ASSESSING HEALTH-RELATED QUALITY OF LIFE IN PEOPLE WITH APHASIA

by

Katerina Hilari

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

City University

Department of Language and Communication Science

May 2002

Volume 1

PAGE

NUMBERING

AS ORIGINAL

Στη Γιώτα και στον Γιάννη Χείλαρη
που μας δίδαξαν το νόημα της ζωής

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGMENTS

This project was funded by the Stroke Association and the Dunhill Medical Trust. Their financial support is much appreciated.

Many people have helped in the completion of this project and my sincere thanks go to all of them. I would like to mention and thank the following, in particular:

The people with aphasia who took part in this study, for their interest in the project, their willingness to help research and other people with aphasia and their generosity in sharing their experiences.

My supervisor Professor Sally Byng for being the most inspiring person and for offering the most insightful guidance. And for being there each step of the way.

The Speech and Language Therapy teams at the Whittington Centre in Streatham, the Streatham Hill Clinic, the Pulross Centre in Brixton, St Giles Hospital in Southwark, Queen Mary's Hospital in Sidcup and the whole team at Connect- the communication disability network in London. Rachel, Jo, Selena, Sarah, Dawn, Beryl, Norma, Hugh, Teresa, Hetty, Natasha, Tom and Sally McV, in particular, were always willing to help and did much more than I ever expected.

My advisory committee members Drs Donna Lamping and Sarah Smith and Professor Dick Wiggins for their insightful advice and their assistance throughout this project. Sarah, in particular, for being so particular in the best sense of the word.

Dr Sarah Schroter for her indirect support and for setting a great example to follow.

Dr Jane Marshall for encouraging me to undertake a PhD, for being my 'raw' model and for all the laughs.

The research support group members Kate, Simon and Lindy for all the sharing and bearing; and Deborah and Sharon for their friendship, their tireless listening to me, their reassurance and, again, the laughs. The unfailing enthusiasm about research and the perseverance of this group keep us all going.

My friends, for the support and the fun. Iphigeneia for being the first to share with me what it is like to do a PhD. Camelia, Rosalie, Anna and Lutfi Farah for their warmth and for becoming my adopted family in London!

Lastly, my family Bård, Sofia and Dimitris and mum. Mum for being the best mum in the world, for always being there and for her faith in me. Sophia for showing the most keen interest, asking the best questions and always knowing exactly how I feel and the best thing to say to me. And Bård mou, agapi mou, for everything, for my sanity and my happiness.

DECLARATION

The author grants powers of discretion to the University Librarian to allow the thesis to be copied in whole or in part without further reference to the author. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

ABSTRACT

Background: Health related quality of life (HRQL) measures are becoming increasingly used in the evaluation of health care interventions. They allow us to better understand the impact of disease on a patient's life as a whole and to incorporate the patient's perspective in clinical decision making and in the evaluation of health care. A number of studies have explored the HRQL of people with stroke. Still, due to a number of conceptual and methodological issues, there is no clear understanding of the HRQL of a stroke subgroup: people with aphasia.

Aims: The broad aim of this research was to explore the HRQL of people with chronic aphasia following stroke in a way that could be replicated in clinical practice. Thus, a single stroke-specific scale (the SS-QOL) was chosen for the assessment of HRQL. The specific research questions that were addressed were: A) Can an acceptable, reliable and valid version of the SS-QOL be developed for people with chronic aphasia? This involved: i) development of an aphasia-adapted version of the SS-QOL and ii) evaluation of its psychometric properties. B) What are the predictors of HRQL in people with chronic aphasia, as measured by the aphasia-adapted version of the SS-QOL?

Methods: The development of an aphasia-adapted version of the SS-QOL involved consultation with professionals with experience in measure development, language and aphasia, and pilot testing for the modification of the instrument, and a pre-test of the adapted version with 18 people with aphasia. This process resulted in the Stroke and Aphasia Quality of Life Scale (SAQOL). A cross-sectional interview-based survey study was undertaken to evaluate the psychometric properties (acceptability, reliability and validity) of the SAQOL and to determine the predictors of HRQL as measured by the SAQOL. Convenience sampling was used in the pilot and pre-test studies and cluster sampling in the survey study.

Measures: HRQL was measured with the SAQOL. In the construct validation of the SAQOL, the following measures were used: for emotional distress the GHQ-12, for cognition the RCPM, for activities the FAI, for social support the SSS and for language the FAST and the ASHA-FACS. Potential predictors of HRQL included demographic, stroke-related variables and variables implicated in previous research or of theoretical interest measured with the following instruments: the GHQ-12, the FAI, the SSS, the ASHA-FACS, the RCPM and the PSI (patients' satisfaction with stroke care).

Results: A) i) Development of an aphasia-adapted version of the SS-QOL resulted in the SAQOL, an interview administered self-report measure. People with moderate or mild receptive aphasia (as determined by a score of ≥ 7 in the receptive domains of the FAST) found the SAQOL acceptable and were able to self-report to it.

A) ii) Psychometric evaluation: 83 out of 95 participants self-reported on the SAQOL. The results supported the reliability and the validity of the overall SAQOL, but not of its subdomains' structure. A shorter 39-item version was derived through factor analysis (SAQOL-39). This instrument had a stable, conceptually clear 4-factor structure (physical, psychosocial, communication and energy) and high acceptability, internal consistency [scale (a= .93) and subdomains' (a=.74-.94)], test-retest reliability [scale (ICC=.98) and subdomains' (ICC=.89-.98)] and construct validity [corrected domain-total correlations (r=.38-.58), subdomains' convergent (r=.55-.67) and discriminant (r=.02-.27), and scale's discriminant (r=.19-.31) and correlated measures (r=.45-.58)].

B) Predictors of HRQL: High emotional distress, reduced involvement in home and outdoors activities, high communication disability and ≥2 comorbid conditions predicted poorer HRQL (adjusted R^2=.52). Stroke type (infarct vs haemorrhage) and demographic variables (age, gender, ethnicity, marital status, employment status and socioeconomic status) were not significant predictors of HRQL in these participants.

Conclusions: The SAQOL-39 is an acceptable, reliable and valid measure for the assessment of HRQL in people with chronic aphasia. Further testing is needed to establish the usability of this measure in evaluative research and routine clinical practice. Poor HRQL is predicted by distress, reduced involvement in activities, communication disability and comorbidity. Service providers need to take these factors into account when designing intervention programmes.

1 INTRODUCTION: THE ASSESSMENT OF HEALTH-RELATED QUALITY OF LIFE

1.1 Evaluating health care provision and patient-based outcomes

In recent decades there has been a paradigm shift in the way health and health care provision are conceptualised and evaluated. Traditionally, a medical conception of health was freedom from disease and abnormalities. In 1948 however, the World Health Organisation (WHO) indicated that health is 'a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity'. Baumann (1961) looked into lay definitions of health and reported three main orientations: a general feeling of well-being, the absence of illnesses and the ability to perform social roles. More recently and from a humanistic perspective, Heyrman and van Hoeck (1993) view optimal autonomy, personal strength and positive meaning of life as central components of health. They also point out that health is subject to cultural relativism. All these are broad conceptualisations of health and they highlight that different definitions stem from different philosophical approaches, value systems and different perspectives (e.g., lay people versus health professionals). Still, although there is no consensus on an exact definition of health, it is generally accepted in health-related research, that health is related to well-being and that it incorporates physical, mental and social components (Berzon et al., 1993).

This broader conceptualisation of health is reflected in the way health care interventions are evaluated. Evaluation has moved beyond the measurement of traditional clinical outcomes such as morbidity and mortality. It is now thought that the effectiveness of interventions

should be based on critical, objective and rigorous scientific evidence using a wide range of outcome measures (evidence-based practice) (NHS Executive, 1996).

Another reason for the shift towards incorporating a broad range of outcomes is the changing effect of health conditions. With advances in medical treatments and technologies, people are less likely to die from diseases and more likely to live with various degrees of long-term disabilities. Most of the care provided in such cases aims to relieve symptoms, reduce pain and discomfort, restore function and help patients in coping with the aftermath of the disease (Wenger et al., 1984). Thus, mortality is no longer an adequate measure of outcome. The increasing prevalence of chronic disability is reflected in health care evaluation with an increase in measures that can capture even small changes in the physical and mental well-being of the users of health care services.

Another change in recent years is that patients have become increasingly involved in treatment decisions (NHS Executive, 1999). The *Patient Partnership Strategy* aims to improve service delivery in the NHS by providing patients with information enabling then to make informed decisions about their health and health care (NHS Executive, 1999). There is, also, general consensus that patients and carers are 'experts' in their own conditions. Patients are the best informants about symptoms, feelings and the ways in which illness affects what is important to them (Mayou & Bryant, 1993). For these reasons, measures of outcome from the patient's perspective (patient-based outcomes) are increasingly used in the evaluation of health care interventions.

1.2 Health related quality of life (HRQL)

Commonly used patient-based outcome measures include measures of quality of life, HRQL, health status, well-being (subjective, psychological, emotional), functional status and patient

satisfaction. These terms are often defined loosely or not defined at all and they are frequently used interchangeably, causing considerable confusion in the area. This thesis is concerned with HRQL and here this term and the related term of quality of life will be discussed and defined.

Nowadays, it is generally accepted that HRQL measures focus on the impact of a perceived health state on a person's ability to live a fulfilling life (Bullinger et al., 1993). The question that arises is what does the concept of HRQL actually incorporate and how is it distinguished from the more common term of quality of life?

Quality of life was introduced as a heading by Medline in 1975 and was accepted as a concept by Index Medicus in 1977 (Bowling, 1995a). Since then there has been an explosion of interest in the area. A review of the literature by Fayers & Jones (1983) found over 200 papers published between 1978 and 1980 with the phrase 'quality of life' in the title. Today, a search in PubMed for papers with the phrase 'quality of life' in the title between 1998 and 2000 retrieves 3000 references. Over 1000 new articles each year are indexed under 'quality of life' (Muldoon et al., 1998).

Bowling (1995a) points out that quality of life is "an amorphous concept that has a usage across many disciplines - geography, literature, philosophy, health economics, advertising, health promotion and the medical and social sciences (e.g. sociology and psychology). It is a vague concept; it is multidimensional and theoretically incorporates all aspects of an individual's life".

The WHO has a working party undertaking a ten-country study of quality of life. They have provided the following definition:

> *"Quality of life is defined as an individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns. It is a broad ranging concept affected in complex ways by the person's physical health, psychological state, level of independence, social relationships, and their relationships to salient features of their environment".*
> (WHOQOL Group, 1993)

This is a broad definition incorporating individuals' perceptions, philosophy, culture and their relationships to the environment. It also seems to adequately represent the broad spectrum of potential consequences of disease. However, in order to evaluate the validity of its domains, we need to consider what people themselves see as essential domains of quality of life and as factors affecting it.

In Britain, Bowling (1995b) reported on a large scale study, which aimed to obtain population defined domains of quality of life and the relative importance of these domains in people's lives. In this study a random sample of 2000 adult members of the population of Great Britain (from the OPCS Omnibus Survey) was used, with a response rate of 77%. In response to an open question about the 5 most important things in their lives, respondents were most likely to mention as the first most important thing *relationships with family or relatives*, followed by their own *health*, the health of another (close) person and *finances/ standard of living /housing*. Elderly people (>65) were more likely to mention *health* as the first most important thing in life.

The results were different for people with chronic disabilities. For those who reported a longstanding illness the first most important effects of the illness on their lives were ability to *get out and about/ stand/ walk/ go out shopping*, being able to *work/ find a job* and effects on *social life/ leisure activities*. This second set of domains could be summarized as the physical,

functional and social aspects (including performing roles) of life that can be affected by health states.

Farquhar (1995) looked at elderly people's definitions of quality of life. The participants in this study were two groups of old people (65-85 year olds) one living in Hackney and one living in Essex and one group of very old people (85+) living in Hackney. In response to the question "What things give your life quality?" the most frequently mentioned answer for all three groups was *family* (children). Other frequently mentioned answers included *activities* (e.g. reading or watching television, going to clubs or the park, etc.), *social contacts*, *health*, *material circumstances*. In response to the question "What things take the quality away from your life?" the most frequently mentioned answer for the 65-85 year olds living in Hackney was *material circumstances* whereas for the other two groups it was *reduced social contacts*. *Ill health* was also high in the hierarchy. It would be interesting to see whether and how the relative importance of the factors mentioned as affecting quality of life in this study might vary depending on the health status of the individuals. However, no information is provided on the health status of the subjects.

In the USA, Pearlman & Uhlman (1988) explored the quality of life perceptions of elderly people with chronic diseases (arthritis, ischaemic heart disease, chronic pulmonary disease, diabetes mellitus and cancer). The factors that were most frequently mentioned as recently affecting the patients quality of life were *health related problems*, *medical care* and *interpersonal relationships*. The same factors came up in a further study (Pearlman & Uhlman, 1991), where a larger, more representative sample of elderly chronically ill outpatients was asked about the factors that have affected their quality of life over the past 12 months. The authors also noted that although *financial circumstances* were frequently acknowledged as affecting quality of life, the patients seldom explicitly volunteered them. Priorities seem to have shifted in this

sample of people with chronic diseases, where health related problems and medical care are the most frequently mentioned factors affecting quality of life.

In terms of HRQL, Bech (1993) quotes Joyce who defines HRQL as *"what the patient says it is"*. The subjective nature of HRQL is here emphasised. He also discusses a multidimensional model in the assessment of HRQL, which includes physical, cognitive, affective, social, economic and ego functions.

In reviewing the relevant literature, a number of authors have concentrated on what are the *essential* domains that should be included in a HRQL measure. An international group of HRQL researchers from the International Society for Quality of Life Research (Berzon et al., 1993) reached a consensus on the following: *physical, mental/psychological,* and *social health,* as well as *global perceptions of function and well-being* .

Bringing all this together, quality of life and HRQL are often used interchangeably in health-related research. In this thesis, quality of life is seen as a related but broader term than HRQL, incorporating a person's culture and value systems (WHOQOL, 1993) and encompassing factors like a safe environment and material well-being. Research into the views of lay people has shown that these factors are essential components of quality of life for healthy people but not as essential for people with chronic diseases and disabilities (Bowling, 1995b; Pearlman &Uhlman, 1988;1991). Moreover, the health care system and its providers usually do not assume responsibility for these more global human concerns although they may be adversely affected by disease (Patrick & Erickson, 1993). Evaluation of health care is mostly concerned with HRQL, which is seen as reflecting *the impact of a health state on a person's ability to lead a fulfilling life* (Bullinger et al., 1993). *It incorporates the individual's subjective evaluation of his/her physical, mental/emotional, family and social functioning* (Berzon et al.,

1993; Hays et al. 1993; de Haan et al. 1993). In the existing literature in the area, often the term quality of life has been used for what is here defined as HRQL. Throughout this thesis the term HRQL is used to operationilise the above definition and to refer to work that assumes this or a very similar version of this definition.

1.3 Applications of HRQL measures

Patient-based HRQL measures have a wide range of applications. There are becoming increasingly used in various areas of health-related research including the prioritisation of health care treatments and containing costs (Maynard, 1993); the evaluation of health policy programmes (Williams, 1993); the audit, quality assurance and evaluation of health care interventions (e.g., Jenkinson et al., 1994) and clinical decision making (Patrick & Erickson, 1993). The use of HRQL measures in the areas of clinical decision making and evaluation of health care interventions are discussed further, as they are of particular interest to health professionals working with people with long-term disabilities.

In terms of clinical decision making, Patrick & Erickson (1993) point out that HRQL is relevant both to individual patients and groups of patients. With individual patients the health professional assesses their lives and the impact of a disease and of treatment on them. With groups of patients the health professional assesses the current and future status of aggregates of patients and evaluates the overall impact of interventions. HRQL outcomes are used in clinical decision making in four ways:

1. To assess client status; for example a stroke patient is assessed on his/her physical abilities, daily and leisure activities, emotional distress and social functioning.

2. To select treatments; for example the effects of different anti-platelet drugs are explored.

3. To monitor the effects of treatments that have been selected.

4. To develop a shared view of the disease and of treatment outcomes with clients. They can help develop a shared understanding and a shared language, among health professionals and clients, of the impact of the disease on the patient.

HRQL measures are also used in the evaluation of health care interventions. They are becoming increasingly used in clinical trials as they provide a more comprehensive assessment of treatment from the client's perspective. Some clinical trial organisations have introduced the assessment of HRQL as a standard part of new trials (Fayers & Machin, 2000).

HRQL measures are particularly useful in the evaluation of health care interventions for people with chronic diseases and disabilities. Rehabilitation of people with chronic disabilities has traditionally focused on compensatory programmes (Frey, 1984) but in recent years it has begun to concentrate more on facilitating adaptation to disability and social and community integration (Royal College of Physicians, 2000; Turner, 1990; Wood-Dauphinee & Williams, 1987). Patient-based HRQL measures are particularly suited for the evaluation of health care provision in people with chronic disabilities as they can quantify the magnitude and duration of problems and experiences of people and the extent to which such problems and experiences affect their everyday life. By exploring a broad range of areas they allow us to better understand and measure the impact of disease on the person's life as a whole (e.g., Patrick & Erickson, 1993). They also help identify unmet needs and key areas in which further rehabilitation or additional support may be needed (Mayou & Bryant, 1993).

Subsequent sections of this chapter raise conceptual and methodological issues related to the assessment of HRQL. Then, information is provided on stroke and aphasia and the challenges in assessing HRQL in people with aphasia after stroke are presented.

1.4 Issues in the assessment of HRQL

1.4.1 Conceptual complexity

In this thesis some definitions of quality of life and HRQL have already been discussed and the operational definition of HRQL that has been followed in this work has been presented. Still, there is an ongoing debate in the literature on the conceptualisation of quality of life and HRQL. Some of the relevant perspectives will be presented briefly in order to highlight the diversity and complexity of the area.

Some theorists have examined the philosophical basis of the concept of quality of life. Megone (1994) drew influences from Aristotle where the best state for a being is the fulfilment of its function. The function of a being is determined by its essence and for human beings, according to Aristotle, the essence is the capacity for rational action. Therefore good life for a man would be a fully rational active life. The upshot from all this is that good life/quality of life is derived from an account of human nature. Thus, its components are fixed rather than determined by different things for different individuals. Hodge (1994) discussed this basic assumption that there is a structure called human nature, and contrasted it to existentialism. In the existentialist frame, identity is not permanent/constant but constituted and reconstituted through the daily decisions and responses produced by individual human beings. So people can have a sequence of identities as a result of radical shifts in orientation and evaluation of priorities. This changeable nature of identity has implications for the stability of quality of life assessments.

Bech (1990) discussed a nomothetic versus an ideographic approach. The first purports the selection of the most unbiased scale tailored to the disorder under investigation. The latter emphasises the construction of a hermeneutical or meaningful scale for the individual patient (using for example a repertory grid technique, where the patient determines the domains to be assessed). The main advantage of the ideographic approach compared to the nomothetic approach is its superior content validity, as the HRQL measures used are patient-derived. The main disadvantage of it is the challenge of comparing patients with one another due to the variability of the measured constructs.

