

**An Australian MAU/QALY Instrument:  
Rationale and Preliminary Results**

**Dr Graeme Hawthorne**

Senior Research Fellow, Program Evaluation Unit, CHPE

**Professor Jeff Richardson**

Director, Health Economics Unit, CHPE

September, 1995

ISSN 1325 0663

ISBN 1 875677 47 X

## CENTRE PROFILE

The Centre for Health Program Evaluation (CHPE) is a research and teaching organisation established in 1990 to:

- undertake academic and applied research into health programs, health systems and current policy issues;
- develop appropriate evaluation methodologies; and
- promote the teaching of health economics and health program evaluation, in order to increase the supply of trained specialists and to improve the level of understanding in the health community.

The Centre comprises two independent research units, the Health Economics Unit (HEU) which is part of the Faculty of Business and Economics at Monash University, and the Program Evaluation Unit (PEU) which is part of the Department of Public Health and Community Medicine at The University of Melbourne. The two units undertake their own individual work programs as well as collaborative research and teaching activities.

## PUBLICATIONS

The views expressed in Centre publications are those of the author(s) and do not necessarily reflect the views of the Centre or its sponsors. Readers of publications are encouraged to contact the author(s) with comments, criticisms and suggestions.

A list of the Centre's papers is provided inside the back cover. Further information and copies of the papers may be obtained by contacting:

The Co-ordinator  
Centre for Health Program Evaluation  
PO Box 477  
West Heidelberg Vic 3081, Australia  
**Telephone** + 61 3 9496 4433/4434      **Facsimile** + 61 3 9496 4424  
**E-mail** CHPE@BusEco.monash.edu.au

# ACKNOWLEDGMENTS

We would like to thank the Australian Research Council and the Vice-Chancellor's Fund at Monash University for funding this project. We would also like to thank Ms Johanna Cook for her valuable input into the early stages of the research, Dr Richard Osborne and Ms Helen McNeil for their dedicated efforts during data collection and analysis, and the medical staff at St Vincent's Hospital and The University of Melbourne who gave so generously of their time and expertise. Without their support this research would not have been possible. Finally we would like to thank the patients who so willingly filled in questionnaires at times which would have been personally inconvenient.

## ABSTRACT

The paper presents the preliminary stages in the construction of an Australian Multi-attribute Utility Instrument for obtaining preferences regarding quality of life (QoL) for economic evaluation of health services.

The project was undertaken following a review of existing MAU and generic QoL instruments, none of which had been validated for use with an Australian population and most of which did not meet standard psychometric criteria.

The definition of 'health' adopted by the research team came from the World Health Organization's definition. Following a review of 14 MAU/QoL instruments, a model of health-related QoL was constructed. This involved defining five dimensions: illness, independent living, physical ability, psychological state and social interaction.

Focus groups of doctors were used to elicit the content and items appropriate for each dimension. An iterative procedure was followed, during which items were progressively refined until there was an overall consensus about adequate coverage, item content and item language. The item bank was then submitted to a patient sample.

The bulk of the paper outlines the procedures which were used to construct a preliminary version of the MAU-instrument, and describes the psychometric properties of this research tool.

The paper concludes with a brief review of the further research in progress to verify the structure of the preliminary instrument, to develop appropriate weights and to undertake standard tests of validity and reliability.

# TABLE OF CONTENTS

## **1 Introduction**

## **2 Why an Australian MAU?**

2.1 Multi-attribute Non-utility Based Instruments

2.2 MAU Instruments

## **3 Constructing the Australian MAU Instrument**

3.1 Conceptualizing 'Health-related Quality of Life'

3.2 The Preliminary Synthesis

The Dimensions

The Items

## **4 The Preliminary MAU Instrument**

4.1 Preliminary Scale Construction

Initial Item Selection: Sensitivity

Scale Construction

## **5 Future Planned Work**

## **6 Conclusion**

## **References**

## **Appendix Multi—attribute Utility Instrument**

# An Australian MAU/QALY Instrument:

## Rationale and Preliminary Results

### 1 Introduction

Before the development of cost utility analysis, economic evaluation of health services either ignored quality of life (QoL) or treated QoL as an 'intangible' that could be noted and described but not quantified or included as an integral part of the health outcome. Cost utility analysis (CUA) has attempted to overcome this deficit by adopting the quality adjusted life year (QALY) as the unit of output for health benefits in cost effectiveness studies. One of two approaches have been adopted.

First, in a 'holistic' or scenario based approach to measurement, the health states relevant to the evaluation of a health program are described in a series of scenarios. These scenarios are then rated using a scaling device such as the time trade-off or standard gamble to obtain a 'utility' index; an index of the strength of a person's preference for a health state. The index is then used to obtain QALYs. The construction of the health scenarios and the rating exercise both require surveys. Normally, patients who have experienced the health states are consulted for scenario construction and a random population sample used for the rating.

The second approach requires the preliminary construction of a generic multi attribute utility (MAU) QoL instrument which is capable of describing numerous health states and assigning a utility—preference—score to each of these. MAU-instrument construction involves the creation of a descriptive system describing multiple health states, the construction of appropriate items and scales representing the system, and the scaling of the instrument for the determination of utility preference scores for each of the health states described. As these instruments are often capable of describing a very large number of combinations of health states the techniques of multi-attribute modelling are commonly used to infer scores for particular health states. The attraction of the MAU instrument is, *inter alia*, that it obviates the need for the two surveys required by the holistic approach, it imposes a standard, non-arbitrary description on the health states and it allows for the continuous collection of data in longitudinal studies.

To date, Australian CUA's have either adopted the first, holistic, approach to QALY measurement or have assumed that one of the MAU instruments developed overseas can be validly applied in Australia.

The project described in this paper is the construction of an Australian MAU-QALY instrument capable of use in CUA and therefore capable of assisting with decisions about the allocation of resources between competing health related interventions. More specifically the objective is to develop an instrument that can be used in the evaluation of a wide range of programs for the treatment of acute and chronic conditions, ambulatory and institutional care, and public health and prevention programs. The project is not simply to replicate an existing instrument but to create a new descriptive system for an instrument and to scale it using values from a random cross-section of the Australian population.

The purpose of the present paper is threefold. First, it justifies the development of yet another QoL instrument (Section 2). Second, it describes the work already completed and presents some preliminary results from the analysis of patient interviews which were carried out as part of the construction of the instrument's descriptive system (Section 3). The paper ends by briefly outlining the further steps planned for the creation of the first version of the Australian instrument (Section 4).

## **2 Why an Australian MAU?**

The objective of creating a new instrument and not simply rescaling the descriptive system of an existing instrument implies dissatisfaction with the existing instruments. These can be divided into two categories; *viz*, non-MAU instruments that purport to measure QoL in some sense but do not measure the strength of individual preferences, and MAU instruments that explicitly seek to attach utility values to the health states described by the instrument.

