2)

11. Agitation

0 I am no more restless or wound up than usual.
1 I feel more restless or wound up than usual.
2 I am so restless or agitated that it's hard to stay still.
3 I am so restless or agitated that I have to keep moving or doing something.

12. Loss of Interest

0 I have not lost interest in other people or activities.
1 I am less interested in other people or things than before.
2 I have lost most of my interest in other people or things.
3 It's hard to get interested in anything.

13. Indecisiveness

0 I make decisions about as well as ever.
1 I find it more difficult to make decisions than usual.
2 I have much greater difficulty in making decisions than I used to.
3 I have trouble making any decisions.

14. Worthlessness

0 I do not feel I am worthless.
1 I don't consider myself as worthwhile and useful as I used to.
2 I feel more worthless as compared to other people.
3 I feel utterly worthless.

15. Loss of Energy

0 I have as much energy as ever.
1 I have less energy than I used to have.
2 I don't have enough energy to do very much.
3 I don't have enough energy to do anything.

16. Changes in Sleeping Pattern

0 I have not experienced any change in my sleeping pattern.

1a I sleep somewhat more than usual.
1b I sleep somewhat less than usual.

2a I sleep a lot more than usual.
2b I sleep a lot less than usual.

3a I sleep most of the day.
3b I wake up 1–2 hours early and can't get back to sleep.

17. Irritability

0 I am no more irritable than usual.
1 I am more irritable than usual.
2 I am much more irritable than usual.
3 I am irritable all the time.

18. Changes in Appetite

0 I have not experienced any change in my appetite.

1a My appetite is somewhat less than usual.
1b My appetite is somewhat greater than usual.

2a My appetite is much less than before.
2b My appetite is much greater than usual.

3a I have no appetite at all.
3b I crave food all the time.

19. Concentration Difficulty

0 I can concentrate as well as ever.
1 I can't concentrate as well as usual.
2 It's hard to keep my mind on anything for very long.
3 I find I can't concentrate on anything.

20. Tiredness or Fatigue

0 I am no more tired or fatigued than usual.
1 I get more tired or fatigued more easily than usual.
2 I am too tired or fatigued to do a lot of the things I used to do.
3 I am too tired or fatigued to do most of the things I used to do.

21. Loss of Interest in Sex

0 I have not noticed any recent change in my interest in sex.
1 I am less interested in sex than I used to be.
2 I am much less interested in sex now.
3 I have lost interest in sex completely.

_____ Subtotal Page 2
_____ Subtotal Page 1
_____ Total Score

APPENDIX L
Demographic Questionnaire
Study 5

1 Date of Birth _____

2. Gender (circle): Male Female

3. Marital Status (circle):
   1. Single/Never Married
   2. Married
   3. Divorced
   4. Common-Law
   5. Widowed

4. Highest level of school completed (circle):
   1. Junior High School
   2. High School (Grade 12)
   3. Technical or Business School
   4. 1 or 2 years University
   5. Undergraduate University Degree
   6. Graduate School

5. Usual type of work, even if not working now (may list more than one occupation): _____

6. Living situation (circle):
   a. In own home/apartment
   b. In own home/apartment with outside assistance (e.g., HomeCare)
   c. Senior's Housing (please name) _____
   d. Nursing Home (please name) _____

5. Ethnicity _____

APPENDIX M
Information Sheet and Informed Consent Form

Title of Study

Quantification of severity of illness in geriatric research: An adaptation of the Severity of Renal Disease Scale (SORDS)

Name of Researcher(s)

Lisa D. Berg, MA
Doctoral Candidate
Psychology Department
University of Saskatchewan
(306) 966-6657

Dr. Darryl Rolfson
Specialist in Geriatric Medicine
Northern Alberta Regional Geriatric Program
Glenrose Rehabilitation Hospital
Edmonton, AB T5G 0B7
780-474-8800

David A. Scott, PhD
Dissertation Supervisor
Psychology Department
University of Saskatchewan
(306) 966-6673

What is the purpose of the project?