In 1995, a series of articles appeared in an issue of *Social Science and Medicine (SSM)* raising issues about the conceptualisation of quality of life. Rosenberg contrasted the naturalistic with the hermeneutic approach. He argued that quality of life research is dominated by the empiricist psychometric tradition, which follows a naturalistic perspective. It is based on the assumption that man can be comprehensively studied by the empirical methods of natural science, psychology or sociology and that his behaviour can be causally explained from biological, psychological or social processes. According to hermeneutics, however, any area striving towards a comprehensive view of the individual must integrate essential aspects of man such as self-reflection, interpretation of life events and philosophical analyses of morals, norms, human dignity and rights. In other words, it must integrate those aspects of human existence, which cannot be captured in the naturalistic frame of reference. Rosenberg proposed that in the study of quality of life the naturalistic and hermeneutic approach should complement one another.

In the same issue of *SSM* (1995), Rogerson looked at the way that quality of life has been conceived and measured in both environmental and health related research. He defined HRQL on the basis of measures of attributes of health status and their characteristics. He

saw it as incorporating the patients' views and as related to their perceived levels of satisfaction and well-being. Jenkinson (1995) discussed the applications of quality of life measures and highlighted their limitations and the requirements (psychometric) they should meet. Ebrahim (1995) saw HRQL as "difficult to define but may be thought of as those aspects of self-perceived well-being that are related to or affected by the presence of disease or treatment". He indicated that HRQL indicators were of limited use for many clinical and public health tasks due mostly to their inappropriately tested reliability and validity (e.g., population repeatability being measured when an indicator is planned for use in examining changes in individuals; predictive validity being neglected). He pointed out that the most important reason for using HRQL measures was the evaluation of the effects of treatments. He also saw them as useful in exploring the subjective feelings of patients as an adjunct to clinical interview.

From a more pragmatic perspective, Muldoon et al. (1998) proposed two operational definitions of HRQL: HRQL as an individual's behaviour or level of functioning or as an individual's perceived health status or well-being. He pointed out that measuring someone's ability to perform common tasks or activities is putatively objective, while asking people to rate the effects of health status on personal well-being is explicitly subjective.

It is hard to draw any overall conclusions from these varied insights and perspectives. Despite the theoretical discussion devoted to the concept and the measurement of quality of life and HRQL, no unified approach has been devised for their measurement and little agreement has been attained on what they mean. On the contrary there seems to be a number of different interpretations each urging the adoption of a different approach (Hunt, 1997). Hunt emphasises the need for pure research, which would attempt to define, refine and understand the concepts of quality of life and HRQL so that a consensus is reached in

the scientific and clinical community. This way all concerned would know exactly what is being evaluated and priorities in medical care would be set on a standard basis.

In the meantime, the debate on these issues is continuing. In this thesis it is suggested that, for conceptual clarity, an essential criterion should be met. The investigators should conceptually identify what they mean by quality of life or HRQL and state clearly the domains they measured as components of the concept (Gill & Feinstein, 1994). They should also indicate clearly the rationale behind their methods.

1.4.2 Measurement issues

1.4.2.1 Types of measures

There are many different types of HRQL measures and the choice of what measure to use depends primarily on the purpose of the research. The most commonly used ones are multi-item scales. Early measures were developed and rated by clinicians and they were mostly limited to functional abilities (e.g., Karnofsky et al., 1948). Following this, a number of activities of daily living (ADL) scales were developed and gradually in the 1970's and 1980's measures of health status started emerging. Such measures included the Sickness Impact Profile (SIP) (Bergner et al., 1981) and the Nottigham Health Profile (NHP) (Hunt et al., 1981). These measures incorporate physical and psychosocial domains and are now often regarded and interpreted as HRQL measures.

Since then numerous HRQL scales have been developed. In 1992, the Short-form 36 Health Survey (SF-36) was developed (Ware et al., 1993), which is today the most commonly used health status measure (Wood Dauphinee, 1999). Unlike many other scales, the SF-36 has been extensively tested for its psychometric properties and evaluated in many populations. It is a generic measure that covers the domains of physical functioning, role limitations due to

physical health problems, bodily pain, social functioning, mental health, role limitations due to emotional problems, vitality/energy/fatigue, and general health perceptions. It has been pointed out that the SF-36 reflects the challenges inherent in any general health measurement: an instrument should be broad in scope but not unwieldy; and a trade-off has to be made between covering many topics superficially and achieving detailed coverage of a few, i.e., comprehensiveness versus precision (McDowell & Newell, 1996).

HRQL scales are commonly distinguished in generic scales versus disease-specific scales. Generic measures are overall scales of health status or HRQL, which have not been designed with a specific population in mind and can be used with different population groups. Their main advantage is their generalisability. Generic measures allow comparisons between different disease groups and they provide a common denominator or common unit of outcome by which to judge the relative severity of health outcomes and the relative effectiveness of interventions (Patrick & Erickson, 1993). Moreover, interventions can affect outcomes that are not condition specific and generic measures may pick up quality of life changes that were not anticipated and thus not included in disease-specific measures (Ware in Ware & Guyatt, 2001).

Disease-specific measures are developed with specific populations in mind and they are not intended for general application. Still they should be general enough to apply to different subpopulations under the same disease (Patrick & Erickson, 1993). Their main advantage is their increased validity and sensitivity (Bech, 1993). They are more likely than generic measures to detect small but clinically significant changes in health status or severity of disease (e.g., Patrick & Deyo, 1989). Moreover, they can reduce respondent burden by including only relevant questions (e.g., Bergner & Rothman, 1987). Ideally health outcomes assessment should incorporate both generic and disease-specific measures since they

complement each other (e.g., Patrick & Deyo, 1989; Muthny et al., 1990; Fletcher et al., 1992). Still, some believe that the use of disease-specific measures avoids asking irrelevant questions and maximizes the chance of detecting clinically significant changes, which is essential in clinical and policy-oriented research (Guyatt et al., 1986).

Another type of HRQL measures developed by health economists are utility assessments that are designed specifically for use in economic evaluations. The cost effectiveness of interventions can be evaluated by calculating quality adjusted life years (QALYs) (e.g., Weinstein & Stason, 1976). In QALYs improvements in length of life and HRQL are amalgamated into one single index. Each life year is quality adjusted with a utility value, where 1=full health. Bowling (1995a) points out that QALYs are not really measures of HRQL but measures of units of benefit from a medical intervention, combining life expectancy with an index of e.g., disability and distress. Other utility measures include the Rosser Index of Disability (Kind et al., 1982), the Kaplan's Index of Well Being Scale (Kaplan et al., 1976), the standard gamble technique (Torrance et al., 1982), the time trade-off (Torrance et al., 1972) and the EuroQol (EuroQol Group, 1990). Utility measures are increasingly used in clinical trials to evaluate the cost-effectiveness of interventions.

Another approach in the measurement of HRQL has tried to take individuals' meaning into account. In one technique (O'Boyle et al., 1992) human judgement analysis was used. The respondents were not given a set questionnaire but were asked to nominate the five most important areas of their lives and rate their function in these areas. This technique is known as the Schedule for the Evaluation of Individual Quality of Life (SEIQoL). In a similar approach, respondents were asked to rate the most important areas of their lives affected by their condition and rate how badly affected each one was (Ruta et al., 1994). The resulting instrument was the Patient Generated Index (Ruta et al., 1994). These measures represent

advances in developing more patient-centred HRQL outcomes (e.g., Staniszewska, 1999; Bowling, 1995a), but more work is needed on their acceptability (Bowling, 1995a). These measures have not yet been widely used.

Qualitative approaches have also been used in the assessment of HRQL. In particular with people with communication problems, in-depth interviewing (Parr et al., 1997) and semi-structured interviewing (Le Dorze & Brassard, 1995) have been used with people with aphasia. Observational techniques (non-participant observation) have also been used with elderly people (Clark & Bowling, 1990; Bond & Bond, 1990). Other approaches that have been proposed include the content analysis of verbal behaviour (Gottschalk & Lolas, 1992) and conversation analysis in people with learning disabilities (e.g., Antaki & Rapley, 1996).

In short, a variety of different measures exist for the assessment of HRQL. The decision on what measure to use will depend primarily on the purpose of the research and also on practical considerations (e.g., respondent burden, respondent communication skills, time constraints, resources, etc.). If, for example, a researcher wants to compare the cost-effectiveness of two interventions then a utility measure would be the most likely choice. If a clinician wants to use a measure to routinely assess the HRQL of groups of patients in clinical practice, then a scale would be a more likely choice than an individualised assessment like the SEIQoL or a qualitative technique which are more time consuming both in terms of data collection and data analysis. The approach followed in this study and the measure chosen for the assessment of HRQL in people with aphasia are discussed in chapter 3.

1.4.2.2 Psychometric properties

When scales are used for the measurement of HRQL they should have been subjected to rigorous psychometric testing. In order to be useful for research and clinical practice HRQL

measures need to be reliable and valid (e.g., Hays et al., 1993). Studies that include HRQL measures need to briefly report on the evidence of the measures' reliability and validity. Hays et al. (1993) point out that reports should also describe the conditions under which the study was completed, including the instructions to the participants, methods of administration, their demographic characteristics and range of illnesses experienced in the sample. This information will help interpret the study's findings and will help researchers in selecting measures for future studies in the clinical area. The psychometric evaluation of the measure used in this study is described in chapters 5 (methods) and 6 (results).

1.4.2.3 Content of measures

Kline Leidy (2001) discussed the effect of symptoms on HRQL. Patients' symptoms intuitively have an effect on their HRQL. They do not however constitute their HRQL. HRQL is a person's subjective evaluation of the impact of a health state on his/her life. As such HRQL measures should not be a list of symptoms or observable behaviours. They should instead reflect the impact of symptoms on patients' lives (e.g., Ware in Ware & Guyatt, 2001). For example asking a patient whether or not s/he can walk cannot tell much about their HRQL. Asking them, however, whether they have trouble walking can reflect an aspect of their HRQL, as it derives their subjective evaluation of their walking.

Similarly, given the subjective nature of the concept of HRQL the content of measures should be largely derived through consultation with patients, in order to include their perspectives (e.g., Streiner & Norman, 1995). Only those having a trait or disorder can report on its most subjective elements. Procedures used to elicit the patients' viewpoints in a rigorous and systematic way are primarily qualitative and they include focus groups and interviews (e.g., Willms & Johnson, 1993). An extra advantage of developing measures in

close contact with patients is that they are more likely to be easy to understand and acceptable to users (Muthny et al., 1990).

1.4.2.4 Who should rate HRQL?

There are a number of HRQL measures that are rated by health professionals rather than the patients themselves. Examples include the Physical and Mental Impairment of Function Evaluation (Gurel et al., 1972), the Disability and Distress Scale (Rosser & Kind, 1978) and the Quality of life Index (Spitzer et al., 1981). Proponents of health professionals rating the HRQL of patients believe that health professionals can judge more objectively and that certain parameters of HRQL, like complex functional indices, should (or can only) be rated by medical experts (Muthny et al., 1990). Still the subjective nature of the HRQL concept suggests that the most accurate raters are the patients themselves, and there is now general recognition that they should be the primary source of HRQL data (Berzon et al., 1993). Only they can make value judgements about the impact of a health state on their lives (e.g., Brock, 1993). Moreover, in terms of implications for service provision, it has been shown that the utilization of health services is more closely associated with the perception of symptoms and people's feelings than with their actual medical condition (e.g., Goldberg & Huxley, 1980).

Bringing all this together, in assessing HRQL investigators should operationalise HRQL indicating clearly what they mean and stating the domains they measure. Their choice of measure should be explicated, depending on the purpose of the research and practical considerations. When psychometric measures are used their psychometric properties need to be reported for the findings to be interpretable. Given the subjective nature of the concept of HRQL, the measures used should be patient-derived and the patients themselves should be the primary source of HRQL data.

1.5 Stroke and aphasia

1.5.1 Epidemiology, long-term disability and costs

Stroke is the most common cause of long-term adult disability in the western world. In Britain, the prevalence of stroke is 1.2% in men aged 45-54, 6.2% in those aged 65-74 and it rises to 10.3% in those aged 75+. In women the prevalence is 0.7%, 5.0% and 8.8% in these 3 age groups respectively (Department of Health, 1998).

Aphasia is a language disability caused by organic damage to the brain, most commonly a stroke. It can affect all language modalities, i.e., speaking and expressing oneself, understanding what other people say, reading and writing. In Britain, Wade et al (1986) looked at 976 patients who suffered a stroke over a 28-month period, from a community survey of 215,000 people. At three weeks post onset, 90% of survivors were tested and 20% had aphasia. At 6 months, 12% of survivors had aphasia, but 44% of patients and 57% of carers thought speech was abnormal. Scarpa et al. (1987) looked at people who were right handed and had suffered a stroke in the left hemisphere and estimated that one month post stroke 55.1% were aphasic.

The prevalence of stroke is not uniform across the population. In the *Health Survey for England* (Department of Health, 1998) there was a social class gradient in both sexes for ischaemic heart disease (IHD) and stroke. The social classes most likely to be associated with IHD or stroke were IIINM (13.1%) and IV (13.4%) in men ($p<0.01$, statistically significant difference from Social Class I), and Social Class V in women (14%, $p<0.001$ statistically significant difference from Social Class I). Men in the lowest income quintile had almost twice as high a prevalence of IHD or stroke as those in the highest quintile (16.0% versus 9.4%). The results were similar for women (11.3% in the lowest and 7.3% in the highest quintile).

Stroke has been reported, also, to be more common in the black population in Britain. Stewart et al. (1999) studied a prospective stroke register (1995-1996) of a multi-ethnic population of 234,533 in South London of whom 21% are black. 612 strokes were registered and incidence rates adjusted for age and gender were significantly higher in black compared with white people ($p<.0001$), with an incidence rate ratio of 2.21 (1.77 to 2.76).

There is also a geographical variation in the incidence of stroke. The mean incidence per 100,000 of the population is 356 in England, 448 in Northern Ireland and 497 in Scotland (Gibbs et al., 1998).

A number of studies have looked at the case fatality and disabilities resulting from stroke. Bamford et al. (1990) looked at the case fatality rates and 1-year outcome of stroke in a prospective study of the Oxfordshire Community Stroke Project, which covers a population of 105,000. The overall 30 day case fatality was 19%. One year post-stroke 23% of people who had suffered an ischaemic stroke were dead and 35% were dependent on others functionally. The rates for those who had suffered an intracerebral haemorrhage were 62% dead and 32% dependent, and for those who had suffered a subarachnoid haemorrhage 48% dead and 24% dependent at 1 year after the stroke.

In Northern England, Geddes et al. (1996) reported that 23% of respondents who had self-reported a stroke in a postal survey had made a full recovery of their stroke. The most common residual impairments were cognitive impairment (33%), lower limb disabilities (27-33%) and speech and language difficulties (27%). O' Mahony et al. (1999) reported that the prevalence of stroke associated dependence in Northern England was 11.7/1000, in an age- and gender-stratified sample of the population aged 45+.

Wolfe et al. (1993a) looked at case fatality of first ever strokes in those under 75 in Southeast England (South London and Tunbridge Wells) and reported it to be 26% at three weeks after stroke. 78% of all cases were treated in hospital with a median stay of 21 days. Three months later 30% of all cases had died and 26% of the surviving cases were moderately or severely disabled (Wolfe et al., 1993b). One year after the stroke the case fatality was 36% and 11% of surviving cases were still moderately to severely disabled and 23% mildly disabled (Wolfe et al., 1995). The authors estimated that the average cost per case was £3,800 in London and £2,650 in Tunbridge Wells. 93% of these costs were for inpatient care. The authors concluded that the cost of stroke care to the health services was considerable and largely reflecting nursing costs in hospitals rather than effective rehabilitation packages.

It is clear from all this, that stroke and its resulting disability, including aphasia, have a considerable impact in modern society and in health service provision. The *National Clinical Guidelines for Stroke* (Royal College of Physicians, 2000) identify as a major goal of rehabilitation for stroke the maximisation of the patient's sense of well-being/quality of life. The assessment of HRQL is therefore most pertinent in stroke and it has drawn considerable research interest (see chapter 2). The next section looks at the challenges aphasia poses on the assessment of HRQL after stroke.

1.5.2 Assessing HRQL after stroke and aphasia

In previous sections of this chapter the importance of patients being the primary source of HRQL data was emphasised. A key methodological challenge in the area of stroke HRQL is that people with aphasia may have difficulty completing self-report assessments. They may have difficulty understanding some of the items or expressing their responses. As a result, in some of the stroke HRQL studies, people with aphasia were excluded (e.g., Duncan et al.,

1997; Jonkman et al., 1998; Clarke et al., 1999). In some it is unclear whether they were included or not.

In the studies that did include people with aphasia, aphasia often resulted in missed assessments (Ebrahim et al., 1986; Kwa et al., 1996; Wilkinson et al., 1997). Alternatively, proxy respondents were used (e.g., Astrom et al., 1992; de Haan et al., 1995). The use of proxies is always less preferable than self-reports and the nature of HRQL may mean that the validity of proxy reports is further compromised. There tends to be a significant difference in proxy and self-report assessments of functional status and quality of life after stroke (Knapp & Hewison, 1999; Sneeuw et al., 1997), and of quality of life in patients with chronic disease in general (Sprangers & Aaronson, 1992). Analysing proxy-reported HRQL findings alongside self-reported findings is therefore questionable.

In some studies, no information is provided on how people with aphasia coped with the whole procedure (Foster & Young, 1996; King, 1996; Lofgren et al., 1999; Bethoux et al., 1999). For example, if people with aphasia were given a questionnaire and were asked to fill it in, how did people with reading difficulties cope and how did people with writing difficulties indicate their responses? Clinical experience of people with aphasia suggests that they would require at least some modification of the testing materials. In interview formats, they would require special skills on behalf of the interviewer in order to give their experience of stroke. The validity of these assessments is therefore in doubt.

These observations indicate that the very nature of aphasia as a language and communication impairment poses a serious challenge in assessing HRQL in an optimal way, i.e., through participants' self-report. This necessitates special attention on the mode of administration, the selection and presentation of the materials to be used and the skills of the interviewer,

when involved. Chapter 3 discusses the approach followed in this study in order to make the assessment accessible to the participants.

1.6 Summary

This chapter highlighted the changes in how health and consequently health care provision and evaluation are viewed and measured in recent years. The focus is on evaluating the effectiveness of interventions using a wide range of outcome measures. Such measures include patient-based measures in order to include the patient's perspective in the evaluation process. HRQL measures are one type of patient-based measures, which can be particularly useful in clinical decision making and in evaluating interventions especially for people with chronic disabilities. By exploring a broad range of areas they allow us to better understand and measure the impact of disease on the person's life as a whole. Issues in the assessment of HRQL were raised and the case of stroke and aphasia was presented. Stroke is the most common cause of long-term disability in adults in the modern world. A common sequel of stroke is aphasia, which affects a person's ability to understand and use language and to communicate effectively. By its very nature, aphasia poses challenges in the assessment of HRQL of the people who have it.

The broad aim of this research was to assess HRQL in people with aphasia. Chapter 2 discusses the existing literature in the area of HRQL in stroke and aphasia. Chapter 3 presents the research questions and the methodology of the current study. It, also, explicates the approach that was followed in order to address the assessment issues raised in this chapter and the next.