### **2.1 Multi—attribute Non—Utility Based Instruments**

Many non-MAU instruments are described in McDowell & Newell (1987) and Bowling (1995). Most are disease or condition specific. In addition some generic instruments have been developed for population health surveys, for the prediction of resource use or clinical outcomes (a review of many instruments is given in Walker & Rosser 1993). Such instruments may produce a health profile in which the different dimension scores remain disaggregated. Alternatively, they may yield a single QoL score through the use of a simple combination rule such as the unweighted addition of positive responses in each of the dimensions. Neither approach satisfies the requirements of economic evaluation. The use of health profiles may result in ambiguous conclusions. A health related intervention may result in an improvement in one dimension of the profile but a deterioration in another. Under these conditions it is difficult to draw firm conclusions about the overall QoL.

Aggregate, non—utility scores may provide an unambiguous indication of the movement in the QoL, but economic decision making also requires a consideration of the impact of the intervention upon the quantity of life—life expectancy—and upon cost. In fortuitous circumstances the change in these three project attributes might lead to an unambiguous ranking of two interventions, A and B. This would occur, for example, when project A was superior to project B with respect to all three attributes. However, if project A led to a greater QoL improvement and project B to a greater increase in life expectancy then the overall improvement

would be ambiguous. If the costs of these two projects were the same, then the choice between them would depend upon the relative importance of the incremental improvement in QoL and life expectancy. Non—utility instruments are not constructed in a way which permits the combination or comparison of QoL and life expectancy and are, therefore, insufficient for the task of economic evaluation. If the cost of A and B differ then economic evaluation requires a comparison of the cost difference and the difference in health outcomes. This again implies the need for outcomes to be expressed in unambiguous and readily comprehended units.

## 2.2 MAU Instruments

Efforts to overcome this difficulty led to the development of the quality adjusted life year (QALY). While there have been numerous criticisms of the concept and its measurement (Burrows & Brown 1993; Richardson 1991) its rationale is simple. Life years are weighted by an index of utility—a preference weight—which acts as an exchange rate between the quantity and quality of life. This implies that a “strong” interval property must be met by the utility index. Not only must a given reduction in its value mean the same across the entire range of the index but it must mean the same as an equivalent reduction in the quantity of life. Thus, for example, a 0.2 reduction in the utility index from 0.8 to 0.6 must be equivalent to a reduction from 0.5 to 0.3, and this must be equivalent to a corresponding reduction in life expectancy.

Achieving this property is highly problematic, but in economic evaluation a comparison of changes in QoL and life expectancy cannot be avoided. The only way to achieve this property is through using one of the preference based scaling techniques; viz, time trade-off, standard gamble or the person trade-off. Each asks respondents to compare QoL and life expectancy and to trade off one against the other. The validity of the scale values obtained from the use of these techniques cannot be determined by direct comparison with actual behaviour: people cannot be easily observed trading QoL against life expectancy. Consequently scale values rely primarily upon construct validity: their acceptability depends primarily upon the belief that people may make meaningful judgments about the trade—off.

Few validity tests of MAU generic instruments have been reported in the literature. The result of one test of two instruments is reported in Table 1. Published values for the original McMaster and the Quality of Well-Being (QWB) instruments (column 2) were used to calculate the number of people whose full cure (utility index returns to 1.0) would be equivalent to saving a life (gaining 1.0). This is reported in column 3. Thus, for example, according to the QWB, curing one person from a ‘cough’ would increase utility by  $1 - 0.74 = 0.26$ . Four such cures would increase utility by  $4 \times 0.26 = 1.03$  and therefore be equivalent to saving a life. The implausibility of this and the other results in Table 1 casts serious doubt upon the existence of this strong interval property for these two instruments.

**TABLE 1**  
Number Cured Equivalent to Saving One Life—Implied by Two MAU Instruments

State	Published Value of State	Number Cured Equivalent to Saving a Life (Approximately)
<b>• McMaster Health Index Questionnaire</b>		
Some limitations in physical ability to lift, walk, run, jump or bend	0.870	8
Needing a hearing aid	0.870	8
Having pain or discomfort for a few days in a row every month	0.870	8
Needing mechanical aids to get around, but not needing help from others	0.730	4
<b>• Quality of Well—Being Scale (QWB)</b>		
Stuffy, running nose	0.830	6
Pimples	0.800	5
Lisp	0.763	4
Headache	0.756	4
Spells of feeling upset	0.743	4
Trouble with sleeping	0.743	4
Cough	0.743	4

Source: Nord *et al* 1993.

There are few generic MAU instruments which purport to produce utility scores for the construction of QALYs. While it is beyond the scope of the present paper to offer a full critique of these instruments it is the contention here—and the motivation of the present project—that these instruments have shortcomings. Three have particularly simple, and therefore rather insensitive, descriptive systems (the EuroQoL, the original Rosser/Kind and the original McMaster instruments [Drummond *et al* 1987, Williams 1995]). The validity of at least two of the instruments is open to question (see above; Nord *et al* 1993). While purporting to produce utility scores for the calculation of QALY values, several of the instruments were constructed without the use of preference scaling techniques (the 15D, Rosser/Kind, and QWB). In addition, the validity of the modelling procedures used to obtain utility scores for each of the multi-attribute health states is largely untested (Sintonen 1995). These issues indicate the present situation regarding the measurement of QALYs for economic evaluation is unsatisfactory.

### 3 Constructing the Australian MAU Instrument

In broad terms the steps involved in the Australian MAU project are the same as those with any MAU instrument: a conceptual model of health-related quality of life (HRQoL) is constructed, and the hypothesized dimensions and elements are determined. Once this preliminary structure is refined, an item bank is constructed, as far as possible faithfully reflecting element and dimension content. Items may be selected from the literature or written *de novo*. In either case, they may be refined through several re-writes to ensure they possess standard psychometric properties. Once finalized, they are piloted and data used to eliminate unsatisfactory or redundant items. Finally the constructed instrument is scaled and undergoes validation.

Given that HRQoL measurement in the multi-attribute paradigm elicits responses which are functions of preferences pertaining to theoretical constructs ('quality of life') it follows that scales

and instruments should meet recognized psychometric criteria, particularly with respect to construct validity and reliability. Construct validity "... is the extent to which the test may be said to measure a theoretical construct..." (Anastasi 1976;151), a point elaborated upon by Brown & Burrows (1992;44) who argued that: "The burden of proof that health status measures measure health status rests with construct validity... [which] is concerned with the extent to which the relationship between a particular measure and other measures are consistent with theoretically-driven hypotheses about the concepts (or constructs) being measured". Reliability, on the other hand is concerned with the stability of the measurement.

Hawthorne (1991) advanced six requirements to be met in instrument construction for there to be substantial evidence of construct validity and reliability under these conditions:

- The measurement should be based on a sound theoretical construct.
- Each dimension should be measured separately by a scale possessing good content validity, which is internally consistent and unidimensional.
- Scale items should be relevant to the dimension they are measuring, and should be free from obvious defects.
- Each scale should generate scores which are independent of scores obtained on scales measuring other dimensions.
- The sensitivity of both individual items and scales should be such that they differentiate between subjects.
- The instrument as a whole (which may comprise one or more scales) should be clearly relevant to the desired measurement.