We are inviting you to take part in a research project about the health of older adults. The information from this project will be used to complete a doctoral degree. It is hoped that what we learn from this project will assist professionals who work with elderly persons coping with physical health problems. Of course, we know that many elderly persons are active and healthy. Therefore, we are inviting all persons over the age of 65 years at Northern Alberta Regional Geriatric Program (NARGP) to take part in the project, not just you and not just other elderly persons who are ill.

What will you be asked to do?

- You are being asked to allow a nurse to collect information from your medical file.
- You are also being asked to allow a trained interviewer to talk with you for about 30 minutes. You will be asked questions about your health (e.g., how much has your health interfered with your usual activities), lifestyle (e.g., level of education), and questions about your mood (e.g., do you sometimes feel sad?). If you would like, you can see the questions before you agree to take part.
- You decision to participate will not affect your ability to receive services at NARGP.
- You will be given any new information that might affect your choice to take part in the study.

What about privacy and confidentiality?

- All information will be held confidential (or private), except when professional codes of ethics or legislation (or the law) requires reporting.
- Your name will not be used. We will use a code number instead of your name.
- Because only group results will be reported, no one person can be identified. Group information will be published in the dissertation paper and may be submitted to professional conferences and journals.
- The information you provide will be kept for at least five years after the study is done. The information will be kept in a locked filing cabinet. Your name or any other identifying facts will not be attached to the information you gave. Your name will also never be used in any presentations or publications of the study results.
- The information gathered for this study may be looked at again in the future to help us answer other study questions. If so, the University of Alberta ethics board will first review the study to ensure the information is used ethically.
- We will give you information about the group results when the study is completed, if you wish. In order to protect the privacy of everyone who joined this project, we will not be able to give information specific to any one person.

What if you change your mind about participating?

- Your participation in this project is voluntary.
- At any time, even after you sign the consent form, you may refuse to answer any questions, may withdraw from the study, or ask that information collected not be used.
- If you decide to withdraw, all information collected will be destroyed.
- If you become tired at any time, you may stop until you feel able to continue.

What if you are bothered by the questions?

- If you have questions about the study, please contact:

  Dr. Darryl Rolfson, Specialist in Geriatric Medicine
  Northern Alberta Regional Geriatric Program
  Glenrose Rehabilitation Hospital
  Edmonton, AB T5G 0B7
  (780) 474-8800

  Or

  Dr. David Scott (Dissertation Supervisor)
  or Lisa Berg-Kolody (Doctoral Candidate)
  Psychology Department
  University of Saskatchewan
  Saskatoon, SK S7N 5A5
  (306) 966-6657

- Concerns regarding this study should be directed to:

  Ms. Karen Turpin, Administrative Assistant
  Health Research Ethics Board
  3-48 Corbett Hall
  Faculty of Rehabilitation Medicine
  University of Alberta
  Edmonton, AB T6G 2G4
  (780) 492-0839

CONSENT FORM

Part 1: Researcher Information	
Name of Principal Investigator:	Lisa Berg-Kolody, Doctoral Candidate
Affiliation:	Psychology Department, University of Saskatchewan
Contact Information:	306-966-6673
Name of Co-Investigator/Supervisor:	Dr. Darryl Rolfson, Specialist Geriatric Medicine
Affiliation:	Northern Alberta Regional Geriatric Program
Contact Information:	780-474-8800

Part 2: Consent of Subject	Yes	No
Do you understand that you have been asked to be in a research study?		
Have you read and received a copy of the attached information sheet?		
Do you understand the benefits and risks involved in taking part in this research study?		
Have you had an opportunity to ask questions and discuss the study?		
Do you understand that you are free to refuse to participate or withdraw from the study at any time? You do not have to give a reason and it will not affect your care.		
Has the issue of confidentiality been explained to you? Do you understand who will have access to your records/information?		

Part 3: Signatures

This study was explained to me by: _____

Date: _____

I agree to take part in this study.

Signature of Research Participant: _____

Printed Name: _____

Witness (if available): _____

Printed Name: _____

I would like to receive a letter explaining the results of the study when the project is completed. YES NO

Please send the letter to the following address:
Name:
Address:
City/Province:
Postal Code

I believe that the person signing this form understands what is involved in the study and voluntarily agrees to participate.

Researcher:_____

Printed Name:_____