2 REVIEW OF HRQL STUDIES IN STROKE AND APHASIA

A number of studies have explored the HRQL of people after stroke and the factors affecting it. In this chapter, first the studies looking at the HRQL of people after stroke are reviewed. Then the studies assessing the main factors that are associated or are predicting HRQL are reviewed. The review includes both HRQL studies and studies of conceptually similar variables. These variables included well-being, life satisfaction and handicap. Studies that explored HRQL and related variables in people with aphasia, in particular, follow and the chapter closes with an overview of the reported studies and a summary of the issues arising from this literature review.

2.1 HRQL and related variables after stroke

One of the earliest studies exploring HRQL after stroke was by Lawrence and Christie (1979). The authors started from the premise that "...the dysfunction of longer-term survivors is often greater than would be expected from their physical disability [...]". They did not specify what they mean by HRQL. A semi-structured interview technique was used to elicit information on the HRQL and coping styles of the participants. The participants were 45 out of 170 consecutively admitted patients who survived 3 years after the stroke and had no severe communication disability. A relative or friend was also interviewed. It is not specified what was meant by 'no severe communication disability', but the likelihood is that people with severe and perhaps moderate aphasia were excluded.

They found that more than half of the participants had inappropriate reactions to or coping styles for illness (as judged by the interviewer) and uncertain or pessimistic attitudes towards the future (again as judged by the interviewer). They grouped the respondents in disability groups according to their scores on a physical disability scale. In the moderate disability group, half had withdrawn from active leisure activities and two thirds experienced severely deteriorated home relationships. In the minimal disability group, leisure activities were less affected but about one third experienced severely deteriorated home relationships. In terms of work, pre-stroke about 75-80% of the respondents were fully occupied. After the stroke, only 20-25% were fully occupied and 65-70% were not occupied. This early study highlighted the effects of stroke on home relationships, leisure and work.

In the same year, Gresham et al. (1979) examined the long-term 'functional disability' of long-term stroke survivors. Their participants were from the original 'Framingham Study' of 5,209 persons who were examined between 1948-1952. This cohort was re-examined between 1972-1975 and of the 354 people who had suffered a stroke, 155 were still living and 148 (95%) took part in this study. 148 controls were also included, matched for age and gender. The Donaldson Activities of Daily Living (ADL) Evaluation form (Donaldson et al., 1973) and set interview questions were used to evaluate 9 areas of functional disability: Not living at home, dependent in ADL, dependent in mobility, limited in household tasks, decreased ability to use transport, decreased vocational function, decreased socialisation at home and outside the home and decreased interest in leisure activities.

They found that people who had suffered a stroke were significantly worse in all these areas compared to controls ($p \leq .001$ for all except socialisation at home where $p \leq .01$). To test whether indeed these disabilities were due to stroke rather than other comorbid conditions, the cases where the disabilities could be due to comorbid conditions were identified and

excluded from the analysis. After this, the results were similar, except for vocational function, decreased socialisation at home and not living at home that did not reach significance. This early study did not set out to assess HRQL but gave information on the impact of stroke on the participants' physical and social health, which are now seen as components of HRQL.

More recently, Angeleri et al. (1993) looked at HRQL and return to work after stroke. Their subjects were 180 consecutive stroke patients who were hospitalised for the first time and were discharged at least 1 year before the study, and 167 age matched controls. The stroke subjects had a mean age of 65.29 years (SD 11.22) and were interviewed between 12 and 196 months post stroke (mean 37.5). People with aphasia were not excluded from this study but the authors specify that the Beck Depression Inventory (Beck et al., 1961) was not applied to them. It is unclear how people with aphasia managed some of the other more complex scales that were used in this study e.g., the Social Dysfunction Rating Scale (Linn et al., 1969). In this study, the authors offer no operational definition of HRQL. They used a multiple correlation statistical analysis of the scores on ADL (Northwestern University Disability Scale), depression (Beck Depression Inventory), social dysfunction (Social Dysfunction Rating Scale) and family stress (Greene Scale on family stress)[1] as an expression of HRQL.

They found that high ADL scores on discharge from hospital were a good prognostic indicator for return home. Depression and reduced social activities were both greater in women ($p<.01$). 20.64% of stroke survivors returned to work. In terms of HRQL, stroke survivors had significantly worse HRQL than the controls, as indicated by more ADL problems, depression, family stress and less social activities. There was a correlation between depression, social activity and family stress. The lack of definition and operationalisation of

[1] No reference was provided in Angeleri et al. (1993) for this scale and I was unable to trace it.

HRQL make these findings hard to interpret. Moreover, the selective exclusion of people with aphasia casts doubts on the applicability of the results for people with aphasia.

Hochstenbach et al (1996) used the Sickness Impact Profile (SIP) to assess HRQL with stroke survivors. This paper is in Dutch and only the abstract was reviewed. In this study, 165 patients who had suffered a stroke in the last 5 years and their relatives completed the SIP. The SIP looks at the individual's own perception of illness. It measures the effects of illness on activities, feelings and attitudes. It is behavioural, in that all 136 items are observable behaviours. The results indicated that 52% of the patients had psychosocial problems often to always and 60% had physical problems often to always. The authors reported that their results suggested that psychosocial problems arise independently of the degree of physical problems and that they are chronic.

Tuomilehto et al. (1995) explored psychosocial and health status in stroke survivors after 14 years. The authors felt that most of the stroke outcome studies concentrate on a few months or years after the stroke whereas they wanted to explore the long-term effects of stroke in people's lives. Their subjects were the survivors of the Finnish part of the collaborative WHO stroke study that took place during 1972-1974. 19.4% of them were alive after 14 years and of those 83.4% (201 subjects) were included in this study. The authors did not set out to measure HRQL and thus did not define the concept. They assessed, however, perceived health, ADL, and psychosocial status, which are commonly seen as HRQL components. They assessed these domains by means of a structured questionnaire. This was sent to the participants in order to familiarise themselves with it prior to the data collection. Then a telephone interview was carried out, where a nurse asked the questionnaire questions and filled out the answers. For people with communication problems their main caregiver was interviewed and for those who were hospitalised at the time of the study (15%) their

charge nurse gave the answers. The questions covered the areas of: medical history; neurological deficits (including language problems); socioeconomic status; ADL; psychosomatic status; mental state; and perceived health.

Their results showed that more than 50% of the participants had another major disease, two thirds had various degrees of leg or arm paresis and 25% had language problems. Still, two thirds of them felt that their ADL abilities and their functional status was good. Their psychosocial health, on the other hand, seemed more impaired: 60-70% had one or more psychosomatic symptoms. More than 50% complained of mental confusion, anxiety, irritability and dizziness. About 50% felt they had some degree of depression and 13% felt they had severe depression. In terms of perceived health, 7-12% of those under 65 and 17-18% of those over 65 felt it was bad. The rest felt it was satisfactory or good.

According to the authors, the results of this study suggested that in the long-term after the stroke, the impact of stroke is still considerable. People showed some adaptation to their physical limitations and did not feel particularly disabled (two thirds reported good functional status, and about 85% reported satisfactory or good health). Still, psychosomatic and mood problems were pertinent, affecting more than 60% of the respondents 14 years after the stroke. However, the lack of comparison with matched controls and the lack of control for comorbidity question whether the reported psychosomatic and mood problems can be attributed to the stroke, rather than other medical problems or circumstances.

Mathias et al. (1997) looked among other things at the HRQL of people after stroke. Their subjects were 74 individuals who had experienced an ischaemic stroke in the last 3 months; were over 18 years old; and were competent to participate in a 15- to 30-minute interview. It is not clear whether the last criterion resulted in the exclusion of people with aphasia and

cognitive decline. The authors did not define HRQL. They used the Health Utilities Index (HUI) (Feeny et al., 1996) to measure it. The HUI is a generic multi-attribute system for the assessment of health status, which covers the areas of vision, hearing, speech, ambulation, dexterity, cognition, self-care, and pain. Their results suggested that stroke impact was substantial, as 79% of their participants reported that 3 to 6 areas of the HUI were affected by stroke (18% reported 2 affected and 3% -1 person- none affected).

Wilkinson et al. (1997) conducted a study on the longer term HRQL and outcome in stroke patients. Their main aim was to assess whether the Barthel Index (Mahoney et al., 1958) alone was an adequate measure of outcome. To do this they carried out a battery of assessments with their subjects, which provided interesting information on the HRQL of the stroke participants. This HRQL information is what is presented here. Their subjects were drawn from a 1989-1990 stroke register of people under 75 years, resident in Southeast London when they had their first ever stroke. Eighty-six percent (106) of the survivors 4 to 5 years later were interviewed for this study. Twelve of those (11%) were unable to complete all the HRQL assessments, some because they were "unable to make their responses understood". Those people who did not complete the scales were generally more disabled than those who did. The assessments included the Barthel Index, handicap and disability scales [the Rankin scale (Bamford et al., 1989), the London Handicap Scale (Harwood et al., 1994), the Frenchay Activities index (FAI) (Wade et al., 1985)], the Mini Mental State Examination (MMSE) (Folstein et al., 1975) for cognition, the Hospital Anxiety and Depression Scale (HADS) (Zigmond & Snaith, 1983) for depression, and for HRQL and perceived health the SF-36 and the NHP. Bivariate analyses with non-parametric statistics were used to explore correlations between variables.

Overall, they found that physical disability as measured by the Barthel correlated highly with the depression and the HRQL measures. Depression was evident in about 50% of the ADL dependent stroke survivors and in 1 out of 6 of the ADL independent stroke survivors. The median scores of the SF-36 and the NHP were compared with the values of age and gender matched controls. Stroke survivors had significantly lower scores in all of the NHP and 6 out of the 8 domains of the SF-36. These results suggest that HRQL in stroke survivors is significantly reduced compared to controls.

Bethoux et al. (1999) looked at the effect of time on HRQL by comparing a group of hemiplegic stroke survivors who had completed their rehabilitation within the past 6 months (group 1) with a group that had completed their rehabilitation more than 6 months ago (group 2). Their subjects were 45 outpatients from their clinic, 46% of whom had aphasia. People with cognitive decline (MMSE<24) were excluded. Again the authors in this study did not define HRQL. They used the Reintegration to Normal Living Index (RNLI) to measure HRQL. The RNLI was developed by Wood-Dauphine and Williams (1987) not as a HRQL measure but as a measure of whether people had resumed their normal activities and responsibilities following illness. Siegrist and Junge (1990) criticized the RNLI because it fails to distinguish between physical and social functioning, for instance a person may be physically able but socially disinclined to perform a certain activity. Functional status and ADL were measured with the Functional Independence Measure (FIM) (Granger et al., 1993). Differences between the two groups were tested with the chi-square test, student's *t*-test, and the Mann-Whitney *U* test as appropriate.

They found that although the two groups were not significantly different in terms of functional status still group 2, i.e. the ones who were longer after discharge, had significantly lower RNLI scores ($p<.05$) in the areas of indoor mobility, self care, relationships and

handling of life events. However, the overall RNLI score did not differentiate between the two groups. Still, the authors concluded that HRQL seems to deteriorate with time after stroke even when disability remains unchanged. A limitation of this study is that the authors did not take into account a number of variables, which may have confounded their results such as depression, social support, general health and comorbidity.

Hackett et al. (2000) looked at HRQL 6 years after a stroke and compared it to controls and the general population of New Zealand. The authors did not define HRQL and they used the SF-36 to measure it. Their subjects were the 639 (36%) stroke patients that survived 6 years after the stroke from a sample of 1761 non-hospitalised stroke patients (from the Auckland Stroke Study 1991-1992). The controls were drawn from the General Electoral Roll for Auckland and were matched for gender and 10-year age strata. Data were collected through a phone interview with "each case or control or a close relative or caregiver who were willing to respond on their behalf'. Proxy respondents were used for people with aphasia and overall for 27% of cases and 9% of controls.

Their results suggested the mean SF-36 scores were significantly lower for people with stroke than for controls and the general population. After standardisation for age and gender, cases had significantly lower scores than the general population in the physical functioning, general health, vitality and social functioning domains of the SF-36. The applicability of these results for people with aphasia is unclear due to the use of proxy respondents. As it has been suggested in chapter 1, there tends to be a significant difference in proxy and self-report assessments of functional status and quality of life after stroke (Knapp & Hewison, 1999; Sneeuw et al., 1997).

2.2 Predictors of HRQL and related variables after stroke

Ahlsio et al. (1984) looked at the effects of disablement and emotional factors on HRQL after stroke. Their subjects were those who survived and consented to take part from all stroke patients that were admitted to the Stroke Unit of a specific hospital in Stockholm, during 1979. A representative sample of the patients admitted to the hospital is treated in the Stroke Unit, according to the authors. People with severe aphasia were excluded from the interviews. The total sample was 96 patients with a mean age of 71 years (range 35-90) of whom 22% had suffered a TIA. The subjects were followed up to 2 years by which time, 27% had died, 56% were living at home and 17% were in geriatric hospitals. The researchers defined HRQL as the experienced degree of satisfaction on human needs. In a structured interview, they used two Visual Analogue Scales (VAS) going from "worst possible" to "best possible". On one of them the respondents marked their HRQL before the stroke and on the other their HRQL after the stroke. Interestingly, their measure of HRQL does not necessarily reflect their definition of the concept. They also collected information on physical well-being, psychological situation, living conditions, relationships and opportunities for meaningful activities, in an interview format by means of a structured questionnaire. The Katz index was used for ADL (Katz & Akpom, 1976). The authors only used univariate analyses of their data (chi-square and independent and related-measures *t*-tests) to test what factors predicted HRQL after stroke.

They found that although ADL skills improved with time, HRQL did not even in the ADL independent group. At 2 years post stroke, 77% of the respondents felt that their HRQL had deteriorated. Physical disablement and psychological reactions (anxiety and depression) were the factors identified as influencing perceived HRQL. Age, gender and social factors had no significant effect on HRQL. Apart from socioeconomic group and living conditions it is

unclear what else was measured as social factors. Social support for example does not seem to have been addressed in this study. The authors did not use multivariate statistics and therefore no conclusions can be drawn on what are the predictors of HRQL or what is the relative importance of the different factors they looked at.

Ebrahim et al. (1986) explored the social and psychological problems experienced by stroke survivors. Their subjects were 198 stroke patients who survived to 6 months and consented to take part from 463 patients admitted to Nottingham with acute stroke within a period of 8 months. One hundred and fifty nine of them were assessed at 1 month and all of them at 6 months post onset. People with aphasia were not excluded but "dysphasia and mental impairment were the main reasons for missed assessments". To assess psychosocial outcomes they used the first part of the NHP. The NHP is a perceived health measure but due to its breadth of coverage it has been often used as a HRQL measure. The first part covers the domains of physical abilities, emotional reactions, social isolation, pain, energy level and sleep. In this investigation the physical abilities section was not used at 6 months as the researchers used a self developed ranked ADL scale. Emotional distress at 6 months was assessed with the scaled version of the General Health Questionnaire (GHQ-28) (Goldberg & Hillier, 1979). A control group was selected randomly from an age-gender register of a large group practice. Non-parametric tests were used for the statistical analyses (Mann-Whitney U test, Kruskal-Wallis ANOVA).

As expected stroke patients had significantly worse NHP scores than controls. NHP scores were correlated with extent of disability; length of hospital stay; place of residence (home versus hospital), and emotional distress at 6 months. However, average NHP scores did not change over the study period despite improvement in physical ability. Age, gender and living alone did not correlate with NHP scores. The majority of the patients were elderly and it is

not clear what 'living alone' was supposed to indicate (loneliness or independence?) or what it was compared with (living with a spouse or other relative, living in sheltered accommodation, living in a nursing home?). Overall, this study indicated that HRQL, as measured by the NHP, is associated with physical disability (although the relationship is not linear) and emotional distress. However, there is no indication of the effect of aphasia or cognitive decline. In addition, no information is given on other factors that may have confounded the findings as e.g., comorbidity, socioeconomic, and social variables (e.g., social support).

Niemi et al. (1988) explored quality of life 4 years after the stroke in young stroke survivors (< 65 years). From a stroke register of 255 cases of first ever stroke, 46 of the patients who survived 4 years and who were under 65 and were able to reply to the questionnaire that they used for the assessment of quality of life, took part in this study. People with severe aphasia and cognitive decline were excluded. Quality of life was conceptualised as "... a person's subjective well-being and life satisfaction and [...] it includes mental and physical health, material well-being, interpersonal relationships within and outside the family, work and other activities in the community, personal development and fulfilment, and active recreation". They used a 45-item questionnaire to assess quality of life, which was developed by the authors on the basis of literature review and clinical experience. No information is given on its reliability and validity. The areas covered in the questionnaire were: working conditions; activities at home; family relationships; and leisure activities. Information on personality, behavioural competence, and relationships with friends and relatives were also collected for descriptive purposes. Independent variables included: cognition [measured with Wechsler Adult Intelligence Scale (WAIS) and the Wechsler Memory Scale (WMS)]; CVA type and lesion location; age; gender; residual neurological impairment (presence of hemiparesis or not); co-ordination disturbances; aphasia; and tendency or presence of depression (as

reported by participants). The results were analysed using Yates' corrected chi-square test, student's *t*-tests and multiple regression.

Overall, 4 years after their stroke 98% of the stroke survivors were living at home, 87% were independent in ADL, and 54% of those employed before the stroke had returned to work. Still, 83% reported deterioration in their quality of life following the stroke. Quality of life was as often affected in ADL independent patients as in ADL dependent patients, and as often affected in older patients as in younger patients. The factors that were associated with a more severe reduction in HRQL were: tendency to depression, presence of hemiparesis, ADL dependency, older age, not returning to work, ischaemic (as opposed to haemorrhagic) stroke, and hemispheric (as opposed to brainstem or unspecified) lesion, and lower intelligence and memory quotient. Aphasia did not significantly affect HRQL, but the authors suggest that their results probably underestimate the importance of aphasia as people with severe aphasia were excluded.

In multiple regression, tendency to depression, difficulties in ambulation, ADL dependence and reduced memory quotient explained 73% of the variance in HRQL. The authors do not indicate how they explored the information on social factors they collected. Social support, socioeconomic status and comorbidity have not been investigated as independent variables here. Moreover, their measures of cognition (WAIS and WMS) rely on language and as such they could well identify people with aphasia as people with cognitive problems. Thus, the 'lower intelligence and memory quotient' may mask aphasia in some participants. Lastly, quality of life was not assessed with a patient-based measure. It is not possible to tell whether the questionnaire used adequately covered the concept of quality of life (content and construct validity) in general, and for stroke survivors in particular (they were not consulted during the development of the questionnaire).

Osberg et al. (1988) looked at factors predicting long-term outcome after stroke. They used three broadly conceived outcomes: 1) a composite variable that included functional status, mortality and discharge disposition, 2) life satisfaction (LS) 12 months post discharge, and 3) medical charges. Here only their findings on LS are reported, which is a close concept to HRQL and has been seen by some authors as HRQL (e.g., Ahlsio et al., 1984; Viitanen et al., 1988; Astrom et al., 1992; Astrom, Asplund & Astrom, 1992). Their subjects were 86 stroke and 3 TIA patients. The authors present their exclusion criteria in another paper, thus, it is not known whether people with aphasia were included. LS was measured with a single question "In general, how satisfied are you with your life?" and answers ranged from 1 'not at all satisfied' to 5 'very satisfied'. Independent variables included: severity of illness on admission [measured with the Horn Index (Horn et al., 1983)]; function at admission [measured with the Barthel Index or the Kenny Scale (Iversen et al., 1973)]; ambulation/wheelchair use; age; and income. Other variables were in and out of the house social supports. The former included marital and dating status and number of people in household, the latter included satisfaction with friends, telephone contact and participation in social activities outside the home.