The remainder of this section describes progress to date, based on meeting these requirements.

### **3.1 Conceptualizing 'Health-Related Quality of Life'**

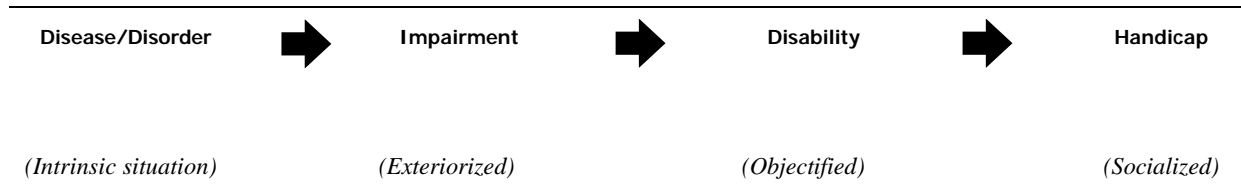
The definition of health accepted for the Australian MAU project is that advanced by the World Health Organization (1958), namely that health is a state of optimum physical, mental and social well-being and not merely the absence of disease or infirmity. In terms of defining an individual's health state, then, it follows the so-called 'gold standard' for any individual is when they are in this optimum state, and that any deviation from this constitutes a less-than optimum health state.

It follows that HRQoL measurement implies a deviation from this optimum, and that it should seek to establish both the type (direction) and magnitude individuals or populations deviate from this. It should also be concerned with the duration of health states.

The WHO definition also implies that health embraces all aspects of life, exemplified by Guyatt *et al's* (1993;622) argument that "the health domain ranges from negatively valued aspects of life, including death, to the more valued aspects such as role function or happiness... when a patient is ill or diseased, almost all aspects of life can become health related". This reflects the progressive shift which has occurred in the past twenty years regarding the conceptualization of

patient focused QoL: a move away from a simple concept of disease or its absence towards the notion of social handicap attributable to an initial disease or disorder, as shown in Figure 1.

**FIGURE 1**  
WHO Classification of Disease, Impairment, Disability and Handicap



The dualism present in this—the impairments/functional states and perception/social opportunities—can be seen in some previous instruments measuring QoL. Broadly these may be divided into two groups. Those concerned with QoL as a consequence of illness or impairment (ie where the measurement focused on the clinical symptoms of illness and impairment; such as movement, ambulation, vegetative functions, vision, hearing, pain etc.) and those concerned with where a person was functionally disabled or handicapped in some way (ie where the normal social role was affected; such as their ability to work, fulfill social obligations, perform family roles, look after themselves etc.).

The WHO (1980;27) definition of impairment defines this as "any loss or abnormality or psychological, physiological, or anatomical structure or function". Impairment may not necessarily affect a person's role functions (eg a person may be short-sighted, but this may be of no practical consequence if it can be easily corrected). It is only when it is severe enough to affect a person's functions that it becomes a disability, defined as "any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner or within the range considered normal for a human being" (WHO 1980;28). That is, the purely clinical or biological concept is replaced by an emphasis on objectified performance. In turn this only constitutes a handicap if "a disadvantage for a given individual, resulting from an impairment or disability, ... limits or prevents the fulfillment of a role that is normal... for that individual" (WHO 1980;29).

Within this scheme utility, or preferences, will be primarily dependent upon perceived handicap on the basis that 'disease/impairment' are physical manifestations of ill-health, whereas 'disability/handicap' reflect functional attributes of ill-health. Since it is functional attributes which impact upon individuals' everyday lives, in this study the concept of ideal HRQoL was defined as *living without handicap attributable to health status*. Consequent upon this definition, the measurement of HRQoL focused upon departures from this idealized health state, specifically where ill-health restricted a person's activities through 'handicap' as defined.

### 3.2 The Preliminary Synthesis

Given this conceptualization, it was necessary to define the various dimensions and their sub components—here referred to as 'elements'—which represented the universe of HRQoL.

The body of research relating to measuring HRQoL was searched to elicit the key dimensions and elements generally agreed upon, together with examples of items measuring these. Given the scope of the literature it was desirable to restrict this review to those instruments relevant to the study. Accordingly, the following inclusion criteria were developed: instruments needed to be based on a generic broad definition of health encompassing physical, emotional and social

dimensions; there needed to be documentation of the conceptual basis for the instrument; instruments had to meet basic testing standards for reliability and validity; they needed to be used in a variety of research settings; to have been reviewed in the published literature; to have used accepted measuring techniques; and to have been published since the early 1970s.

Fourteen instruments were located meeting these inclusion criteria. They were the Sickness impact profile (Bergner *et al* 1981), The Quality of well-being scale (Kaplan & Bush 1982; Kaplan *et al* 1976); The EUROQOL quality of life project (EUROQOL Group 1990); The general health questionnaire (Goldberg 1972); The physical and mental impairment of function evaluation (Gurel *et al* 1972); The Nottingham health profile (Hunt *et al* 1985); Functional status questionnaire (Jette 1980), the McMaster health index questionnaire (Chambers *et al* 1976), Arthritis impact measurement scale (Meenan 1982); The OARS multidimensional functional assessment questionnaire (Fillenbaum & Smyer 1981); The index of health-related quality of life (Rosser 1988); Questionnaire on health-related quality of life (15D) (Sintonen 1981, Sintonen 1995); Multi-Attribute Health Status System (Torrance *et al* 1992); and The MOS 36-item short form health survey (SF-36) (Ware *et al* 1993). Reviews of most of these instruments and their applications can be found in McDowell & Newell (1987), Walker & Rosser (1993) or Bowling (1995).

Following review, all fourteen instruments and their items were placed on a data file, pooled and analyzed. For each instrument, the dimensions around which it was constructed were identified and recorded, either through stated instrument dimensions or through instrument item examination (where dimensions were not explicitly stated). Two difficulties became apparent during the review which materially affected the elicited dimensions. The first was the representation of dimensions and the second the construction of items.

## **The Dimensions**

With respect to the representation of dimensions, many of the instruments reviewed focused on specific aspects of health and ill-health. This was further compounded by the fact that while in some instruments dimensions were covered by an item scale, in others they comprised item responses only. In other instruments it was clear that several of the postulated dimensions were included within an item scale.

Similar difficulties were encountered with respect to items as applied to dimensions. Instruments often contained several types of items producing different types of data. Others had items which measured in the one item several elements. In some cases items referred to different dimensions than the one nominated: although an item/item scale was allocated to a dimension, it contained statements/responses appropriate to other dimensions. Clearly, all these fail to meet the criteria for good instrument construction and valid measurement as outlined above.