They found that life satisfaction at 12 months was associated with severity of illness, life satisfaction at 1 month and in and out of the house social supports. These 4 variables explained almost 40% of the variance, in multiple regression analysis. This study highlights the importance of social support on LS after stroke. Still, there is not enough information on the sample to judge the generalisability of the results. Factors like depression and comorbidity, which could be confounding have not been taken into account in this investigation.

Viitanen et al. (1988) also explored life satisfaction in long-term stroke survivors. Their subjects were 62 stroke patients who survived 4 to 6 years after the stroke and consented to take part, from an initial cohort of 327 stroke and TIA patients admitted in a stroke unit from January 1978 - September 1981. Of those, 96 (33%) were alive 4-6 years later. People with severe aphasia and severe cognitive decline were excluded. The stroke group was compared with a group of 60 healthy controls. The authors presumed that general LS reflects HRQL. Participants were asked to rate their LS before and after the stroke in the following domains: general, self-care/ADL, leisure, togetherness-friends, togetherness-family, marriage, and sexuality. Ratings could vary from 1 'very dissatisfied' to 6 'very satisfied'. Other measures included: motor impairment (Occupational Therapist (OT) assessment); ADL (OT assessment); depression (Montgomery-Asberg Depression Rating Scale, MADRS) (Montgomery & Asberg, 1979); and social integration (OT assessment on a rating scale ranging from 0 'full participation' to 8 'institutionalised, isolated'). For comparisons of groups of data chi-square statistics were used.

Overall, they found that HRQL as reflected by general LS was significantly worse in long term stroke survivors than in healthy controls. Sixty-one percent of the stroke participants felt that stroke had led to decreased general or domain specific LS. Decreases in several aspects of life satisfaction were related to motor impairment and reduced ADL abilities. Still 11 out of the 23 subjects who had poor general LS had no motor impairment. Depression was present in 16% of the stroke survivors and it was significantly associated only with reduced leisure satisfaction. Reduced social integration was identified in 34% of the stroke survivors and was significantly associated with reduced general, ADL and leisure satisfaction. The authors, however, questioned the validity of their measurement of social integration as it did not reflect the participants' own experience with their social relationships but rather that

of the OT. The participants felt there were no significant changes in togetherness with family and friends and satisfaction with marriage post stroke.

Overall, this study indicated that satisfaction with different domains of life is reduced in long-term stroke survivors and it tends to be associated firstly with motor impairment and reduced ADL abilities and to a lesser extent with reduced social integration. Depression was only associated with reduced leisure satisfaction. The authors used only chi-square tests to analyse their findings, whereas a multivariate analysis might have led to a better picture of what is the relative importance of the different factors that were associated with HRQL.

Astrom and colleagues (first study: Astrom et al., 1992; second study Astrom, Asplund and Astrom, 1992) also viewed HRQL as Life Satisfaction (LS) and explored the mental, functional, and social factors associated with it. In the first study their subjects were 80 patients at a 3-month follow up of 98 consecutively admitted stroke patients in the Stroke Unit of a set hospital. Sixteen percent of them had suffered a TIA and 13% had aphasia. In the second study they followed up the 50 patients who survived up to 3 years. Twelve percent had suffered a TIA. The mean age was 71.4 years (SD 10.8). People with severe aphasia and cognitive decline were excluded. Proxy respondents were used for people who were "unable to co-operate" but it is not clear whether this included the remaining people with aphasia. The subjects were interviewed 4-5 days after the stroke regarding their pre-stroke situation and at 3 months, 1 year, 2 years and 3 years after stroke. The Katz index was used for ADL and depression was assessed by a Psychiatrist according to DSM-III criteria. A questionnaire was used to assess living conditions and LS and information was also collected on social networks. Bivariate statistics were used for within stroke group comparisons on different variables and across different times and for between group comparisons (stroke versus the general population).

They found that compared to the normal elderly population, stroke patients had more comorbidity problems and reduced LS even before their stroke. After the stroke they also had increased psychiatric and ADL problems and reduced leisure activities and social support from neighbours and friends. In the first study they reported that almost 50% of the stroke survivors had reduced LS compared to 2% in the national sample. In the second study they reported that, with regard to the effect of time, ADL changed little after 3 months whereas psychiatric symptoms continued to change between 3 months and 1 year. Still, the prevalence of depression at 3 months and 3 years was the same (25%). LS and leisure activities improved from 3 months to 1 year and remained unchanged thereafter. Social support remained unchanged from 3 months to 3 years. Overall, the 20% of stroke survivors who reported poor LS were older, lived alone and had reduced social support and ADL function and increased anxiety and tiredness.

De Haan et al. (1995) looked at the effects of stroke type and lesion location on HRQL after stroke. Their subjects were the 441 6-month stroke survivors who consented to participate from an original cohort of 760 consecutively admitted stroke patients in a multi-centre study in the Netherlands. People with dementia were excluded. All people with aphasia were included and proxy respondents were used for people with severe aphasia. HRQL was seen as including at least four dimensions: physical, functional, psychological, and social health. The SIP was used to measure HRQL. Differences in SIP scores of stroke patients in relation to reference data, hemispheric lesion laterality and stroke types were analysed with related and independent group *t*-tests. Chi-square tests were used to study the relationship between specific HRQL patterns and patient and clinical features.

Overall, they found that patients with infratentorial strokes reported better functioning than patients with supratentorial strokes. Type of (sub) cortical stroke, in terms of infarct versus

haemorrhage, was unrelated to HRQL. Lesion laterality and type of stroke had no significant effect on level of emotional distress. Cluster analysis revealed that stroke patients' HRQL could be described in three clusters. Sixty percent had mildly impaired HRQL scores. These patients tended to be younger, male and with infratentorial and lacunar strokes. Seven percent had high levels of psychosocial dysfunction. One third (33%) of the subjects had severely impaired HRQL, both physically and psychosocially. Severely affected HRQL was associated with older age, comorbidity and stroke variables (initial severity and a supratentorial lesion). The relative contribution of other potentially confounding variables like depression and social support was not explored in this study.

People with left-sided lesions and speech and language problems did not differ significantly in their HRQL scores from people with right-sided lesions, who also reported communication problems. This finding, however, may be an artefact of the measurement used. More than half of the people with a left-sided lesion have aphasia and in this study proxy respondents were used for people with severe aphasia. Due to the lack of agreement between proxy and self-report HRQL, analysing proxy alongside self-report data is not recommended (see chapter 1).

A surprising finding of this study is the low level of emotional distress reported (7%), as quoted rates of depression following stroke range from 18% to 61% (House, 1987). The authors suggest that "...stroke per se may not result in emotional problems but that such problems result from a complex interaction between patients' personal traits, social circumstances, living arrangements and functional abilities." The question that arises is why these factors did not 'operate' in this sample of stroke survivors. Perhaps this finding (low levels of emotional distress) is also an artefact of its measurement. The relevant subscales of

the SIP may not be the best way to identify emotional distress, due to the behavioural nature of the measure.

King (1996) examined quality of life in long-term stroke survivors and sought to identify variables that predict it. Her subjects were 86 stroke survivors who consented to participate from a group of 121 persons who met the inclusion criteria from a pool of 698 records of consecutively discharged stroke patients. The inclusion criteria were: 1-3 years post discharge from rehabilitation for first stroke; no other neuromusculoskeletal condition; adequate cognitive and language function (i.e., people with cognitive decline –MMSE<24- and severe aphasia were excluded); and residing in a noninstitutional setting. HRQL was defined as satisfaction with aspects of life that are important to the individual. It was measured with the Ferrans and Powers Quality of Life Index – Stroke Version (Ferrans & Powers, 1985), which rates 38 items for satisfaction and importance. It includes 4 sub-scales: health and functioning, socio-economic, psychological-spiritual, and family. Independent variables included age, socio-economic status (SES), comorbidity, aphasia (20% had aphasia), functional status (measured with the FIM), motor impairment (judged by investigator), depression [measured with the Centre for Epidemiological Studies Depression Scale (CES-D) (Radloff & Locke, 1986)], and perceived social support [measured with Social Support for the Elderly (SSE) scale, (M. Powers & J. Miller, 1986 unpublished data]. Data were also collected on gender, marital status, race, education, duration of stroke, and location of lesion for descriptive purposes. Student's *t*-test, chi-square statistics and correlation coefficients were used to compare groups of subjects and examine relations between variables. Stepwise multiple regression was computed to predict quality of life.

A surprising finding was that the mean quality of life Index score (22.9 out of 30, where 30 is the best possible score) was similar to that of a sample of 339 subjects drawn randomly from

a telephone directory (mean score 23). However, quality of life, as measured with the Ferrans and Powers instrument, does not seem to correlate with subjects' life satisfaction. In the question "how satisfied are you with your life in general?" 23% of the subjects reported dissatisfaction or slight satisfaction.

In terms of factors affecting quality of life, 38% of the variance of quality of life scores was explained by depression, perceived social support and functional status. Aphasia and comorbidity were not significant predictors of quality of life. It should be noted however, that only people with "adequate language" were included in the study, which probably reflects only mild aphasia. The findings overall seem to reflect milder strokes, since people with more than one stroke, cognitive decline, and living in institutions were also excluded.

From a statistical point of view, stepwise regression relies only on statistics computed from the particular sample in order to decide which variables are included in the regression equation. It is, thus, recommended that the cases to independent variables ratio is 40 to 1, as the solution will not generalise beyond the sample unless the sample is large (Tabachnick & Fidell, 2001). Cross validation is also recommended in stepwise regression (deriving the solution with some of the cases and testing it on the remaining cases) to test the generalizability of the solution. These criteria were not met in this study, which casts doubts on the generalizability of the results.

In terms of the way quality of life was conceptualised and assessed, this study is the first in this review to take into account the participants' opinion on the importance of different quality of life domains. The fact, however, that the instrument used did not seem to differentiate stroke survivors from randomly selected subjects, in combination with its low

correlation with life satisfaction, raises some questions on its sensitivity and validity with stroke survivors.

Kwa et al. (1996) investigated the role of cognitive impairment in HRQL after stroke. Their subjects were 129 stroke survivors with ischaemic stroke out of 252 consecutively admitted stroke patients, who were interviewed an average of 2.3 years (SD 0.8 years) after the stroke. Their mean age was 63.2 (SD 14.6 years). The authors did not define HRQL. They chose a simple way to measure it, that is a vertical visual analogue scale (VAS) going from "worst possible quality of life" at the bottom to "best possible quality of life" at the top, due to the inclusion of people with cognitive decline and aphasia. Their independent variables were: infarct volume; location of lesion; comorbidity; arm and leg function [measured with the Motricity Index (Demeurisse et al., 1980)]; ADL (measured with the Barthel); global functional status (measured with the Rankin scale); aphasia [measured with the Boston Diagnostic Aphasia Examination (BDAE) (Goodglass & Kaplan, 1983). 38% of the subjects had aphasia]; and cognition [measured with the CAMCOG (Roth et al., 1986). 36% of the subjects had serious cognitive decline]. Independent samples *t*-tests were used in univariate analyses to assess the relationship of participants characteristics and clinical factors on HRQL. Significant variables ($p \leq .2$) were entered in a forward stepwise regression analysis.

Despite the use of the VAS, 25% of the subjects were non-assessable due to serious communication problems. In the regression analysis, residuals analysis showed no violations of the assumptions of linearity, equality of variance, independence of errors and normality. The strongest predictors of diminished HRQL were reduced global functional status, larger infarct volume and severity of aphasia (explained 22% of the variance). Cognitive decline did not reach significance in the multiple regression analysis. Still, it should be noted that a

quarter of the subjects could not have their HRQL assessed and this group included people with the most severely affected cognition.

Overall, the strengths of this study include a comprehensive assessment of aphasia and cognition and an attempt to measure HRQL in a way that would be accessible to people with cognitive and communication difficulties. Still, HRQL was assessed in a rudimentary way by the use of a VAS, which does not indicate what aspects and dimensions of the concept the respondents took into account when giving their ratings. The presumed simplicity of the VAS is also questioned, as 25% of respondents could not do it. Perhaps the abstract, non-specific nature of the measurement is more demanding for people with cognitive or communication difficulties than a set of specific questions. Lastly, this study has not taken into account potentially confounding variables, such as depression and social support.

Duncan et al. (1997) explored the predictors of health status and HRQL of individuals with mild stroke. 304 people with mild stroke were recruited from three sources in the USA: the American Medical Centre Consortium with files of five academic medical centres; the United HealthCare with files of five independent practice associations; and a community-based sample from the Cardiovascular Health Study. People with aphasia or cognitive decline were excluded. The 304 subjects with mild stroke were compared with 184 TIA patients and 654 subjects with an elevated risk for stroke (asymptomatic group). As an indication of health status they used the SF-36. This measure is most often used as a measure of HRQL as it covers a broad spectrum of domains (general health, mental health, role limitations due to emotional problems, role limitations due to physical problems, social function, vitality, pain, and physical function), but in this study it was used just as a health status measure. The authors did not define HRQL and they used two 'utilities' to measure it. These were a time trade-off (TTO) (whether participants would prefer to live 10 years in their current state of

health or 9 years in excellent health) and a rating on a VAS of their current HRQL with 0 representing death and 100 representing excellent health. The conceptual confusion here is apparent, as 'death' and 'excellent health' seem more like the end points of health status rather than HRQL. Other 'functional measures' in this study included the Barthel Index for physical function and ADL and the CES-D for depression. Sociodemographic variables, including social support, were also taken into account. Groups were compared with chi-square statistics for categorical variables and ANOVA for continuous variables. Regression analysis was used to determine whether patient group (stroke, TIA, asymptomatic), comorbid conditions and Barthel Index scores were predictive of responses to any of the eight different domains of the SF-36 and the CES-D.

Overall, HRQL as measured by the TTO and VAS scale was lower for the stroke group than the other two. Health status, as measured by the SF-36, was similar between stroke and TIA (except for the physical function sub-scale) and significantly lower than the health status of the asymptomatic group. This was the case, despite the fact that 66% of stroke subjects and 81% of TIA subjects had a Barthel Index of 100, i.e. no difficulties in ADL. The stroke and TIA subjects also suffered from higher comorbidity and higher incidence of depression.

In the regression analyses, the Barthel Index and a history of stroke were the strongest predictors of health status (the models accounted for 6-36% of the variance in the different domains of health status) and depression (8% of the variance explained). The authors concluded from these findings that in addition to stroke, reduced ADL/physical abilities as measured by the Barthel Index was a consistent predictor of health status and depression. It is interesting that they did not explore the effect of depression on health status or HRQL. In any case, the poor conceptualisation and measurement of HRQL limit the applicability of these results. The authors also do not address questions such as: what factors better

predicted reduced HRQL within each symptom group; or whether any of the social variables (e.g., socioeconomic status, social support) were associated with HRQL or health status.

Jonkman et al. (1998) looked at HRQL after first ischaemic middle cerebral artery (MCA) stroke in the period of 3-12 months post onset. Their subjects were 35 consecutively admitted patients ranging in age between 25 and 70 years. Patients with severe aphasia, a pre-stroke mental disorder, alcohol and drug abuse, comorbidity that could lead to reduced cerebral functioning and expected survival less than 4 years were excluded. Participants were assessed at 3, 6, and 12 months post onset. Twenty controls were used, matched for age, last occupation and educational level. The authors did not offer a HRQL definition. They used the SIP to measure it. They evaluated the effect of neurological deficit, mood and cognitive function on HRQL. These were measured respectively with the 'Stroke Databank Neurologic History and Neurologic Examination' form (Shinar et al., 1985); the HADS; and the revised WAIS, the WMS, together with a reading test and 4 tests for skilfulness in naming, writing , calculating and visual construction, and reaction times. Non-parametric statistical tests were used to compare the stroke group to the control group and to assess changes over time and stepwise regression analysis was used to explore the predictors of HRQL (SIP total score).

As expected the stroke patients HRQL was significantly reduced compared with the controls. The SIP total and the SIP physical scores improved overtime for the stroke survivors but the SIP psychosocial scores did not. There was a correlation between the SIP scores and total weakness, depression and reduced IQ. In multiple regression, depression and total weakness explained 55% of the variance of total SIP scores, whereas cognitive decline did not reach significance. The stroke patients also had increased depression incidence and reduced cognitive function compared to the controls. This last finding is questionable, however, as all

the 'intelligence' tasks required language. A language impairment or even just a speech impairment would affect performance.

Overall, this study emphasised the effects of depression and physical disabilities on HRQL. It also highlighted the long-term nature of the psychosocial sequelae of stroke. However, its methodology for evaluating cognition after stroke is flawed. Moreover, the use of stepwise regression with such a small sample and the extended exclusion criteria limit the generalisability of the results.

Neau et al. (1998) explored the HRQL of young stroke survivors and in addition they looked at the factors that were associated with return to work. Of 75 consecutively admitted patients (15-45 years old) with ischaemic stroke, 67 were included in the study and 65 were followed up at a mean of 31.7 (SD 13) months after the stroke. The authors do not provide a conceptual definition of HRQL. They use the psychosocial domains of the SIP to measure HRQL. They collected information on demographic variables, socioeconomic status, social and family conditions, risk factors/comorbidity; aetiology of stroke; infarct territory; depression (measured with the MADRS and DSM-IIIR); aphasia; return to work; neurological deficit on admission (measured with the National Institutes of Health (NIH) Stroke Scale; post-stroke seizures; functional disability (measured with the Barthel Index); handicap [measured with the Rankin Scale and the Glasgow Outcome Scale (GOS) (Jennet & Bond, 1975)]. With people who had aphasia a proxy was interviewed together with the patient.

Overall, psychosocial HRQL was poor in 31.7% of the subjects and moderate in a further 11.6%. Depression was present in 48.3% of the respondents. In univariate analysis (chi-square statistics) HRQL was associated with professional and educational level, vascular

territory, aphasia, NIH score, seizures, depression, no return to work and reduced GOS, Rankin and Barthel scores. In multivariate analysis (forward stepwise logistic regression) a good HRQL score was significantly associated only with good NIH score on admission ($p<.01$), return to work and absence of depression.

It is unclear why the authors chose to dichotomise the continuous SIP scores and to use logistic regression. They also performed stepwise regression with 11 independent variables on a group of 65 people, which gives a case to variables ratio of 6 to 1 (40 to 1 recommended in stepwise regression). This has serious implications for the generalizability of the findings.

Seventy three percent of the people previously employed returned to work and of those 73.9% returned to the same occupation. In multivariate analysis, return to work was associated with good NIH and GOS scores, which reflect reduced impairment on admission. In short, two and a half years after the stroke, a third of young stroke survivors had poor psychosocial outcomes and almost half of them suffered from depression. Still, the concept of HRQL is problematic in this study as it includes only psychosocial domains.