During construction of the HRQoL model from this literature, the selection criteria for dimension and element inclusion were that they obviously contributed to the broad concept 'HRQoL' as defined above, that they emphasized handicap, and that a dimension was not obviously and fully subsumed by some other dimension. With respect to item selection, the criteria were that items were logically related to elements, that they were free of obvious blemishes, that they gave precise descriptions of outcomes and that they emphasized handicap.

This enabled the identification and classification of dimensions, the various elements within each dimension, and a set of items nested within each element. This classification was carried out by

three researchers, each initially working independently. When the results were pooled, a synthesized model of the components of the universe of HRQoL was constructed. The resultant model was subsequently modified by comparison with the dimension and item pool created cross-culturally by the World Health Organization for the WHOQOL (WHO 1993).

The application of the criteria resulted in the inclusion of some dimensions, elements and items concerned with impairment and disability, as these were not always fully subsumed by dimensions concerned with handicap. However, where possible items were selected which emphasized functioning. Thus, for example, pain—which is not itself a handicap—can be described in terms of its impact upon functioning. Consequent upon these decisions the initial synthesized model intentionally included a broader range of dimensions and elements than envisaged for the final instrument.

The constructed model of HRQoL by dimension and element is given in Figure 2.

### **The Items**

Based upon examination of the instrument item pool, draft items were written covering the content of each element within each dimension. To assist with understanding, a common format for all items was adopted where item responses formed either Guttman or Likert scales (see Figure 3 for a Guttman example). Wherever possible, and in the interests of simplicity, items were constrained to four item responses. The first response represented the best possible health state and the fourth response the hypothesized worst health state. For each element three or four items were initially constructed exploring different aspects of the defined content. Through a process of iteration these items were progressively refined by the research team.

All items were written in the first person. While it is expected that the final instrument may be used by researchers without recourse to a patient survey (eg by researchers with detailed and independent information on patient outcomes) it is also intended that the final instrument may be used in large-scale studies involving population samples. Under these conditions the only feasible method of administration is through patient self-administration.

**FIGURE 2**  
Proposed MAU Dimensions and Elements

Dimension	Elements
Illness	Pain Receiving medical treatment Reliance of medical aids
Independent living	Activities of daily living Bodily care Safety and security
Physical ability	Mobility Physical performance Rest and fatigue Sensory function Sexual relationship
Psychological state	Anxiety/depression Cognitive ability Emotional fulfillment Memory
Social interaction	Communication Family role Intimacy Social function Work satisfaction

**FIGURE 3**  
Item Example: Mobility

Thinking about my ability to get around.

- A. I am able to get around my home and community or travel by myself without any difficulty.
- B. Compared to others in my age group I am less able to get around the community by myself. Eg *I have difficulty using the available public transport or driving a car.*
- C. Compared to others in my age group I am severely limited in getting around the community by myself. Eg *I am not able to use the available public transport or drive a car.*
- D. I have great difficulty in getting around my own home by myself. Eg *unless helped by others I cannot more outside my home.*

Although no timeframe is included within each item, it is intended subjects in large scale studies will be instructed to respond by reflecting over the previous week. Thus items include comments like ‘generally’, or ‘only occasionally’. This was an attempt to get over the obvious difficulty that subjects’ QoL health status—and their assessment of that status—may vary quite considerably within a very short timeframe, may be affected by the use of drugs, or by recall bias where the timeframe is a long one. For example, people are in pain now or are generally pain free. When people are pain-free the impact of pain may be suppressed (in the present) but enhanced (in the past). Thus asking for assessment of their health state now or in the past may provide very misleading assessments. An item example—the mobility item from the final instrument version (see Tables 3 & 4)—is given in Figure 3.

The constructed model—comprising dimensions, elements and items—was subject to qualitative analysis through a series of focus groups and interviews with medical specialists and general practitioners affiliated with St Vincent’s Hospital (Melbourne) and the Department of Public Health and Community Medicine at The University of Melbourne. In most cases doctors were interviewed twice, and in some cases several times. The reason for approaching medical personnel was to draw upon their experience of patients. This ensured item responses properly reflected the best and worst health states, and adequately described the range of health states generally encountered in health care. Despite the obvious potential for bias, we found that after

explanation of the project's objectives the doctors were able to offer considerable insights into patients' perspectives. Contrary to expectation, the initial model was criticized for its emphasis on impairment rather than handicap (see below). The objective of this analysis was principally to test the structure and scope of the model (ie to determine if it omitted important HRQoL dimensions or included redundant dimensions, elements or items). Interviewees were also asked to modify, reject or add items in view of their knowledge of patient experiences and perceptions.

On the basis of this qualitative research, the synthesized model was revised and a smaller more focused model constructed. Poor items were discarded, and others re-written in the interests of clarity. The revised model dimensions and elements are given in Figure 2.

## **4 The Preliminary MAU Instrument**

The third stage of the study involves the construction of an initial descriptive system from the synthesized model. This work is currently in progress, and this paper provides interim results only.

The results reported here are data provided by 131 subjects recruited from St Vincent's Hospital. These subjects will be matched by data from 130 community subjects currently being interviewed. When the combined data are analyzed, the interim results presented here will change.

The purpose of the hospital patient interviews was to test the model and develop a descriptive system which reflected a wide range of health states. The objectives of interviews were, first, to test individual items and item responses (that they were relevant to patients, that they were clear and could be understood, that they were easy to answer, and that items discriminated between patients) and, second, to determine the content of the MAU instrument (through identifying which items were appropriate for inclusion). Thirdly, interview data were used to test whether the hypothesized dimensions could be constructed in such a way that each is orthogonal as required by multi-attribute modelling theory (Winterfeldt & Edwards 1986).

Hospital patients were recruited as they experience a wide range of the health states (which the instrument seeks to measure) and because populations with which the MAU will be ultimately used will include subjects receiving hospital care. The sample was a cluster sample selected through a listed process across all wards where new cases were reported within May/June 1995;  $131/149$  (88%) eligible cases agreed to participate. While it is possible this procedure resulted in selection bias, during recruitment every attempt was made to representatively cover a broad range of diagnostic related groups (DRGs), although at the time of the interim analysis these data were not available for all cases. Among the 53 cases for which DRG data were available there were 43 different DRGs. The demographic profile by gender and age is given in Table 2.

**TABLE 2**  
Subjects by Gender and Age

Gender	Age Group						Total
	≤29	30–39	40–49	50–59	60–69	70+	
Female	7	8	7	6	11	11	50
Male	14	6	6	12	7	11	56
Total	21	14	13	18	18	22	106

Note: Data missing from 7 cases.

Eligible patients were approached shortly after admittance to the ward or during a recovery period following their procedure, and consent obtained. They were offered the choice of self-completion or verbal administration of the questionnaires. Where subjects chose self-administration, the researcher stayed with the subject clarifying any difficulties. Few were reported and most subjects indicated they enjoyed filling in the questionnaire. Where subjects indicated they were getting tired, they were offered the option of self-completion in their own time and collection the next day by the researcher. No particular difficulties were identified where this procedure was followed (eg missing data).

The items were presented to subjects in random order: neither the MAU dimensions nor elements (or their descriptors) were given. To ensure control over instrumentation, two different random order versions were used.

Following data entry and cleaning, the data were examined and a preliminary instrument constructed following the procedures described below. It should be noted that this preliminary instrument is an *unweighted* instrument in the sense that a simple additive model has been used. In the final instrument version a *weighted* model will be developed.

## 4.1 Preliminary Scale Construction

Initially, items were examined, and poor items discarded. The remaining items were then subject to a variety of tests, resulting in the retention of fifteen items in five scales.

### Initial Item Selection: Sensitivity

This is a measure of how well an item can discriminate between subjects, ie it is a measure of the 'spread' of responses.

Two measures of item sensitivity were applied here: the actual range of responses across the possible range of responses expressed as a percentage, and the standard deviations. The formula for calculating the first measure is:

$$I_s = \left\{ \frac{AR}{PR} \right\} \cdot 100$$

Where AR and PR represent the actual and possible ranges respectively and  $I_s$  is the item sensitivity. All items with an obtained value less than 100% were discarded.

With respect to item standard deviations, all items where  $S_x \leq 0.50$  were also discarded on the grounds that they did not discriminate sufficiently between subjects.

### Scale Construction

The remaining items were then examined using an iterative process involving principal component, factor analysis and tests of reliability.

Factor analysis provides a statistical tool for examining the relationship structure between sets of variables through the identification of underlying patterns among responses to the variables. Where this underlying relationship is such that it can be demonstrated the variables belong together (ie they are, for example, measuring the same idea), then it can be asserted that the set of variables are 'unidimensional'. Factor analysis enables identification of this state by loading variables on common factors. Where the variables load on a single factor, a scale is said to be unidimensional. To clarify the relationship between variables, the model may be rotated. Where a VARIMAX model rotation is used, each unique factor is said to be orthogonal to other factors, ie the variance explained by each factor is maximised and the between-factor variance minimised.

The steps in constructing the initial MAU-model were:

- **Step 1.** All items for each element of each of the five proposed dimensions were pooled and examined using factor and reliability analyses. Poor items were discarded. The criteria were where an item failed to load ( $\geq 0.30$ ) on the principal component analysis, or obtained low item-rest-of-test values ( $r_{(it)} \leq 0.20$ ). Items deemed to have repetitive content (ie were too similar) were also discarded.
- **Step 2.** The surviving items were then pooled and subjected to a two-stage factor analysis (principal components and varimax) and further reliability analyses. Exclusion criteria were where an item failed to load ( $\geq 0.20$ ) on the principal component analysis, loaded ( $\geq$

0.30) on a second or third factor following a varimax rotation, or obtained low item-rest-of-test values ( $\leq 0.20$ ). Through introspection, items deemed to have repetitive content or which loaded on a factor where there was no logical scale construct (ie they logically didn't 'fit' with other items on the scale) were also discarded.

**Step 3.** Step 2 was repeated iteratively until the most parsimonious solution was derived, where—consistent with the theoretical model—there were the fewest significant factors ( $\lambda_k > 1.0$ ) explaining the greatest cumulative variance, where all items uniquely contributed to the concept HRQoL by loading ( $\geq 0.20$ ) on the principal component, and on varimax rotation loaded on the scale factor identified in Step 2 ( $\geq 0.30$ ).

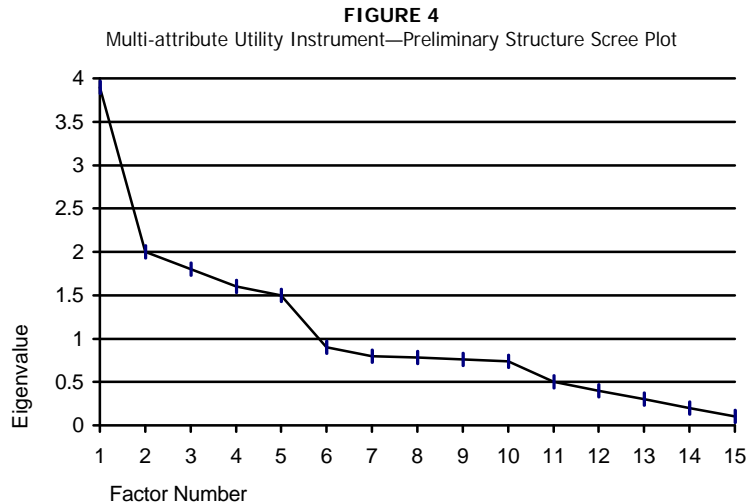
- **Step 4.** Each factor identified in Step 3 was treated as a scale, and analyzed accordingly for principal component loadings, and internal consistency. Redundant items were removed.
- **Step 5.** At this stage there were fifteen items remaining, in five scales. To check on the validity of the model all items were pooled and re-analyzed. This also provided a check on the unidimensionality of each scale. This involved principal components analysis (to check they all contributed to the theoretical concept of HRQoL), a varimax factor analysis (to verify that remaining items loaded on their appropriate scale and not on some other factor, and that the factors were orthogonal), and a reliability analysis involving all items (through estimating internal consistency).

The results of these procedures are given in Table 3.

The final analysis (Step 5) confirmed the presence of the five orthogonal factors shown in the varimax analysis in Table 3. This is graphically illustrated in the scree plot of factors in Figure 4. The cumulative proportion of variance explained by the model was 69%, while the reliability of the instrument, as assessed through internal consistency, was good, at Cronbach's standardized  $\alpha = 0.78$ .

**TABLE 3**  
Multi-attribute Utility Instrument—Preliminary Structure of the Instrument

Scale and Item Content	Principal Components Analysis			Factor Analysis <i>Varimax Rotation</i>					Reliability Analysis	
	PC	Eigen value	% variance	I	II	III	IV	V	$r_{(it)}$	$\alpha$
<b>Scale 1: Illness</b>										
1 Drug use	.48			.86					.39	
2 Medical aids use	.58			.86					.52	
3 Doctor treatment	.64	3.8	25.5%	.85					.48	
<b>Scale 2: Independent Living</b>										
4 Self-care	.62				.84				.43	
5 Mobility	.64				.84				.49	
6 Activities of daily living	.65	1.99	13.3%	.81					.50	
<b>Scale 3: Social Relationships</b>										
7 Social relationships	.38					.80			.27	
8 Family role	.51					.76			.39	
9 Purpose in life	.57	1.62	10.8%			.73			.45	
<b>Scale 4: Psychological State</b>										
10 Sleep	.27						.77		.20	
11 Cognition	.42						.70		.33	
12 Anxiety	.46	1.48	9.9%				.70		.39	
<b>Scale 5: Physical Senses</b>										
13 Bodily senses	.40							.79	.40	
14 Vision	.51							.75	.43	
15 Communication ability	.29	1.35	9.0%					.66	.20	
<b>Whole instrument summary</b>			68.5%							.78
<b>Additional items considered for inclusion on logical grounds</b>										
Sex (Scale 3)					.30	.37			.13	
Pain (Scale 4)					.30		.52			
Note:	Factor loadings <.30 not shown.									