Wyller et al. (1998) investigated the correlates of subjective well-being (SWB) in stroke survivors. The authors felt that the content of the HRQL concept varies widely and therefore decided to focus on emotions such as satisfaction and happiness. They, thus, considered the term SWB more appropriate for these aspects. Four items were considered a priori to reflect SWB, that is 'satisfaction', 'strength', 'calmness', and 'cheerfulness'. One of the aims of this study was to examine variables explaining SWB in a large, population-based sample of stroke and stroke-free individuals. Their sample was drawn from an area with 127,000 inhabitants in the middle of Norway. All inhabitants aged ≥20 (85,100) were invited to participate in a health survey and 88.1% (74,977) took part. Of those 1417 reported having suffered a stroke

and a stratified random sample of 1,439 individuals with the same age distribution were used as controls. With regard to the validity of self-reported stroke, a previous study using the same question had 20% false positive and 3% false negative self reports compared with stroke diagnosed according to the WHO criteria (the coefficient of agreement being κ=.79).

The subjects had to complete a postal questionnaire (31 items) covering: perceived health; functional abilities; contact with health care system; general well-being; working conditions and chronic diseases. A nurse visited and collected the questionnaire and also assisted those who needed help to fill it in. No further information is given regarding people with aphasia. The subjects also had to do a second questionnaire (42 items) on: lifestyle; housing; educational level; working conditions; medical symptoms; social support and well-being. Linear regression models for each explanatory variable were used, each including the grouping variable as well (stroke/no stroke). Variables with p<.1 were considered for inclusion in a separate multivariate regression analysis.

A model with 12 explanatory variables explained 50.3% of the variance in SWB. The strongest explanatory variables were stroke, gender (higher SWB in women), age (higher SWB in older age), perceived health, social support, and loneliness and sleep problems (which reflect mood/mental health).

The authors could not explain the gender effect on SWB. They also acknowledged that as this was a cross-sectional study they could not assess whether increased SWB with increased aged was an effect of age or cohort. They quote a study by Mastekaasa & Moum (1984) where the effect of age depended on whether the dependent variable-HRQL- reflected satisfaction more or happiness: self-reported happiness seemed to decline with age whereas satisfaction seemed to increase. As far as the rest of the explanatory variables are concerned

this study corroborated the findings of other studies that well-being is closely associated with perceived health, social support and mood/mental health. The authors, however, did not report on the validity of their self-developed measures for the assessment of SWB and other variables (e.g., social support).

Clarke et al. (1999) used the RNLI as a measure of handicap in stroke survivors. The authors viewed HRQL as a "broad, ubiquitous term that is often undefined and loosely measured" and chose to look at handicap instead as defined by the International Classification of Impairments Disabilities and Handicaps (ICIDH) (WHO, 1980). Their aim was to investigate the effects of impairment and disability following stroke on handicap. Of 450 consecutively admitted stroke patients, almost 60% were excluded due to death, recurrent stroke, residence out of town, or because they were too ill or had inadequate cognitive of language function to participate in the assessment. Of the remaining, 78% (n=145) consented to follow up and 135 were reassessed a year later. As indicated, the RNLI was used for the assessment of handicap. The authors' conceptualisation is interesting here, as the RNLI is a measure of reintegration to normal living and thus an inverse measure of handicap. Impairment was measured with the Adams Hemispheric Stroke Scale (Adams et al., 1987) and the Zung Self-Rating Depression Scale (Zung, 1965) and disability was measured with the FIM. Other factors taken into account included living arrangements, marital status, receipt of rehabilitation, and age, gender, site and type of lesion, pre-stroke disability and handicap and comorbidity. The results were analysed using MANOVA.

They found that increased disability and depression explained 41% of the variance in handicap at 3 months post stroke and with the addition of cognitive disability, impairments from a previous stroke, marital status and gender 44% of the handicap variance at 1 year post stroke was explained. The presence of a spouse, which is an indication of social support, was

found to benefit male survivors at 1 year. The results of this study corroborate the findings of other investigators that depression and functional limitations increase handicap after stroke. It would have been interesting to see what the findings would be if people with more severe strokes had been included.

Lofgren et al. (1999) explored psychological well-being 3 years after stroke. Their subjects were 55 stroke survivors at 3 years who consented to take part from a pool of 100 subjects who were discharged from rehabilitation within a year from their strokes. Psychological well-being was measured with the Philadelphia Geriatric Centre Morale Scale (PGCMS) (Lawton, 1975) which views morale as "a generalised feeling of well-being with diverse specific indicators". These include "freedom from distressing symptoms, satisfaction with self, feeling of syntony between self and environment, and ability to strive appropriately while still accepting the inevitable". Independent variables included: depression (assessed with the MADRS); ADL (assessed with the Katz Index); and motor and sensory function. People with severe aphasia were excluded but still 40% of the subjects had aphasia (not clear how assessed). Information on age, gender, living condition, vision, hearing and diabetes were also collected. Spearman correlation matrices were performed to discover correlations between the PGCMS and subjects' characteristics. Variables showing correlations of >.30 and variables that, a priori, could be presumed to affect well-being were entered into a hierachic cluster analysis to see which variables interacted with the PGCMS.

Overall, 64% of the participants had middle or high morale. In a study of the general population, 91% reported middle or high morale. In terms of which factors predicted morale/well-being, only depression (strongest predictor) and reduced motor function reached significance. Factors like aphasia, social support and comorbidity were not

meaningfully explored in this study. Moreover, the reduced response rate (55%) limits the generalisability of the results.

Carod-Artal et al. (2000) sought to identify variables that could predict HRQL 1 year after stroke. They saw HRQL as incorporating physical, functional, psychological and social health and they used the SF-36 and the SIP to measure it. A cohort of 118 consecutively admitted stroke patients in a stroke unit was followed up and at 1-year follow up, from the 91 living and located patients, 90 took part. Independent variables included age, gender, comorbidity, functional status/disability (Barthel Index, FAI), motor impairment (number of falls) and depression [Hamilton Rating Scale for Depression (Hamilton, 1960)]. Seventeen respondents (19%) had aphasia and proxy respondents were used for 3 of them. The authors reported that a regression model was used to correlate variables but no information is given on the type of regression used and no multiple regression numerical data are reported in the paper.

The authors reported that the strongest predictors of HRQL as measured by the SF-36 were severe disability and depression and of HRQL as measured by the SIP were female gender, depression and severe disability. With regard to the effect of female gender, the authors indicate that the mean age of women at the onset of stroke was 71 years, which is 6 years later than that of men. Women also had worse handicap (Rankin scores) at discharge. With regard to the effect of disability and depression on HRQL their findings corroborate those of other studies. The lack of information on the regression analysis used and on the resulting data makes it hard to judge the applicability and generalisability of their results.

The next section looks at the studies that particularly focused on people with aphasia. Then an overview of the key findings of all the reported studies and the issues arising from their review is presented.

2.3 HRQL and related concepts in people with aphasia

A different strand of research, the field of aphasiology, is devoted to the study of aphasia and people with aphasia. In this area very few studies have looked at HRQL. In reviewing the area LaPointe (1999) pointed out that although various papers and books refer to the area, very few empirical studies exist. Gainotti (1997) attributed the problem of lack of studies on the psychosocial aspects of aphasia to '..the extreme complexity of this field of investigation and to the poverty of research tools enabling investigators to explore it effectively.' Here, the research studies that have looked at various types of psychosocial outcomes are briefly covered before looking at the studies that addressed HRQL with people with aphasia.

In a series of studies, Code, Müller and their co-workers looked at psychosocial adjustment to aphasia (e.g., Müller et al., 1983; Hemsley & Code, 1996; Code, Müller & Herrmann, 1999; Code et al., 1999). Using the Code Müller Protocols (Code & Müller, 1992) they looked at the expectations and the optimism of people with aphasia, their significant others and their Speech and Language Therapists (SLTs) in relation to various psychosocial states and situations. Overall, they noted that these three groups have different perceptions of what are the most relevant issues for psychosocial adjustment. They also noted that these perceptions can change markedly over time. These findings have implications for the use of other people for the assessment of HRQL and emotional outcomes in people with aphasia.

Some studies have looked at psychological well-being after aphasia but they have used scales such as the Ryff Psychological Well-being Scales (Hoen et al., 1997), the Psychological Well-being Index (Lyon et al., 1997) and the *How I feel about myself* measure (Thelander et al., 1994) which have not been extensively tested for their psychometric properties. Similarly, Salonen (1995) looked at the physical, functional and social changes in the life of people with aphasia. She used a self-developed questionnaire without providing any information on its content

and its psychometric properties. These studies reported reduced psychological well-being and HRQL after aphasia and improvements following SLT interventions (Hoen et al., 1997; Lyon et al., 1997; Thelander et al., 1994), but the validity of their findings is questionable.

Two studies used qualitative methodologies to explore the consequences of aphasia. Le Dorze & Brassard (1995) explored the handicap associated with aphasia based on an analysis of the experience of the persons affected by it, i.e., people with aphasia and their relatives or friends. One of their main aims was to understand the consequences of aphasia in the terms used by the people with aphasia and their significant others. They carried out semi-structured interviews with 9 pairs of aphasic person and close relative/friend. The main themes/questions of the interviews included the main consequences of aphasia on their lives and in relation to work and changes in interpersonal relationships. They used Grounded Theory to analyse their results based on the ICIDH model. They found that for people with aphasia their language disabilities lead to considerable handicap in the following ways. They influenced negatively situations involving communication, restricted their activities, altered their interpersonal relationships, lead to loss of autonomy and triggered stigmatisation.

Along the same lines, Parr et al. (1997) addressed the long-term consequences and significance of aphasia through analysis and interpretation of in-depth interviews with a substantial sample of aphasic people. The intention of the study was to explore 'insider' views of aphasia. The study was carefully designed to ensure that the diversity and complexity of the experience of aphasia would be captured and therefore the subjects were purposefully selected. Fifty people with at least five years' experience of aphasia, living in different parts of the United Kingdom were selected and took part in this study. The topic guide included among other things questions on the impact of aphasia on different aspects of life, perceptions and understanding of aphasia and experiences of health, social care, voluntary

and other services. The data were charted on a thematic basis, allowing a sense of the individual participants and their stories to be retained. On the basis of the initial index, a number of matrices were drawn up, each using a set of thematically-linked headings and sub-headings. These were designed to allow respondents' comments on various topics and issues to be juxtaposed in a systematic fashion, thus allowing between-case and within-case analysis.

They found that aphasia made it difficult, not just to communicate with family and friends, but to continue with work and education, to find sufficient money, to pursue interests and maintain lifestyles, to sustain relationships and identities, to access information and to understand and negotiate rights and responsibilities. The authors concluded that aphasia impacts upon the aphasic person as an individual, as a partner and family member, as a part of various institutions, communities and networks (neighbourhood, workplace, church for example), and as a citizen.

Two studies have incorporated HRQL assessments specifically with people with aphasia. Sarno (1997) reported on a study that did not look at HRQL per se but at the influence of age on recovery in aphasia. Among the other areas that were assessed as outcomes was HRQL. The author saw HRQL as reflecting a broad spectrum of consequences of disease, incorporating elements of impairments, disabilities and handicap as well as patients' perceived health-status and well-being (de Haan et al., 1993). HRQL was measured with the Geriatric Evaluation of Relatives Rating Instrument (Schwartz, 1983), the Functional Life Scales (Sarno et al., 1973) and the Caregiver Burden Interview (Zarit et al., 1980). Their subjects were from a cohort of consecutively admitted people with aphasia due to a left-hemisphere stroke, who were right-handed and had no previous history of substance abuse, pre-existing speech and language disorder, severe cognitive decline, dementia, psychiatric disorders, cerebral neoplasm or previous stoke. Patients with transcortical motor,

transcortical sensory and conduction aphasia were also excluded because according to the author "their numbers in our clinical population tend to be small". 59 out of 107 eligible participants took part in the study (48 dropped out) and were classified in three groups: people with fluent aphasia, people with non-fluent aphasia and people with global aphasia (the most severe). They all had intensive SLT interventions and were assessed at 3-month intervals from 3 months to 1 year after stroke. The results were analysed with MANOVA.

Sarno found that the Functional Life Scales scores significantly improved for the fluent and non-fluent groups and the Caregiver Burden Interview for the global group. No significant changes were found for the Geriatric Evaluation of Relatives Rating Instrument. These findings suggest that time and the intervention programme provided resulted in improvements in individual areas that can be seen as HRQL aspects. Such improvements however do not reflect an overall HRQL improvement. It is unclear how the author formed the following conclusion "when the post-stroke aphasic patients are provided with intensive, long-term aphasia rehabilitation services which address language, communication strategies, functional communication, coping skills and psychosocial issues for the first year, all of these areas show continuous improvement with a consequently positive impact on quality of life". The extensive exclusion criteria and the low response rate limit the generalisability of these results.

Cruice et al. (2000b) looked at the performance and usability of HRQL assessments in people with aphasia. These included the Dartmouth COOP charts (Nelson et al., 1987), the SF-36 and the SIP. The instruments were administered to people with aphasia by Speech and Language Therapists (SLTs). After the administration, the SLTs judged the measures on the following usability factors: length, wording of instructions, understanding of instructions, wording of questions, understanding of questions, format of assessment, amount of

assistance required and overall appropriateness. Still, none of the raters had administered all three of these instruments, as three parallel studies were run each using different instruments. The number of participants on each study ranged from 10 to 20, and no information is given on how they were sampled. The authors also looked at the distribution of scores across response categories and inter-rater reliability as further evidence of the accessibility of the assessments to people with aphasia. The rationale behind using inter-rater reliability this way was that the responses of people with aphasia should be sufficiently clear to be interpreted with accuracy by different raters.

Overall, both the SF-36 and the SIP got low ratings on many of the usability factors raised above. Still, the varied distribution of scores and the high inter-rater reliability (Spearman's rho > .98 for all the measures, $p < .01$) were seen as evidence of the accessibility of the measures to people with aphasia. The authors concluded that people with aphasia could respond reliably to HRQL assessments that were administered in an interview format by a SLT familiar with the assessment and experienced in facilitation techniques. They felt that "there is a continuing need for a HRQL assessment that is relevant, practical, robust and 'user-friendly' to people with aphasia".

The results of this study are quite interesting as it is the first study that actually looked at the usability of generic HRQL scales with people with aphasia. Still, the small sample sizes, the lack of information of how the participants were sampled and the lack of information on the severity of their aphasia hamper any judgement on the applicability and generalisability of these results.

In the next section, the main findings of this literature review are summarised. Then, the issues that arise from this review, which hamper the interpretability of the findings, the repeatability of the studies and the generalisability of the results, are presented.

2.4 Overview

2.4.1 Main findings of literature review

2.4.1.1 HRQL in stroke and aphasia

Overall, studies indicated that stroke and aphasia impact on all aspects of life. They lead to worse HRQL than age matched controls (Angeleri et al., 1993), age and gender matched controls (Hackett et al., 2000), age, last occupation and educational level matched controls (Jonkman et al., 1998), healthy controls (Viitanen et al., 1988), controls with an elevated risk for stroke (Duncan et al., 1997), and the general population (Hackett et al., 2000).

Stroke and aphasia result in reduced physical abilities and participation in activities (e.g., Lawrence & Christie, 1979; Gresham et al., 1979; Le Dorze & Brassard, 1995; Parr et al., 1997; Hackett et al., 2000); emotional distress (e.g., Angeleri et al., 1993; Tuomilehto et al., 1995; Wilkinson et al., 1997); and affected family and social relationships and role fulfilment (e.g., Lawrence & Christie, 1979; Angeleri et al., 1993; Le Dorze & Brassard, 1995; Parr et al., 1997; Bethoux et al., 1999; Hackett et al., 2000). Qualitative assessments, in particular, with people with aphasia also revealed affected identity, loss of autonomy and stigmatisation (Le Dorze & Brassard, 1995; Parr et al., 1997).

2.4.1.2 Predictors of HRQL

The main predictors of HRQL in people after stroke include the following

2.4.1.2.1 Physical/functional disability

The majority of the studies reviewed found a strong association between reduced ADL/functional status and diminished HRQL (Ahlsio et al., 1984; Ebrahim et al., 1986; Niemi et al., 1988; Viitanen et al., 1988; Astrom et al., 1992; Astrom, Asplund & Astrom 1992; Kwa et al., 1996; King, 1996; Wilkinson et al., 1997; Jonkman et al., 1998; Neau et al., 1998; Clarke et al., 1999; Lofgren et al., 1999; Carod-Artal et al., 2000).

The relationship between ADL/functional disability and reduced HRQL is not a linear one. Firstly, even in ADL independent people HRQL may be diminished (Ahlsio et al., 1984). Secondly, even when disability remains unchanged HRQL may continue to deteriorate (Bethoux et al., 1999). A number of studies, also, indicated that unlike ADL/functional status, which tends to improve with time, psychosocial problems or overall HRQL may not improve with time (Ebrahim et al., 1986; Tuomilehto et al., 1995; Hochstenbach et al., 1996; Jonkman et al., 1998).

2.4.1.2.2 Depression

Depression has also been repeatedly associated with reduced HRQL (Ahlsio et al., 1984; Niemi et al., 1988; King, 1996; Duncan et al., 1997; Jonkman et al., 1998; Neau et al., 1998; Clarke et al., 1999; Lofgren et al., 1999; Carod-Artal et al., 2000). Other factors reflecting affected mood have also been associated with reduced HRQL, like anxiety (Ahlsio et al., 1984; Astrom et al., 1992; Astrom, Asplund & Astrom 1992) overall distress (Ebrahim et al., 1986) and loneliness and sleep problems (Wyller et al., 1998).

2.4.1.2.3 Social Support

A number of studies have found that aspects of social support seem to affect HRQL after stroke (Osberg et al., 1988; Viitanen et al., 1988; Astrom et al., 1992; Astrom, Asplund & Astrom 1992; King, 1996; Wyller et al., 1998). Very specific associations have also been reported, for example, the presence of a spouse was found to benefit male survivors at 1 year post onset, in Clarke et al. (1999). Yet, a few studies found that certain social factors, that are commonly seen as indicators of social support, did not have an effect on HRQL (i.e., living arrangements in Ahlsio et al., 1984; and social and family conditions in Neau et al., 1998).

2.4.1.2.4 Age

The relationship between age and HRQL after stroke is not a straightforward one. Ahlsio et al. (1984) and Ebrahim et al. (1986) found no age effect on HRQL. Still, Ebrahim et al. (1986) acknowledged that the majority of their subjects were elderly. Some authors found that reduced HRQL after stroke was associated with older age (Astrom et al., 1992; Astrom, Asplund & Astrom 1992; de Haan et al., 1995). Niemi et al. (1988) found that although HRQL was affected in young patients as often as it was affected in older patients, in older patients it was more severely affected.

However, the opposite trend i.e., increasing HRQL with increasing age has also been reported. King (1996) found that older age was associated in specific with socioeconomic HRQL, which was one of the HRQL domains in her study. Wyller et al. (1998) found an association between older age and better well-being. This disagreement may at least partially reflect differences in the way HRQL is conceptualised and measured in these studies, as in some, for example, it is seen as life satisfaction (Niemi et al., 1988; Astrom et al., 1992; Astrom, Asplund & Astrom 1992) and in others as perceived overall health (Ebrahim et al., 1986).