Each of the five factor scales comprised three items. In all cases items loaded only  $>0.30$  on their factor. The separate analyses of each factor scale are given in Table 4. Principal component analysis of each factor scale revealed the presence of just one factor per scale—suggesting unidimensionality—consistent with the combined pooled varimax analysis (Table 3). The reliability of the five factor scales varied, from  $\alpha = 0.83$  for factor scales 1 & 2 to  $\alpha = 0.57$  for factor scale 4. These analyses, however, suggest each factor comprised a short scale with excellent or good psychometric properties.

While factor analysis identifies factors underlying data, it does not provide any interpretation of the meaning of each factor. It is, however, accepted that factor meanings can be ascertained through examination of the pivotal items' content, ie those items with the highest loadings. Where the pivotal items' content is consistent with the hypothesized conceptual structure, it provides evidence that the scale validly represents the theoretical model describing the concept under investigation.

In the current example where there were just three items per factor scale, the content of all three items was reviewed. Brief item descriptors are given in Tables 3 & 4, together with the interpretation placed on these factor scales.

When these interpretations are viewed against the synthesized model of HRQoL shown in Figure 2 it can be observed there is 'close match' of factor scales with dimensions.

For **illness**, it was hypothesized this would comprise pain, reliance on medical aids and receiving medical treatment. The analysis shows this construct was supported with three items, each covering one of the hypothesized aspects.

**TABLE 4**  
Multi-attribute Utility Instrument—Preliminary Structure of Scales

Scale and Item Content	PC Analysis			Reliability Analysis	
	PC	Eigen value	% variance	$r_{(it)}$	$\alpha$
<b>Scale 1: Illness</b>					
1 Drug use	.86			.67	
2 Medical aids use	.90			.74	
3 Doctor treatment	.83	2.2	74.2%	.64	.83
<b>Scale 2: Independent Living</b>					
4 Self-care	.86			.69	
5 Mobility	.85			.67	
6 Activities of daily living	.87	2.2	74.3%	.70	.83
<b>Scale 3: Social Relationships</b>					
7 Social relationships	.78			.49	
8 Family role	.80			.51	
9 Purpose in life	.76	1.8	60.6%	.46	.67
<b>Scale 4: Psychological State</b>					
10 Sleep	.78			.41	
11 Cognition	.70			.36	
12 Anxiety	.75	1.6	54.4%	.40	.58
<b>Scale 5: Physical Senses</b>					
13 Bodily senses	.79			.46	
14 Vision	.79			.48	
15 Communication ability	.67	1.7	56.8%	.35	.62
Whole instrument			68.5%		.78
Additional items considered for inclusion on logical grounds					
Sex (Scale 3)	.56			.32	
Pain (Scale 4)	.58			.31	

For **independent living**, it had been hypothesized this would refer to bodily care, activities of daily living (ADLs) and safety & security. The analysis showed this comprised items covering bodily care (self-care) and ADLs. The inclusion of mobility on this factor can be explained by the item content which referred to mobility around the local community: essential for independent living. (See Figure 3.)

**Physical ability** referred to physical performance, sensory functions, mobility, rest & fatigue and sexual relationships. The analysis showed this theoretical model was slightly inconsistent with the data: the correlations showed mobility was closely aligned with issues of independent living, and that sexual relationships were more closely aligned with social relationships. The items from the physical ability dimension which loaded together were items concerned with physical senses: the senses in general, vision and communication ability. Accordingly this dimension is labeled **physical senses** in Tables 3 & 4.

Regarding **psychological state**, it had been hypothesized this comprised emotional fulfillment, anxiety/depression, cognitive ability and memory. During analysis, items concerned with anxiety (sleep and anxiety proper) and cognition loaded together, while fulfillment was shown to be more closely correlated with social relationships.

Finally, **social relationships**. It was postulated this comprised intimacy, family role, communication, social function and work satisfaction. The final items loading on this factor covered relationships with non-family members and within the family, together with an item on personal fulfillment. This gives rise to speculation that social relationships provide a framework within which personal fulfillment is achieved.

These interpretations generally supported the synthesized theoretical model: five factors were identified and shown to be orthogonal, each comprised a unique factor. Each factor scale was clearly interpretable and obviously related to the five hypothesized dimensions, although with some minor differences in content.

The constructed instrument provided the most parsimonious solution to representing the synthesized model of HRQoL. There was, however, some concern regarding two aspects of QoL generally present in the literature, but omitted during the analysis: sexual function and pain. The analysis suggested that these were subsumed within other items or dimensions and that inclusion of these in the model added little to it.

For sexual function, analysis suggested this was most strongly correlated with social relationships, loading on this factor (0.37). It also loaded on the factor covering independent living (0.30), although not so strongly. An analysis of the social relationship scale incorporating the sexual function item showed this added little to the scale ( $r_{(1)} = 0.13$ ), and made little difference to its reliability estimate, from a standardized  $\alpha = 0.66$  to  $\alpha = 0.67$ .

**TABLE 5**  
Multi-attribute Utility Instrument—Estimated Scale and Instrument Properties

	Range	$I_s$	X	$S_x$	% scale range
Scale 1	0—9	100%	4.5	3.0	50%
Scale 2	0—8	89%	1.8	2.1	80%
Scale 3	0—8	89%	2.3	1.9	74%
Scale 4	0—8	89%	2.9	2.1	68%
Scale 5	0—7	78%	1.5	1.6	83%
Instrument	2—26	53%	12.6	6.6	72%

Note: % scale range = mean score expressed as a percentage of the potential scale score.

A similar situation was identified regarding pain. The three original items relating to pain initially loaded on different factors. As through iteration the instrument became more parsimonious and focused, the final surviving pain item loaded onto the psychological state scale. This was not expected, for it had been hypothesized pain would be associated with a person's physical condition; this loading may reflect the illness level of subjects, for these were all hospital inpatients. As shown in Table 3, pain ultimately loaded on two factors: psychological state (0.52) and to a lesser extent on the independent living scale (0.30). When it was included in the

psychological state scale, the analysis showed that although this strengthened the scale this was a marginal effect:  $r_{(i)} = 0.31$ , and standardized  $\alpha = 0.59$  as opposed to  $\alpha = 0.58$  without the item.

Given the inclusion of sexual function and pain made very little difference to the properties of either the whole instrument or the factor scales on which these loaded—and that these differences were trivial—the analyses suggested there were no strong statistical grounds for their inclusion. Despite these findings, we believe a logical case can be made for their inclusion, although we recognize this would add little (perhaps nothing) to the instrument or scales. It may be that when the final instrument is constructed, using a weighted model, a clearer picture of the role of sexual function and pain will emerge.