2.4.1.2.5 Stroke variables

A number of studies looked at whether stroke related variables (e.g., type, site, extent of lesion) were associated with HRQL. Some studies have reported that ishaemic strokes lead to worse HRQL outcomes than haemorrhagic strokes (Niemi et al., 1988; Carod-Artal et al., 2000; Hackett et al., 2001). Niemi et al. (1988) also found that hemispheric as opposed to brainstem or unspecified strokes were associated with reduced HRQL. De Haan et al. (1995) found that infratentorial strokes were associated with better HRQL outcomes than supratentorial strokes. However, they found that type of (sub)cortical lesion (infarct versus haemorrhage) and laterality (left versus right hemisphere stroke) had no significant effect on HRQL outcomes. Kwa et al. (1996) found that larger infarct volume was associated with reduced HRQL and Neau et al. (1998) found that vascular territory was associated with HRQL outcome.

In addition, worse initial stroke severity, as indicated by e.g., consciousness at onset (de Haan et al., 1995), National Institute of Health (NIH) score (Neau et al., 1998) or Horn Index score (Osberg et al., 1988) has been associated with poorer HRQL outcomes.

2.4.1.2.6 Comorbidity

Some studies have found that other coexisting health problems, i.e., comorbidity tend to be associated with reduced HRQL after stroke (de Haan et al., 1995; Duncan et al., 1997). Clarke et al. (1999) found that impairments from previous strokes were associated with reduced HRQL outcome.

2.4.1.2.7 Aphasia

From all the studies that included people with aphasia and explored its impact on HRQL, only two found aphasia to be associated with reduced HRQL. These were Kwa et al. (1996)

where severity of aphasia was associated with reduced HRQL and Neau et al. (1998) where aphasia was associated with reduced HRQL in univariate but not multivariate analysis.

2.4.1.2.8 Cognitive Impairment

Cognitive impairment was associated with reduced HRQL in three of the reviewed studies (Niemi et al., 1988; Jonkman et al., 1998 –in univariate analyses but not in multiple regression-; and Clarke et al., 1999). However, in the first two studies cognition was assessed with the WAIS and the WMS, which rely heavily on language. Although people with severe aphasia were excluded, other people with aphasia were included in these studies. The validity of these assessments of cognition is questioned, since they rely on language and thus they could well identify people with aphasia as people with cognitive decline. The third study (Clarke et al., 1999) did not attempt to differentiate between aphasia and cognitive decline. Rather the authors measured "cognitive disability" with the communication and cognition sub-scales of the FIM.

The single study that specifically investigated the role of cognitive impairment on HRQL after stroke (Kwa et al., 1996) did not find a significant impact of cognitive impairment on HRQL. Still, 25% of the subjects could not complete the HRQL measure used, due to communication problems. These subjects were significantly more likely than those tested to have a larger infarct volume, aphasia, and cognitive decline.

2.4.1.2.9 Socioeconomic Status (SES)

Few studies explored the effect of SES on HRQL after stroke. Ahlsio et al. (1984) found no effect of SES on HRQL in their stroke survivors. King (1996) found that increased SES was associated with increased socioeconomic HRQL. Neau et al. (1998) found that higher

educational and professional level were associated with increased HRQL (in univariate analysis only).

2.4.1.2.10 Gender

The majority of the studies did not find a significant gender effect on HRQL. Wyller et al. (1998) found a higher subjective well-being in women. They characterised their finding as surprising and did not offer any explanation for it. Carod-Artal et al. (2000) found worse HRQL in women. They acknowledged that the women in their sample were older than the men at the time of their stroke and also had worse handicap at discharge from hospital. These factors could have contributed to the gender effect.

In summary, the strongest predictors of poorer HRQL after stroke are depression/emotional distress and reduced activities/physical abilities. Other factors that have been associated with poorer HRQL outcomes in some studies include reduced social support and increased comorbidity. The evidence for reduced cognition and presence and severity of aphasia is limited. Various stroke and demographic variables have also been implicated but the evidence is not conclusive.

2.4.2 Issues arising from the literature review

A number of issues limit the interpretability and generalisability of the findings of these studies and the repeatability of the studies in other settings or populations.

2.4.2.1 Lack of conceptual clarity

Many of the reviewed studies did not define the concept of HRQL (Lawrence & Christie, 1979; Angeleri et al., 1993; Kwa et al., 1996; Duncan et al., 1997; Mathias et al., 1997; Bethoux et al., 1999).

A number of studies saw HRQL as life satisfaction (Ahlsio et al., 1984; Viitanen et al., 1988; Astrom et al., 1992; Astrom, Asplund & Astrom, 1992). Niemi et al. (1988) conceptualised HRQL as "a person's subjective well-being and life satisfaction". King (1996) defined HRQL as satisfaction with aspects of life that are important to the individual.

A few studies reviewed did not actually set out to assess HRQL but related concepts that have been seen by others as HRQL. Two of them looked at subjective well-being (Wyller et al., 1998; Lofgren et al., 1999) and Osberg et al. (1988) looked, among other things, at life satisfaction. Ebrahim et al. (1986) looked at social and psychological problems following a stroke. Tuomilehto et al. (1995) assessed perceived health, ADL, and psychosocial status. Clarke et al. (1999) viewed HRQL as "a broad, ubiquitous term that is often undefined and loosely measured" and looked at handicap instead.

Lastly, some studies have used generic measures of HRQL like the NHP, the SF-36 and the SIP (de Haan et al., 1995; Hochstenbach et al., 1996; Wilkinson et al., 1997; Neau et al., 1998; Jonkman et al., 1998; Cruice et al., 2000b; Carod-Artal et al., 2000). It is generally accepted that these studies assume the operational definition accepted in this study, which sees HRQL as incorporating physical/functional, social and emotional health.

This variability in the main concept assessed makes it difficult to draw comparisons between the studies and interpret their results. Moreover, the lack of conceptual clarity in some studies leads to questioning of the validity of the assessments.

2.4.2.2 Measurement variability

A number of approaches have been used in the measurement of HRQL after stroke. Researchers have used a single VAS (e.g., Kwa et al., 1996; Duncan et al., 1997); a questionnaire based interview (e.g., Lawrence & Christie, 1979; Gresham et al., 1979;

Tuomilehto et al., 1995); a generic scale like the NHP (e.g., Wilkinson et al., 1997), the SIP (de Haan et al., 1995; Neau et al., 1998; Jonkman et al., 1998; Cruice et al., 2000b) and the SF-36 (Wilkinson et al., 1997; Hackett et al., 2000; Cruice et al., 2000b; Carod-Artal et al., 2000); or a battery of different tests (e.g., Angeleri et al., 1993; Sarno, 1997). In aphasiology, qualitative methodologies, like semi-structured and in-depth interviewing have also been used (LeDorze & Brassard, 1995; Parr et al., 1997).

Some of these approaches, namely the use of a VAS or the use of a non-psychometrically tested questionnaire, result in confusion as to what the concept of HRQL actually reflects. The latter also prevents other researchers from replicating the study in order to see whether the results will hold in different populations.

The qualitative approaches have provided rich, conceptually broad information and useful insights on the impact of aphasia on people's lives. Still, their approach cannot be routinely used in clinical practice or even incorporated in large clinical trials on stroke. This criticism also applies to the use of a battery of tests, so that each test taps on one domain of HRQL.

Lastly, some studies followed one of the currently dominant approaches in the assessment of HRQL, by using generic HRQL measures. In all of these studies, the researchers encountered problems with people with aphasia, due to their language problems. In de Haan et al. (1995) and in Carod-Artal et al. (2000) proxy respondents were used for people with severe aphasia and in Neau et al. (1998) and Hackett et al. (2000) proxy respondents were used for all people with aphasia. In Jonkman et al. (1998) people with severe aphasia were excluded and in Wilkinson et al. (1997) people with aphasia resulted in missed assessments. In Cruice et al. (2000b) people with aphasia were able to complete the SIP and the SF-36 but a specialist SLT facilitated the administration of the measures. These results indicate clearly

that people with aphasia have difficulty completing generic HRQL measures with no facilitation.

None of the studies reviewed used a patient-based disease-specific measure for the assessment of HRQL in people with stroke. This is probably because the first patient-based stroke specific HRQL measures were published in 1999 (see chapter 3). The advantages of disease-specific measures have been highlighted in chapter 1 and most importantly include increased validity and sensitivity in detecting small but perhaps clinically significant changes (e.g., Patrick & Deyo, 1989). They are also highly appropriate for use in routine clinical practice and clinical trials as they can reduce respondent burden, compared to generic measures, by asking only relevant questions (Guyatt et al., 1986; Bergner & Rothman, 1987).

2.4.2.3 Applicability of findings to people with aphasia

In many of the stroke studies reviewed people with severe aphasia were excluded (Lawrence & Christie, 1979; Ahlsio et al., 1984; Viitanen et al., 1988; Niemi et al., 1988; Astrom et al., 1992; Astrom, Asplund & Astrom, 1992; King, 1996; Duncan et al., 1997; Jonkman et al., 1998; Clarke et al., 1999 and Lofgren et al., 1999). In 3 studies, it is unclear from the available information whether people with communication problems were included or not (Osberg et al., 1988; Hochstenbach et al., 1996; and Mathias et al., 1997).

In the studies that did include people with aphasia, often proxy respondents (usually the main caregiver or a nurse) were used instead of the person with the communication problem (Astrom et al., 1992; Astrom, Asplund & Astrom 1992; de Haan et al., 1995; Tuomilehto et al., 1995; Neau et al., 1998). Due to the highly subjective nature of the concept of HRQL and the documented disagreement between stoke patients themselves and their proxies on

HRQL and psychosocial domains (Knapp & Hewison, 1999; Sneeuw et al., 1997; Code et al., 1999), the applicability of these assessments for people with aphasia is questionable.

Communication problems also resulted in missed assessments (Ebrahim et al., 1986; Kwa et al., 1996 -25% could not do the HRQL assessment; Wilkinson et al., 1997 -11.3% could not do the HRQL assessment). In the Angeleri et al. (1993) study people with aphasia did not complete the BDI. HRQL in this study was derived from a multiple correlation analysis of 4 scales, one of which was the BDI. Missed assessments lead to bias in the responses and, therefore, it cannot be assumed that the results of these studies apply to people with aphasia.

In other studies, quite complex instruments were used to measure HRQL. Some of these studies included relatively high proportions of people with aphasia: in the King (1996) study 20% of the subjects had aphasia, in the Lofgren et al. (1999) study 40% had aphasia and in the Bethoux et al. (1999) study 47% had aphasia. Still, neither these three nor Niemi et al. (1988) give any information on how people with aphasia managed the complex HRQL instruments that were used (Ferrans and Powers HRQL Index; RNLI; PGCMS; and a 45 item questionnaire respectively). People with aphasia probably had difficulty understanding at least some of the items and also had difficulty expressing their responses. They would probably require at least some facilitation from the interviewer in order to understand the items and give their responses (Cruice et al., 2000b). The validity of these assessments in people with aphasia is, therefore, questionable. Only Wyller et al. (1998) indicated that a nurse assisted those who needed help to fill in the questionnaire they used to assess HRQL. Although it is not clearly specified, people with aphasia were probably among those needing help.

In short, drawing overall conclusions on the HRQL of people with aphasia or the factors predicting it and comparing aphasic versus non-aphasic stroke survivors is problematic, due to the above conceptual and methodological issues.

2.5 Summary and conclusions

This chapter reviewed the existing studies on the HRQL of people with stroke and aphasia and on the factors predicting HRQL after stroke. Overall, the studies suggest that stroke and aphasia have a considerable effect on the lives of people, impacting on their physical, emotional and social well-being. Commonly identified predictors of HRQL in people with stroke include physical disabilities and depression to a great extent and comorbidity, social support, aphasia, cognitive decline and various stroke-related and demographic variables to a lesser extent. Still, the clarity of these conclusions is clouded by the conceptual confusion on what the term HRQL is supposed to reflect. Moreover, a number of measurement issues (e.g., partial exclusion of people with aphasia, partial use of proxy respondents with people with aphasia, aphasia resulting in missed assessments, no facilitation of people with aphasia) limit the applicability of these results to people with aphasia.

Another issue arising in this review is the following. Some measurement approaches, namely qualitative interviews and use of a battery of assessments for the different HRQL domains, may provide rich HRQL data but they are unwieldy for use in clinical trials or routine clinical practice. Generic HRQL measures, which are easier to use in such settings seem to be difficult for people with aphasia to use, with no facilitation.

Current research on the HRQL of people with aphasia needs to address these issues. The next chapter describes how they were tackled in this study.

3 THE ASSESSMENT OF HRQL IN PEOPLE WITH CHRONIC APHASIA FOLLOWING STROKE: APPROACH, RESEARCH QUESTIONS AND METHODOLOGY

The broad aim underlying this research was to assess the HRQL of people with chronic aphasia following stroke. In the process of defining the aims more precisely and assessing the feasibility of the overall research, certain considerations were taken into account. Criteria were set in order to address some of the challenges in existing literature in the area. This chapter begins by describing these criteria under approach. The specific aims of the study are then presented under research questions and the chapter closes with the overall methodology of the study.

3.1 The approach in this study

The measurement approach that was followed aimed to address some of the challenges identified in existing literature on HRQL in general and in stroke and aphasia in particular, by meeting the following criteria:

3.1.1 Conceptual clarity

In this study HRQL was conceptualised as reflecting the impact of a health state on a person's ability to lead a fulfilling life (Bullinger et al., 1993). It incorporated the individual's subjective evaluation of his/her physical, mental/emotional, family and social functioning (Berzon et al., 1993; Hays et al., 1993; de Haan et al., 1993). This operationalisation reflects the subjective nature of HRQL and identifies the main domains of the concept that need to be assessed.

3.1.2 *Usability in a broad range of clinical practice*

Measurement approaches of HRQL in stroke and aphasia have included a single Visual Analogue Scale (e.g., Kwa et al., 1996); a scale (e.g., Wilkinson et al., 1997; de Haan et al., 1995; Neau et al., 1998; Jonkman et al., 1998; Dorman et al., 1999; Hackett et al., 2000); a battery of different tests (e.g., Angeleri et al., 1993); or semi-structured or in-depth interviewing techniques (e.g., Lawrence & Christie, 1979; LeDorze & Brassard, 1995; Parr et al., 1997).

This study aimed to assess HRQL in a way that could be replicated in clinical practice. A viable way of assessing HRQL in people with aphasia in clinical practice is by use of a single HRQL measure. Administering a single scale is less time consuming and causes less burden to the respondents than a battery of tests or assessments. Recording and analysing the data of a scale is also less time consuming than recording and analysing the data of semi-structured or in-depth interviews. Extra advantages of scales include the following: they allow better checks of their reliability and validity than Visual Analogue Scales or interviews and they are easier to replicate than interviews (e.g., Singleton & Straits, 1999). The use of a single scale for the assessment of HRQL was seen as the method that was most feasible and least susceptible to error for use in clinical practice, and it was, thus, the preferred method for use in this study.

3.1.2.1 *Choice of HRQL scale*

In chapter 1 the two main types of HRQL scales, generic and disease-specific, were discussed. Ideally research on HRQL should incorporate both generic and disease-specific instruments as they complement each other (e.g., Fletcher et al., 1992; Hawker et al., 1995). This was not feasible in the present study mainly for two reasons. First, there was not sufficient time or resource to modify two instruments related to the same concept in order to

make them communicatively accessible to people with aphasia. Second, the administration of two measures for the same concept would increase respondents' burden, a major consideration for this group. It was decided to choose a disease-specific measure as disease-specific measures have greater content validity (e.g., de Haan et al., 1993) and can have greater sensitivity and responsiveness to change for specific population groups than generic measures (e.g., Patrick & Deyo, 1989;). Such a measure would, thus, allow for a more comprehensive assessment of the HRQL of people with aphasia.

The targeted population was people with aphasia following stroke and as there was no single measure for the assessment of HRQL in people with aphasia, a stroke-specific HRQL scale had to be selected. At the time of the last literature search, prior to the data collection commencing (September 2000) there were three stroke-specific HRQL scales in the English language. These were the SA-SIP30 (Stroke-adapted 30-item version of the SIP) (van Straten et al., 1997), the SS-QOL (Stroke-specific Quality of Life scale) (Williams et al., 1999a) and the SIS (Stroke Impact Scale) (Duncan et al., 1999). The choice between these three measures was based on the following criteria:

1. the conceptualization of HRQL of the measure should be identical or as close as possible to the operational definition of HRQL of this study

2. as the concept of HRQL is highly subjective, the measure should be patient-based, i.e., its content should be derived through consultation with stroke patients

3. the measure should have been tested at least for reliability and validity and should have good properties in these areas

4. in terms of accessibility to the population under study, the measure should meet as many as possible of the following criteria: be linguistically simple, be relatively short to reduce respondent burden, have one or maximum two response formats to avoid confusing people with aphasia and have the same direction throughout the response format(s), again to avoid confusing people with aphasia.

The SA-SIP failed criterion 2. The SS-QOL and the SIS both met criteria 1, 2 and 3. The SS-QOL met more of the criteria under 4 as it was shorter than the SIS (49 as opposed to 64 items), had only 2 response formats (as opposed to 3 in the SIS) and the direction of the response formats was the same throughout the instrument. The SS-QOL was, therefore, selected as the scale for the measurement of HRQL in this study. The SS-QOL is described fully under 'measures' in this chapter.

3.1.3 Accessibility

Given the subjective nature of HRQL, the best people to report on it are those whose HRQL is being assessed. For this reason, assessments used should be accessible to the population under study. People with aphasia would require at least some modifications of the testing materials and special skills on behalf of the interviewer in order to access the assessments and express their responses.

The author, a Speech and Language Therapist (SLT) experienced in working with people with aphasia, carried out all the assessments in an interview format, in order to facilitate the understanding and communication of people with aphasia. All materials were shown to participants in an accessible format so that they could read the items while the interviewer said them. To facilitate participants' response, they only had to point to their responses and the interviewer recorded them. Materials used had been previously reviewed for their level of

linguistic complexity. Although their content (in terms of meaning) remained unchanged to avoid invalidation, their presentation was modified to make them more communicatively accessible. In particular, few items were presented per page. Practice items were introduced at the beginning of each questionnaire to make sure the respondent understood the response format and what s/he had to do. Larger font was used (14-16) and key words were presented in bold (Hilari & Byng, 2001).

Despite increasing the accessibility of materials used for people with aphasia, it was anticipated that some participants would have such severe aphasia they would be unable to self-report. In the pre-test of the study (chapter 4) the level of aphasia severity was established below which participants would be unable to self-report on the HRQL measure. For these participants, with their consent, proxy respondents were used for all measures that required language. All proxy results obtained in this study are not included here and will be analysed separately in another study, due to the reported disagreement between self-report and proxy data (Knapp & Hewison, 1999; Sneeuw et al., 1997; Sprangers & Aaronson, 1992).

3.2 Research questions

This project aimed to increase our understanding of the HRQL of people with long-term aphasia following stroke, by addressing the following questions:

1. Can an acceptable, reliable and valid version of the SS-QOL be developed for people with aphasia?

   This involved:

   a. Development of an aphasia-adapted version of the SS-QOL.

b. Evaluation of the psychometric properties of the aphasia-adapted version of the SS-QOL with people with long-term aphasia after stroke.

Following the development and the psychometric testing of the aphasia-adapted version of the SS-QOL, the following question was addressed:

2. What are the predictors of HRQL, as measured by the aphasia-adapted version of the SS-QOL, in people with long-term aphasia after stroke?