All items were scored on a 4-point Guttman or Likert response scale, where the normal health state was scored as '0' and the worst health state as '3'. Based on the sample data, through simple summation, this gave a range for each dimension (factor scale) of '0—9', where the value of '0' represented a normal HRQoL, and '9' the worst possible HRQoL. For the combined instrument, this gave a range of '0—45', where '0' represented the normal HRQoL health state and '45' the worst possible HRQoL state.

Scale and instrument sensitivities were computed using the formula cited above. The results are given in Table 5. This shows the sensitivity of the whole instrument was 53%. For scale 1 (illness) it was 100%, declining to 78% for scale 5 (physical senses). The mean and standard deviation for the whole scale suggest it would provide good measurement of HRQoL. There was some, but not a great deal, of evidence for a ceiling effect. Scales 1 & 4 (illness and psychological state) worked well, while there was some evidence of a ceiling effect (normal HRQoL) in the other three scales. The percentage scale ranges support this assessment. (See Table 5.)

\*

The instrument construction above was premised on a simple psychometric model of HRQoL measurement (in the same manner as many other HRQoL instruments): it assumes all items are of equal importance, that responses formed interval scales, and that the products could be added up to form a single score. If these assumptions are true (and this will be tested during further development work) it can be accepted that the instrument structure is an accurate representation of hypothesized HRQoL. The procedures followed and instrument properties revealed suggest this instrument meets the basic criteria for validity and reliability for this type of measurement outlined at the beginning of the paper.

## **5 Future Planned Work**

The work described above has been experimental, and has involved data collection from hospital subjects only. As a result of collecting subjects' opinions on the items used, revision of the more difficult items will be undertaken, and the refined items re-administered to a new hospital sample together with data from a community sample.

The next stage of the analysis—not yet undertaken—will be the inclusion of data from the both hospital and community subjects, followed by re-analysis. This will produce a 'final' construction. Once this is achieved, it will be followed by scaling of the descriptive system. This requires, first, the scaling of each of the items—obtaining the relative importance weights of each item response

and weighting these by the importance of the whole item as determined by a utility scaling technique. Second, item values must be combined to produce an overall QoL utility score.

As noted earlier the selection of the appropriate model for combining item scores is problematical and has received little attention in the literature. A difficulty that does not appear to have been satisfactorily overcome is that non-orthogonal items have been combined in such a way that the negative effect of one aspect of health has been separately included through two or more items. This results in 'double counting' this negative effect. Although the analysis presented above showed the dimensions (factor scales) in the preliminary MAU instrument are orthogonal—ie double-counting does not occur in this preliminary instrument—, to provide an extra safeguard item responses will be initially combined with a simple additive and multiplicative model; subsequently a 'hybrid' or 'nested' model will be used. In this, dimension non-orthogonal items (ie those which form a factor scale or dimension) will be combined using an additive model and the maximum disutility of that dimension independently determined. The values of the orthogonal dimensions will then be combined using either an additive or multiplicative model. This will give a global utility score which is not confounded by double counting.

The final stage of the project will be to carry out a conventional process of validation. As noted earlier there are few tests to determine whether or not an instrument has the strong interval property described previously. Whether or not the final utility score represents an acceptable exchange rate between life expectancy and QoL depends, in part, upon construct validity and the process of instrument construction. As well, it is proposed validation will include testing this instrument against other validated instruments to determine whether or not it is an acceptable HRQoL instrument in the context of particular diseases.

## **6 Conclusion**

This paper reports on ongoing research into the construction of a valid and reliable Australian MAU instrument.

We have presented the argument that for the economic evaluation of an intervention, the impact on both QoL and life expectancy must be considered. Consequently, measurement which considers QoL but neglects life expectancy is insufficient. This implies the need for a preference or utility based QoL instrument. Further, our review of present multi-attribute instruments purporting to measure utility shows many have serious defects. Hence the need for an instrument meeting the known requirements for validity and reliability. The research results reported here are preliminary. The analysis reported in Section 3 is incomplete as it omits non-hospital subjects, and is based on a simple unweighted additive model. It is therefore possible that the results reported here may be significantly altered during the analysis of the full sample.

However, these initial findings are very encouraging. They suggest that a wide range of health states may be described by a limited number of dimensions and that these broadly correspond with those found in the literature. The orthogonality of the dimensions supports the belief that a valid utility instrument may be based upon multi-attribute modelling of the health states.

## REFERENCES

- Anastasi A, Psychological testing. 4th Edition. New York, MacMillan Publishing, 1976.
- Bergner M, Bobbitt R, Carter W & Gilson B, The Sickness Impact Profile: development and final revision of a health status measure. *Medical Care*. 1981; 19; 787-805.
- Bowling A, Measuring disease: A review disease specific quality of life measurement scales. Open University Press, Suffolk, 1995.
- Brown K & Burrows C, What is validity? A prologue to an evaluation of selected health status instruments. Research Report 1. Melbourne, National Centre for Health Program Evaluation, 1992.
- Burrows C & Brown K, QALYs for resource allocation: probably not and certainly not now. *Australian Journal of Public Health*. 1993; 17 (3); 278-286.
- Chambers L, Sackett D, Goldsmith C *et al*, Development and application of an index of social function. *Health Services Research*. 1976; 11; 430-441.
- Drummond M, Stoddard G & Torrance G, Methods for the economic evaluation of health care programs. Oxford, Oxford Medical Publications, 1987.
- EUROQOL Group, EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy*. 1990; 16; 199-208.
- Fillenbaum G & Smyer M, The development, validity and reliability of the OARS multidimensional functional assessment questionnaire. *Journal of Gerontology*. 1981; 36; 428-434.
- Goldberg D, The detection of psychiatric illness by questionnaire. Oxford, Oxford University Press, 1972.

Gurel L, Linn M & Linn B, Physical and mental impairment-of-function evaluation in the aged: the PAMIE scale. *Journal of Gerontology*. 1972; 27; 83-90.

Guyatt G, Feeny D & Patrick D, Measuring health-related quality of life. *Annals of Internal Medicine*. 1993; 118; 622-629.

Hawthorne G, Pre-Driver Education: an evaluation of a traffic safety education program for senior students in Victorian schools. PhD thesis. Melbourne, Monash University, 1991.

Hunt S, McEwen J & McEwen S, Measuring health status: a new tool for clinicians and epidemiologists. *Journal of the Royal College of General Practitioners*. 1985; 35; 185-188.

Jette A, Functional status index: reliability of a chronic disease evaluation instrument. *Archives of Physical Medicine and Rehabilitation*. 1980; 61; 395-401.

Kaplan R & Bush J, Health-related quality of life measurement for evaluation research and policy analysis. *Health Psychology*. 1982; 1; 61-80.

Kaplan R, Bush J & Berry C, Health status: types of validity and the Index of well-being. *Health Services Research*. 1976; 11; 478-507.

McDowell, I. & Newell, C, Measuring health: a guide to rating scales and questionnaires. Oxford, Oxford University Press, 1987.