A study was designed to address both these questions, as described below.

3.3 Methodology

3.3.1 Research question 1a

The development of an aphasia-adapted version of the SS-QOL comprised modification of the instrument to make it communicatively accessible to, and increase its content validity with, people with aphasia and pre-testing of the modified version with people with aphasia. The modification involved consultation with professionals with expertise in instrument development, language and aphasia and a small pilot study with people with aphasia. The pre-testing aimed to evaluate the instruments content validity further, to assess its accessibility and initial acceptability by people with aphasia and to identify whether any revisions were necessary before testing its psychometric properties further in a large sample. It involved people with aphasia who attended the groups of the City University Centre for people with aphasia and who were at least one year post-stroke. Further details of the methods used to address research question 1a are given in chapter 4: 'Development of an aphasia-adapted version of the SS-QOL: methods and results'.

3.3.2 Research questions 1b and 2

This section presents the overall design, the participants, the procedure and the measures that were used to evaluate the psychometric properties of the aphasia-adapted SS-QOL (research question 1b) and to assess the predictors of HRQL in people with chronic aphasia after stroke (research question 2). Details of how the data were analysed are presented in chapter 5 for research question 1b and in chapter 7 for research question 2.

3.3.2.1 Design

A cross-sectional survey study was undertaken. A questionnaire-based interview was administered and data were collected on HRQL using the aphasia-adapted SS-QOL, on various concepts against which the construct validity of the aphasia-adapted SS-QOL would be assessed (for research question 1b) and on potential predictors of HRQL (for research question 2).

The potential predictors included demographic, stroke and health related variables and variables that have been identified as predictors of HRQL in people with stroke. These were: age, gender, ethnic background, marital status, socioeconomic status, employment status (demographic variables); type of stroke, time post onset of stroke, and number of other comorbid conditions (stroke and health related variables); and psychological distress, level of activity, communication disability, cognition and social support (previously implicated variables). One extra variable of theoretical interest was also used as a potential predictor. This was satisfaction with services for stroke. This was because, as has been already indicated (chapter 1), some theorists believe that for people with chronic illnesses and disabilities HRQL should incorporate satisfaction with medical care received (Pearlman & Uhlman, 1988; 1991).

3.3.2.2 Participants

Due to the documented long-term impact of stroke and aphasia, the targeted population were people with long-term aphasia following stroke. To find people with long-term aphasia it was necessary to recruit from community settings. A two-stage sampling frame was adopted. In the first stage, clusters were identified, i.e., community services for people with aphasia. In the second stage, potential participants were identified in the approached clusters.

The clusters were 2 SLT Service Providers, one inner city (Community Health South London NHS Trust -Lambeth and Southwark-) and one semi-rural (Oxleas NHS Trust -Queen Mary's Hospital-), and a not-for-profit organisation for people with aphasia (Connect, the communication disability network). The aim was to recruit a varied sample in terms of ethnic background and age. All recruiting sites were in Southeast England, which has a multi-ethnic population (South London in particular). Connect and Queen Mary's were targeted because they offer long-term services to people with aphasia and therefore they are likely to have a larger caseload than other sites. Connect also has services for younger people. The inclusion criteria were aphasia due to a stroke of at least 1-year duration, no known pre-stroke history of severe cognitive decline or mental health problems and living at home prior to the stroke.

3.3.2.3 Procedure

Ethical approval was obtained from City University and the participating sites' research ethics committees (appendix 3.1). The recruitment and data collection period was 11 months with 3-4 months spent on each site. The aim was to recruit at least 80 participants. In the participating sites, review of SLT records was undertaken to identify eligible participants. All eligible participants were given information on the project (both face to face or on the phone and in writing) and were asked to take part in the study. Participants were given an information booklet on the project, which had the consent form at the end, and the

interviewer went through this booklet with each participant (appendix 3.2). Consent was obtained in writing since clinical care was not the primary purpose of the contact with the participants (Department of Health, 2001). Consent was obtained at least 2 days after the main information on the project was given, in order to give time to the participants to absorb the information and make their decision (Department of Health, 2001)[2].

All self-reporting participants were interviewed twice at home or in their SLT site. For people with such severe aphasia that proxy respondents were used, only one interview was held. In this interview the severity of the participant's aphasia was assessed (see measures below, FAST) and the participant completed the only other measure that did not require language (see measures below, RCPM). The proxy respondent (usually the spouse/partner or the main carer of the person with aphasia) then completed the rest of the assessments. At the end of the interview(s) the participants were informed that they would receive a summary of the main findings of this project once it was completed.

3.3.2.4 Measures[3]

3.3.2.4.1 HRQL

HRQL was assessed with the aphasia-adapted version of the SS-QOL. The development of this measure and its content are fully described in chapter 4 and a copy of the instrument is shown in appendix 4.3. Here the original instrument is described.

The SS-QOL is a stroke specific quality of life scale. The authors see HRQL as the physical, psychological, and social aspects of life that may be affected by changes in health states. The

---

[2] Although the Department of Health guidelines for good practice in consent primarily refer to consent to examination and treatment, the authors indicate that "the same principles apply to consent in research as in clinical practice" (www.doh.gov.uk/consent).

[3] The measures' scoring sheets that show the items of the measures have been included in the appendices. The scoring sheets that do not reflect the content of their respective measures have not been included in the appendices.

SS-QOL is a patient-derived measure: to establish domain and item content validity the developers held focused interviews with stroke survivors to identify the domains most affected by their stroke (Williams et al. 1999a). The instrument was specifically designed for use in clinical trials and it is, thus, a relatively easy and quick to administer measure.

The SS-QOL has 49 items and is divided in two parts, which cover 12 domains. Part 1 is a list of questions that ask how much trouble the respondent had in the past week with activities in the areas of self-care, vision, language, mobility, work and upper-extremity function. The response format is a 5-point scale ranging from 'couldn't do it at all' to 'no trouble at all'. Part 2 is a list of statements with which the respondent has to agree or disagree. They cover the areas of thinking, personality, mood, family roles, social roles and energy. Table 3.1 shows examples of two items of the questionnaire and a full copy is presented in appendix 3.3.

Table 3-1: An item from part 1 and an item from part 2 of the SS-QOL and their response formats

	Couldn't do it at all	A lot of trouble	Some trouble	A little trouble	No trouble at all
SC1. Did you have trouble preparing food?	1	2	3	4	5

	Strongly agree	Moderately agree	Neither agree nor disagree	Moderately disagree	Strongly disagree
P2. I was impatient with others	1	2	3	4	5

Scores are calculated separately for each domain and then an average can be calculated from the subdomains' mean scores. Scores range from 1 to 5 with higher scores indicating better function.

The psychometric properties of the SS-QOL were tested in 72 people who had suffered an ischaemic stroke, at 1 and 3 months after their strokes. The SS-QOL domains had good internal consistency (Cronbach's alpha values for each domain $\geq .73$). Most of them also had good convergent validity as indicated by moderate correlations with similar domains of established outcome measures (r^2 range .30 to .50). The exceptions to this were the language (r^2=.10), the social roles (r^2=.01), the thinking (r^2=.00), the upper extremities (r^2=.18) and the vision (r^2=.11) domains. The authors attributed the lack of moderate correlations with the external measures of the language and thinking domains to the fact that people with language or cognitive deficits were excluded; and of the upper extremities domain to a ceiling effect on the external measure. Most domains were responsive to change with standardized effect sizes (SES) >.40, except for the energy (SES=.36), the personality (SES=.20) and the thinking domain (SES=.36). One- and 3- month SS-QOL average scores were associated with patients' self-report of their HRQL compared to before their strokes (p<.001).

3.3.2.4.2 Measures used in the psychometric evaluation of the aphasia-adapted SS-QOL

A number of measures were used against which the validity of the overall adapted SS-QOL and its subdomains was tested. Chapter 5 describes fully the validation process, i.e., how validity was tested and what correlations were expected between external measures and SS-QOL average and subdomains' scores. This section presents the areas in which external measures were needed and describes the measures that were chosen.

To test the validity of the adapted SS-QOL external measures were needed in the following areas: language and communication, cognition, emotional well-being/distress, activities and social support (see chapter 5 for details on why and how these areas were chosen). A number of measures were reviewed in each of these areas for their psychometric properties, their linguistic complexity, their length (to reduce respondent burden) and for their applicability and acceptability for people with aphasia. The following measures were chosen.

Aphasia was screened with the Frenchay Aphasia Screening Test (FAST) (Enderby et al., 1987). The FAST covers the 4 major areas of language that can be affected by aphasia: comprehension and reading (receptive domains) and expression and writing (expressive domains). It has 30 items and the overall score ranges from 0 to 30, with higher scores indicative of better language function. It is a short and quick to administer measure taking three to ten minutes to complete (Enderby et al., 1987). The FAST was validated against established measures of language function, i.e., the Functional Communication Profile (FCP) (Sarno, 1969) and the Shortened Shuell (Thompson & Enderby, 1983) with good results. In terms of measuring severity of aphasia, the FAST correlated with the FCP with $r = 0.96$ ($p<0.001$) for the late post onset group (1 to 3.5 years post onset).

Participants' communication and the extent of their communication disability was further assessed with the American Speech and Hearing Association Functional Assessment of Communication Skills for Adults (ASHA-FACS) (Frattali et al., 1995). The ASHA-FACS asks about communicative activities that people with aphasia perform and whether they perform them independently or with assistance. Scores range from 1 to 7 with high scores indicative of communicative independence. The 44 items of the ASHA-FACS cover the following four domains: social communication, communication of basic needs, reading, writing and number concepts and daily planning. Examples of items include requesting

information of others, explaining how to do something, expressing feelings and writing messages. It is rated by the SLT of the person with aphasia, based on observations of this person or observations by others who are familiar with the person. This was seen as an advantage for this study, as a number of questionnaires had to be administered to the respondents.

The psychometric properties of the ASHA-FACS were tested in a sample of 131 people with aphasia and 54 people with cognitive communication disorders (N=185). The instrument has good inter- and intrarater reliability (*r* range: .88-.99), high internal consistency and good convergent validity against external measures (all $r > .40$). Principal Components Analysis suggested that there was one underlying latent component and Factor Analysis supported the 4-domain structure of the instrument. A copy of the instrument's scoring form is presented in appendix 3.4.

Participants' emotional distress was assessed with the General Health Questionnaire – 12 item version (GHQ-12) (Goldberg, 1972). The GHQ is a measure of distress that has been extensively used as a screening tool for psychiatric disorders. It was designed to identify two main types of problems: "the inability to carry out one's normal functions, and the appearance of new phenomena of a distressing nature" (Goldberg & Hillier, 1979). The original 60-item version covers four elements of distress: depression, anxiety, social impairment, and hypochondriasis. The shortened versions (30-, 28-, 20-, and 12-item) do not include 12 items on somatic symptoms that were answered positively by physically ill people (Goldberg, 1972; Goldberg & Hillier, 1979). This reduces the number of false positive responses. In general practice patients, the GHQ-12 has a sensitivity of 93.5% and a specificity of 78.5% (Goldberg & Williams, 1988) whereas in hospital outpatients, its sensitivity is 74.2% and its specificity 95% (McDowell & Newell, 1996).

Its psychometric properties have been extensively tested with very good results (for reviews see Goldberg & Williams, 1988; Vieweg & Hedlund, 1983). The GHQ has also been used in stroke studies (e.g., Ebrahim et al., 1986; Dennis et al., 1997; Dennis et al., 2000). O'Rourke et al. (1998) compared the Hospital Anxiety and Depression Scale (HADS) (Zigmond et al., 1983) with the GHQ-30 and found they exhibited similar levels of sensitivity and specificity with stroke patients. Johnson et al. (1995) compared the HADS and the GHQ-28 with stroke patients and reported that the GHQ was superior to the HADS in detecting both anxiety and depression. Johnson et al. (1995) also compared the GHQ-28 with the HADS, in terms of screening for depression and anxiety after stroke. The patients were diagnosed with the DSM-III criteria. They found that the GHQ was superior to the HADS and it had the best overall performance for specificity, sensitivity and predictive validity. The GHQ-12 was chosen in the present study as a highly appropriate measure for the assessment of emotional distress in people with stroke. It was preferable to the longer versions of the instrument as it has comparable psychometric properties and yet it is considerably shorter, reducing respondents' burden. A copy of the instrument's scoring form, is presented in appendix 3.5.

To assess cognition, the Raven Coloured Progressive Matrices (RCPM) (Raven, 1962) were used. The RCPM measure uses non-verbal symbols to assess cognition, it does not require verbal responses from the respondents and only minimal verbal instruction is necessary. As such it is, to the best of our knowledge, the most valid instrument for the assessment of cognition in people with language impairments. Research has also demonstrated that it is equally reliable for different ethnic groups (e.g., Carlson & Jensen, 1981). The RCPM has been used to explore cognitive decline in brain damage and aphasia (e.g., Villardita, 1985). The coloured (RCPM) rather than the standard matrices (SPM) were preferred as they are considerably shorter, reducing respondent burden. Smits et al. (1997) highlight two extra

advantages of the RCPM. The matrices themselves are coloured large-print drawings, which are visible for older subjects with modestly impaired eyesight. Each part of the test starts with easy items, which is encouraging for the respondents as they can answer at least some of the items correctly. The psychometric properties of the RCPM have been extensively tested with very good results (e.g., Raven et al., 1995) and normative data exist for different population groups, but not for young adults in Britain. The RCPM scores were therefore converted to SPM grades (Raven et al., 2000) as in the present study younger participants were included. SPM grades range from I to V and they represent percentile ranks (I at or above the 95th percentile, 'intellectually superior' and V at or below the 5th percentile, 'intellectually impaired').

Participation in activities was explored with the Frenchay Activities Index (FAI) (Wade et al., 1985). The FAI is a measure of general (i.e., other than personal care) activities of stroke patients, which has been standardised on a sample of 976 stroke patients (seen just after the stroke, and at 3- 6- and 12-months post onset). The FAI consists of 15 items that cover domestic, social, leisure activities and work. There is one item on work and a few on household tasks. This bias is deliberate as the stroke-prone population is the elderly and routine daily chores take up much of their time and are of importance (Wade et al., 1985). It is interviewer administered and the respondent is asked about the frequency with which s/he performed each activity over the past 3 or 6 months. The emphasis is placed on the frequency of the activities rather than e.g., quality or satisfaction to reduce subjectivity. The overall score ranges from 0 to 45 with high scores indicating frequent participation in activities. The FAI was tested for validity, interrater reliability and sensitivity to change with good results (Wade et al., 1985). A copy of the instrument's scoring form is presented in appendix 3.6.

Social support was assessed with the Social Support Survey (SSS) (Sherbourne & Stewart, 1991). The SSS assesses the perceived availability of four types of support (tangible, emotional/informational, social companionship and affectionate support). It consists of 19 items, i.e., it is brief enough to minimize respondent burden. The response format is a 5-point scale going from 'none of the time' to 'all of the time'. Average scores are calculated ranging from 1 to 5 with high scores indicative of higher social support. It has a sound theoretical basis guided by theory on the most important dimensions of support (e.g., Cohen & Syme, 1985; Cohen & Wills, 1985). Its items were designed specifically to be short, simple and easy to understand, restricted to one idea in each stem. It has very good psychometric properties (Sherbourne & Stewart, 1991), which were tested on a group of chronically ill outpatients. It is, therefore, applicable to patient populations who may have greater than average needs for various forms of social support (Sherbourne & Stewart, 1991). All this makes it particularly suitable for use with people with aphasia. A copy of the scoring sheet of the SSS is presented in appendix 3.7.

3.3.2.4.3 Measures used as potential predictors of HRQL

Potential predictors were: age, gender, ethnic background, marital status, socioeconomic status, employment status (demographic variables); type of stroke, time post onset of stroke, and number of other comorbid conditions (stroke and health related variables); and psychological distress, level of activity, communication disability, cognition, social support and satisfaction with services for stroke (other variables).

Information on demographic, stroke related and comorbidity variables were collected from the participants' SLT notes. They were confirmed and supplemented through a short interview with the participants. This information was recorded on a case history form, which is presented in appendix 3.8.

Participants' socioeconomic status was determined using the revised socioeconomic classification (SEC) proposed by the Office of National Statistics (Rose & O' Reilly, 1997). The short 'collapsed version' was used, which classifies people in eight groups. In terms of its conceptual basis, the developers stress that the revised SEC follows a well-defined sociological position that employment relations and conditions are central to delineating the structure of socio-economic positions in modern societies. As most people after stroke are not in employment, their SEC was determined using their last occupation before the stroke. Recent research suggests that classifying individuals not currently in paid employment by their last main job is a satisfactory procedure even for those who have been out of the workforce for many years (Marshall & Roberts, 1996; Arber, 1997). As Marshall & Roberts (1996) point out "the fact that individuals are not in employment at any particular time does not mean that they have dropped out of the class structure, and are no longer affected by their earlier class experiences". Recording last main job to measure class for all people not currently in paid work, including women who have not worked for many years and unemployed, disabled or retired people is also the position of the Review Committee of the SEC (Rose, 1997). A copy of the SEC is presented in the last page of the case history form in appendix 3.8.

Patient satisfaction with stroke care was measured with the Patient Satisfaction Index (PSI) (Pound et al., 1994). The developers used in-depth interviews with people with stroke to develop the instrument. The PSI was examined for test-retest reliability, internal consistency, content and convergent and discriminant validity and was found to be valid and reliable (Pound et al., 1994; Pound et al., 1999). The instrument consists of 12 items on stroke care, which can be grouped in inpatient care, therapy/recovery and services after discharge. The response format is a 4-point scale, going from 'strongly agree' to 'strongly disagree'. Total

scores are calculated and high scores are indicative of more satisfaction. A copy of the scoring sheet for the PSI is presented in appendix 3.9.

The rest of the potential predictors were measured with the instruments already mentioned, i.e., emotional distress with the GHQ-12, level of activity with the FAI, communication disability with the ASHA-FACS, cognition with the RCPM and social support with the SSS.

3.4 Summary

This chapter described the approach of this study on the assessment of HRQL in people with aphasia, the main aims of the study and the methodology used to address these aims. It was aimed to assess the HRQL of people with long-term aphasia following stroke in a way that can be replicated in clinical practice and that would allow the majority of the participants to self-report. The study addressed two questions: Can a valid and reliable version of the SS-QOL be used for the assessment of HRQL in people with aphasia? This involved development of an aphasia-adapted version of the SS-QOL and testing of its psychometric properties. The second question explored what were the main predictors of HRQL in people with long-term aphasia following stroke. The methodology used in the development of an aphasia-adapted version of the SS-QOL was briefly described here. Further details of this process and the results of the adaptation are presented in chapter 4. The design, the procedure, the participants and the measures used for the psychometric evaluation of the aphasia-adapted version of the SS-QOL and for exploring the predictors of HRQL using this instrument were also presented in this chapter. Further details on the methods and the data analysis of the psychometric evaluation are given in chapter 6 and the results in chapter 7. Chapter 8 describes further the methods used to analyse the data on the predictors of HRQL and the results of these analyses.