Meenan R, The AIMS approach to health status measurement: conceptual background and measurement properties. *Journal of Rheumatology*. 1982; 9; 785-788.

Nord E, Richardson J & Macarounas—Kirchmann K, Social evaluation of health care versus personal evaluation of health states, *International Journal of Technology Assessment and Health Care*. 1993; 9 (4); 463—478.

Richardson J, Economic assessment of health care: theory and practice. *The Australian Economic Review*, 1991; First Quarter; 2-21.

Rosser R, A health index and output measure. In Walker S & Rosser R (Eds), Quality of Life: Assessment and Application. MTP Press. 1988.

Sintonen H, The 15D measure of health related quality of life, feasibility, reliability and validity of its valuation system. Working paper 42. National Centre for Health Program Evaluation, 1995.

Sintonen H, An approach to measuring and valuing health states. *Social Science and Medicine*. 1981; 15; 55-65.

Torrance G, Zhang Y, Feeny D, Furlong W & Barr R, Multi-attribute preference functions for a comprehensive health status classification. CHEPA Working paper series N 92-18. Hamilton, McMaster University, 1992.

Walker S & Rosser R, Quality of life assessment: key issues in the 1990s. Dordrecht, Kluwer Academic Publishers, 1993.

Ware J, Snow K, Kosinski M & Gandek B, SF-36 health survey: manual and interpretation guide. Boston, The Health Institute, New England Medical Center, 1993.

Williams A, The measurement and validation of health: a chronicle. Discussion paper 136. York, Centre for Health Economics, The University of York, 1995.

Winterfeldt D Von & Edwards W, Decision analysis and behavioral research. Cambridge, Cambridge University Press, 1986.

World Health Organization (WHO), International classifications of impairments, disabilities and handicaps. Geneva, World Health Organization, 1980.

World Health Organization (WHO), Measuring quality of life. Geneva, Division of Mental Health, World Health Organization, 1993.

World Health Organization (WHO), The First Ten Years of the World Health Organization. Geneva, World Health Organization, 1958.

## Multi-attribute Utility Instrument

In general, when I think about being purposeful and fulfilled:

- A I feel very purposeful and fulfilled by my life.
- B Generally I feel fulfilled and that there is some purpose in my life.
- C My life does not make me feel particularly fulfilled. There is little purpose in my life.
- D I am not at all fulfilled. There is no purpose in my life.

Thinking about how I generally feel:

- A I do not feel anxious, worried or depressed.
- B I am slightly anxious, worried or depressed.
- C feel moderately anxious, worried or depressed.
- D I am extremely anxious and depressed.

When I think about solving problems—such as making plans for a party, learning how to operate a new computer or car, understanding bank statements, or following instructions in forms—I feel that:

- A Solving problems is an interesting challenge.
- B Familiar problems are enjoyable to solve.
- C Unfamiliar problems are a trial to me.
- D I feel hopeless about solving problems.

Thinking about my physical senses, such as seeing, hearing, touching etc:

- A I am not aware of any problem with any of my senses.
- B I have slight loss or impairment in one or more of my senses.
- C I have a major loss in one of my senses, although my other senses work quite well.
- D I suffer from two or more major sense impairments.

Reflecting on my vision, including when using my glasses or contact lenses if needed:

- A I see normally.
- B I have some difficulty focusing on things, or I do not see them sharply. *Eg small print, a newspaper, or seeing objects in the distance.*
- C I have a lot of difficulty seeing things. Most things are blurred. *Eg I can see just enough to get by with.*
- D I only see general shapes, or am blind. *Eg I need a guide to move around*

Thinking about my ability to get around:

- A I am able to get around my home and community or travel by myself without any difficulty.
- B Compared to others in my age group I am less able to get around the community by myself. *Eg I have difficulty using the available public transport or driving a car.*
- C Compared to others in my age group I am severely limited in getting around the community by myself. *Eg I am not able to use the available public transport or drive a car.*
- D I have great difficulty in getting around my own home by myself. *Eg unless helped by others I cannot move outside my home.*

If I think about how I sleep:

- A I am able to sleep without difficulty most of the time.
- B My sleep is interrupted some of the time, but I am usually able to go back to sleep without difficulty.
- C My sleep is interrupted most nights, but I am usually able to go back to sleep without difficulty.
- D I sleep in short bursts only. I am awake most of the night.

To what extent am I reliant on using drugs or a medical aid prescribed by a health worker? (NOT glasses or a hearing aid.)

- A I do not use any drugs and/or medical aids.
- B I occasionally use drugs and/or medical aids. *(Eg for occasional mild headaches or indigestion).*
- C I regularly use drugs and/or medical aids.
- D I rely on constant use of drugs and/or medical aid

Concerning my use of medicinal drugs.

- A I do not or rarely use any medicinal drugs at all.
- B I use one or two medicinal drugs regularly.
- C I need to use three or four medicinal drugs regularly.
- D I use five or more medicinal drugs regularly.

Do I require medical treatment from a doctor or allied health professional?

- A I do not need ongoing medical treatment.
- B Although I have some ongoing medical treatment, I am not dependent on this.
- C I am dependent on having ongoing regular medical treatment.
- D My life is dependent upon constant ongoing medical treatment.

Do I need any assistance looking after myself?

- A I need no assistance at all.
- B Occasionally I need some assistance with looking after myself.
- C I need assistance with the more difficult personal care tasks.
- D I need daily assistance with most or all personal care tasks.

With respect to doing household tasks, such as preparing food, using simple machines (eg TV, radio, CD player, telephone etc), tidying up the house or making my bed:

- A I have no difficulty in carrying out these tasks.
- B I find some of these tasks difficult, but I manage to do them on my own.
- C Many of these tasks are difficult, and I need help to do them.
- D I cannot do these tasks by myself at all.

Because of my health my relationships with my friends, partner or parents:

- A Are very close and warm.
- B Are sometimes close and warm.
- C Are seldom close and warm.
- D I have no close and warm relationships.

Thinking about my relationship with my family:

- A My role in the family is unaffected by my health.

- B There are some parts of my family role I cannot do because of my health.
- C There are many parts of my family role I cannot carry out because of my health.
- D I feel I have lost my role in the family because of my health.

When I communicate with others:

- A I have no trouble speaking to them or understanding what they are saying.
- B I have some difficulty being understood by people who do not know me. I have no trouble understanding what others are saying to me.
- C I am only understood by people who know me well. I have great trouble understanding what others are saying to me.
- D I cannot adequately communicate with others.

## **Supplementary Questions**

To what extent am I satisfied with my sexual relationship(s)?

- A I have no difficulty with my sexual relationships.
- B I am satisfied overall.
- C I have some difficulties.
- D My sexual relationship(s) is/are difficult for me.
- E I am completely dissatisfied with my sexual relationship(s).  
am not able to engage in or have an active sexual relationship.

Thinking about how often I experience serious pain. I experience it:

- A Very rarely.
- B Less than once a week.
- C Three to four times a week.
- D Most of the time.