4 DEVELOPMENT OF AN APHASIA-ADAPTED VERSION OF THE SS-QOL: METHODS AND RESULTS

The development of an aphasia-adapted version of the SS-QOL was undertaken in two stages. In the first stage, the instrument was modified to make it communicatively accessible to people with mild to moderate receptive aphasia and to increase its content validity with this population group. In the second stage, the modified version of the SS-QOL was pre-tested in a group of people with aphasia. The main aims of the pre-testing were to evaluate the instrument's content validity, accessibility and acceptability with people with aphasia and to identify whether any revisions were necessary before testing its psychometric properties further in a large sample. This chapter covers the modification of the SS-QOL and the pre-testing of the aphasia-adapted version of the instrument.

4.1 Modification of the SS-QOL[4]

4.1.1 Methods: review of the SS-QOL by expert professionals

To identify what aspects of the SS-QOL needed to be modified and how, professionals with related expertise reviewed the instrument. Firstly, two focus groups were held with Speech and Language Therapists (SLTs) specialist in aphasia. Prior to the focus groups the SLTs were given a copy of the questionnaire, a copy of the Williams et al. (1999a) original paper on the development of the SS-QOL and a sheet with some brief information on the instrument and questions to think about for the focus group discussion (appendix 4.1). The questions were:

[4] This work has been presented and published in Hilari (2000) and Hilari & Byng (2001).

1. *What do you think overall of the SS-QOL?*
2. *What difficulties you think people with aphasia may have completing the SS-QOL?*
3. *How can we modify it to make it more aphasia-friendly?*
4. *Are there any areas/questions that you think should be added to the instrument?*

During the focus groups, the participants were encouraged to make any other comments they had regarding the SS-QOL.

Secondly, the instrument was reviewed by a linguist with a specialist knowledge on communication disability, who made suggestions on how the language could be simplified. It was also reviewed by a questionnaire development expert from the National Centre for Social Research who advised on the format and presentation of the SS-QOL.

4.1.2 Results and modifications: review of the SS-QOL by expert professionals

On the whole the reviewers felt that the measure had good face validity, i.e., on the face of it, it seemed to cover the concept it was intended to measure (Singleton & Straits, 1999). This section covers the areas that were raised as needing modification and the modifications that were suggested and implemented.

4.1.2.1 Mode of administration

The administration of the SS-QOL involved giving respondents a paper version of the questionnaire to complete. This would pose problems for people with aphasia with reading difficulties. Alternatively one could ask the questions in an interview format. This would pose problems for people with aphasia with speech understanding problems. Some of the respondents could also have difficulty marking their responses due to writing difficulties. It was therefore decided to administer the aphasia-adapted version of the SS-QOL in the following way. An interviewer would show the respondent each item on a paper version of the instrument (presenter's form), while at the same time reading out loud each item. This

way, people with reading problems could rely mostly on listening to the interviewer whereas people with understanding speech problems could rely mostly on reading the items. The respondent would then just point to their response which the interviewer would mark on a scoring sheet.

4.1.2.2 Presentation of the instrument

The layout of the original instrument had too many items per page, resulting in too much information per page for a person with language difficulties. In addition, a lot of items from different areas were grouped together, which again could be confusing for people with aphasia. The font was so small that it could be a challenge for older respondents with visual problems.

The following modifications were made to address these issues. The items of each of the original 12 subdomains were presented on a separate page resulting in fewer items per page (3-6 depending on the domain). The font was increased to 14. These modifications ensured that the instrument had a simple "uncluttered" format with sufficient space between items as recommended by Woodward & Chambers (1991).

Transitional questions were introduced between different domains to set the context of the questions coming up, e.g., "the next set of questions asks about your family and social life". Practice items were also introduced at the beginning of the two parts of the questionnaire, to ensure that the respondent understood the response formats and what s/he had to do (pp 3, 12 in appendix 4.3).

4.1.2.3 Linguistic and reading complexity

Overall, the majority of the items were straightforward. Still some of them were longer than others and would require greater effort to read and process. Some items would need modification from American to British English.

The following modifications were made to reduce linguistic and reading complexity. Consecutive items in the SS-QOL started in the same way e.g., in the first part

> "Did you have trouble preparing food?"
> "Did you have trouble eating, for example, cutting food or swallowing?"
> "Did you have trouble getting dressed, for example, putting on socks or shoes, buttoning buttons, or zipping?"

To reduce reading demands, the overall question was placed at the top of each page (lead in question) and then the items followed, e.g.,

> "How much trouble did you have"
>
> "Preparing food?"
> "Eating?"
> "Getting dressed?"

The length of the items was also reduced. For items that included examples (e.g., see examples on 'eating' and 'getting dressed' above), the examples were not included in writing in the modified version. Instead, the interviewer would give them orally. Some redundant information was also omitted, e.g., "...when bending over or reaching for things?" was changed to "...when bending over or reaching?"

To further reduce reading demands and to facilitate understanding of the items by focusing on the essential information key words were emboldened. Clinical experience with people with aphasia indicates that their reading comprehension is often facilitated when the key words stand out by being printed in bold. In the SS-QOL key words were identified as those

conveying crucial information, compared to the rest that could be inferred given the context. For example "Finishing jobs that you started?" was presented "**Finishing jobs** that you started?". The respondent is not expected to miss the information carried in the non-bold words, as the instrument is interviewer administered, with the interviewer asking the whole item, while the respondent can at the same time read it.

Examples of changes from American to British English included the following: "seeing the television well enough to enjoy a show" was initially changed to "... seeing the television well enough to enjoy a programme" and then, to reduce length as well to "seeing the television well enough to enjoy it". "Buttoning buttons" and "zipping a zipper" were changed respectively to "doing buttons" and "doing a zip"

4.1.2.4 Content validity

Content validity refers to whether a measure adequately covers all aspects of the concept to be measured (Streiner & Norman, 1995). People with aphasia were not included in the process of developing the SS-QOL, and the reviewers felt that this had implications on its content validity with this group of people. It was pointed out that language minimally involves speaking and understanding what is said and the language domain of the SS-QOL had no items on understanding other people.

It was also noted that the family roles and social roles domains of the SS-QOL had an item each on the effects of physical problems on family life and social life respectively. People with aphasia and no physical problems might think that physical problems included their aphasia. To avoid such confusion it was suggested to include an item in each of these domains about the effects of language problems on family and social life.

Moreover, research on differences in cognition between people with right hemisphere

damage (RHD) and left hemisphere damage (LHD), which can result in aphasia, has indicated that people with LHD have specific difficulties with decision-making (e.g., Tartaglione et al., 1991). The clinical experience of the SLTs involved in the focus groups supported this evidence as they indicated that people with aphasia often complained about difficulties in decision making after their strokes.

All this resulted in the following modifications. One item was added on the language domain on understanding what other people say (stem: "how much trouble did you have", item: "understanding what other people say"). Two items were added, one in the family roles and one in the social roles domain, to reflect the effects of language problems in these domains (stem: "did you", items: "feel that your language problems interfered with your family life", "feel that your language problems interfered with your social life". One item on making decisions was included in the thinking domain (stem: "did you", item: "find it hard to make decisions"). The content validity of the measure was tested further in the pre-test of the study with people with aphasia.

4.1.2.5 Strongly agree-strongly disagree response format

The response format of the second part of the SS-QOL was identified as a potential challenge for people with aphasia. It is a 5-point Likert scale going from 'strongly agree' to 'strongly disagree' (SA-SD). The SA-SD response format is a popular yet controversial response format in social research (e.g., Fowler, 1993). It is generally used to measure attitudes and beliefs and it seems to fit well attitudinal statements, e.g., 'Abortions should be illegal' where the respondents opinion is sought. The second part of the SS-QOL however is a list of problems or feelings that people may have after stroke. It is not their opinion that is sought but rather their experience or feelings. Using a SA-SD response format results in respondents having to agree or disagree with feelings e.g., "I was irritable" or problems like

"I had to write things down to remember them". Moreover, the SA-SD response format is linguistically complex both in terms of word length and word meaning. Thus, the reviewers felt that people with aphasia may have considerable problems answering the second part items.

It was felt essential to involve people with aphasia in the decision making process on whether indeed the SA-SD format was difficult and whether it needed to be replaced with a potentially simpler response format. A small pilot study was, therefore, undertaken with people with aphasia to address this issue (see below).

4.1.2.6 Negative items

Some of the items in the second part of the SS-QOL (that had the SA-SD response format) were negative (e.g., "I didn't go out as often as I would like"). Disagreeing with a negative item resulted in a double negative and thus a positive response (so in the example above: I did go out as often as I would like). This might be confusing for people with language problems. As this issue was linked to the response format it was decided to address it in the pilot study (see below).

4.1.3 Methods: pilot test with people with aphasia

4.1.3.1 Aims

The main aim of the pilot was to determine whether indeed the SA–SD response format was the most challenging and to choose an easier format, if needed. A secondary aim was to see how the respondents would cope with long items and negative items.

4.1.3.2 Design, procedure and participants

An interview-based study was undertaken. Twelve people that attended groups for people with mild or moderate aphasia in the City University Centre for people with aphasia were

approached and asked to take part in the study. They were told that they would be shown some questions and that they would have to choose their responses from set response formats. It was explained that the aim of the study was to see which response format was the best. All approached participants agreed to take part. They were all seen individually by the same SLT who presented them with a set of 5 questions from the SS-QOL, reproduced 6 times with 6 different response formats (see *materials* below). Each one answered the 5 questions with the 6 different response sets presented in random order. They were then asked which one they found harder and why and which one they found easier and why.

4.1.3.3 Materials

4.1.3.3.1 Response formats

Alternative response formats were generated through consultations with the reviewers. The aim was to keep them as close as possible in meaning to the original format. In the context of the second part items (e.g., "I was irritable"), 'agree' and 'disagree' seemed to reflect whether the respondent felt s/he had the problem/feeling or not. 'Strongly' and 'moderately' seemed to reflect the extent to which the respondent felt s/he had the problem/feeling or not. Thus, response formats like faces scales and scales going e.g., from "all the time" to "never" were not included, as they mainly reflected concepts like satisfaction and frequency respectively, rather than agreement.

The alternative response formats were the following:

- *Very true/True/Neither true nor false/False/Very false* (T-F)

This response format was thought to fit well the content of the items. Still, true and false are absolute concepts and the distinction between 'very true' and 'true' is artificial.

- *Yes, a lot/Yes, a little/Neither yes nor no/Not really/No, not at all* with statements (Yes-No) and with questions (Question Yes-no)

This response format was used twice. Once, with the items as they were (statements) e.g., "I was impatient with others". And once with the statements turned into questions e.g., "Did you feel impatient with others?", as 'yes' and 'no' fit better with questions. It was thought to be reasonably close conceptually to what SA–SD conveyed, and was linguistically simple, salient and straightforward. Like the SA-SD format, however, it could be confusing with negative items.

- *That is so right/ I guess so/ I don't know/ I don't think so/ Certainly not* (Comment)

This response format was tried as it was seen as reflecting comments that one might make in a conversation if one agreed or disagreed with statements like "I was discouraged about my future". It was acknowledged, however, that the middle response ('I don't know') was distinct from the middle responses of the other formats.

- ✗✗/✗/✗-✓/✓/✓✓ (✗-✓)

This is a symbolic representation, ranging from two crosses if one has the problem or the feeling a lot, to two ticks if one doesn't have the problem or the feeling at all. This format should be the easiest for people with reading difficulties. It could, however, create some conceptual confusion as it is not clear what it reflects. Due to its symbolic nature it was also seen as more distant conceptually from the original SA-SD response format.

4.1.3.3.2 Items

Five items (md2, md3, fr5, sr4, sr5) were selected from the Mood, Family Roles and Social Roles domains of the second part of the SS-QOL (table 4.1). The first item (md2) was

selected because it was straightforward to ease the respondents into the task. The rest of the items, however, were purposefully selected to see whether they were difficult for the respondents. They were negative (md3, fr5, sr5) or long (including comparative: sr4) items. The 5 items were reproduced 6 times so that there was one set with the SA – SD response format and 5 sets with each one of the alternative formats. Table 4.1 shows the items with the T-F response format.

Table 4-1: Pilot test: SS-QOL items with the T-F response format

DURING THE PAST WEEK:

		Very true	True	Neither true nor false	False	Very false
MD2.	I was discouraged about my future.					
MD3.	I wasn't interested in other people or activities.					
FR5.	I didn't join in activities just for fun with my family.					
SR4.	I did my hobbies and recreation for shorter periods of time than I would like.					
SR5.	I didn't see as many of my friends as I would like.					

Each set was preceded by a page explaining the particular response format and giving an item from the SS-QOL (t3) as an example to practise (table 4.2). An example of the full set of the

six response formats is presented in appendix 4.2 (NB: the order of presentation of the response sets varied between participants).

4.1.4 Results and modifications: pilot test with people with aphasia

All twelve respondents were able to complete the task and give their opinion on the complexity of the different response formats. Table 4.3 summarises their opinions. These results were drawn from a small sample of people with aphasia. They are thus seen as just an indication of what some people may think.

Table 4-2: Pilot test: practice item with the true-false (T-F) response format and explanation of the T-F response format

This is a list of problems or feelings that some people have after their stroke. Possible answers go from:

Very true: If you have the problem a lot, to
Very false: If you don't have the problem at all.

Tick in the box that best says how you felt about each statement **during the past week.**

For example

	Very true	True	Neither true nor false	False	Very false
I had trouble remembering things					

Clearly, there is some variability in the respondents' views. Some of them confirmed our suspicions and found the SA-SD format the hardest. Five out of 12 found the ✗ /✓ format the easiest because of its presumed simplicity, yet two people found it most difficult as its meaning was unclear. The 'yes-no' response format where the items were converted to questions was the only format that nobody found most difficult. Observation of the

respondents and some of the clarifying questions they asked suggested that most of them found the negative items hard. Some of them also had difficulty with the grid format as they had to check back on the top of the page to see what the response options were.

Table 4-3: Pilot test: number of participants and their views on the different response formats

N=12	SA-SD	Yes-No	Question yes-no	Comment	T-F	✗/✓
Easiest	1	0	4 'easy', 'natural', 'like normal conversation' 'straight' 'very clear'	0	2	5 'easy', 'you see it and you know', 'you don't have to read', 'it's quicker'
Most difficult	3 'confusing' 'hard with these (pointing to negative items)'	2	0	3	2	2 'you can't read it, you have to think what it means', 'it's terribly confusing'

To address these issues the following changes were made. First, the second part statements were converted to questions. Converting the items into questions did not affect their meaning. This also resulted in uniformity with the first part. Second, the question yes-no response format was adopted for the second part. It was also slightly modified to: 'definitely yes/mostly yes/neither yes nor no/mostly no/definitely no', following consultation with the questionnaire development experts. The 'yes-no' response format was conceptually similar to the original and highly acceptable to people with aphasia. Third, the negative items were rephrased in the question format to avoid double negatives, e.g. 'I didn't go out as often as I would like' was changed to 'Did you go out less often than you would like?'. Fourth, the grid format was dropped. In each page, the questions were on the left as in the original and the

response categories were presented in the middle of the page, on the right (for example, see p4 appendix 4.3).

In summary, the SS-QOL was modified to make it communicatively accessible to people with moderate or mild receptive aphasia. The modification process included consultation with expert professionals and pilot testing with people with aphasia. This process resulted in an aphasia-adapted version of the SS-QOL designed for interview administration. In terms of presentation, the font was increased to 14, few items were presented per page, transitional questions and practice items were introduced and the grid format was dropped. To increase the measure's content validity with people with aphasia four items were added in the language, thinking, family roles and social roles domains. To reduce the instruments linguistic and reading complexity lead-in questions were used, long sentences were made shorter, key words were presented in bold and some changes were made from American to British English. In the second part of the questionnaire, the items were converted to questions, negation was removed from negative items and the response format was changed from SA-SD to 'definitely yes-definitely no'.

4.2 Pre-test of the aphasia-adapted SS-QOL

Once an instrument is nearly ready to be used, a pre-test by face-to-face interviews with a small number of individuals for whom the instrument is intended is recommended (Fowler, 1993). The main aims of the pre-testing were to evaluate the instruments content validity, accessibility and acceptability with people with aphasia and to identify whether any revisions were necessary before testing its psychometric properties further in a large sample.

4.2.1 Pre-test methods

Face-to-face interviews were carried out with people with aphasia. The interviewer was a specialist in aphasia Speech and Language Therapist. Participants were recruited from the groups of the City University Centre for people with aphasia. Information on the project was given to nineteen people, who were at least one year after their strokes. They were told that they were helping to test a new questionnaire. Two interviews were carried out with each participant, with the second interview being 2-7 days after the first. The first interview comprised a brief language assessment using the Frenchay Aphasia Screening Test (FAST) (Enderby et al., 1987), the administration of the aphasia-adapted SS-QOL and a discussion of the questionnaire based on the following questions:

1. Did this questionnaire cover the effects that stroke and aphasia had on you?
2. Is there anything important to you that was not covered?
3. Would you add any questions to it?
4. Was it a straightforward questionnaire to do?
5. Did you have difficulty understanding any of the questions?
6. Did you find any of the items particularly hard?
7. Is there anything else you want to tell me about this questionnaire?

The interviewer also noted any other comments that were made during the administration of the instrument or during the interview. In the second interview, the SS-QOL was re-administered.

4.2.1.1 Accessibility

The aphasia-adapted SS-QOL was administered twice in order to assess which participants could consistently respond reliably to the questionnaire. This was one way of testing the accessibility of the questionnaire. If the respondents could understand all the questions asked then they were more likely to respond reliably (i.e., give the same or similar responses) on two consecutive administrations of the questionnaire, than if they did not understand some questions and gave chance responses. Intra-class correlations coefficients (ICCs) were

calculated between the two administrations of the instrument and the set criterion was that they should be >.70 for each subdomain and >.90 for the overall mean score.[5]

The FAST was administered in order to see at what level of severity of aphasia the aphasia-adapted SS-QOL was accessible to respondents. Comparing the FAST scores of those who responded reliably on the 2 administrations of the instruments to those who did not, could help determine a cut-off point in the FAST scores above which the instrument was accessible to respondents for self-report.

Questions 4, 5 and 6 were, also, used to assess whether the instrument was accessible to the respondents.

4.2.1.2 Content validity and acceptability

Questions 1, 2 and 3 were used to assess whether the respondents thought that the questionnaire adequately covered the concept under study (content validity). The acceptability of the instrument at this stage was tested, as recommended by Fowler (1993), by the interviewer observing whether certain behaviours occurred. In particular, the interviewer noted whether there were any particular questions that she misread or where the respondents asked for clarification or where the respondents needed prompting to give adequate answers. If these behaviours occur in 15% or more of pre-test interviews, then the questions involved are either highly likely to produce distorted data or distinctively susceptible to interviewer effects (Fowler, 1993). Still, asking for clarification is a common and desirable behaviour for people with aphasia in order to make sure that they have understood the items. For this behaviour the criterion was relaxed to 30%.

[5] For a full discussion of ICCs and acceptable levels see chapter 5: Psychometric evaluation of the SAQOL: methods.

4.2.2 Pre-test results

18 out of the 19 people that were invited to take part agreed to participate.

4.2.2.1 Accessibility

All participants self-reported on the instrument. 17 out of the 18 responded reliably on the 2 administrations of the instrument. For them, the ICC was .97 for the overall mean and for the subdomains it ranged from .73 to .98. These 17 scored 7/15 or more on the receptive scales of the FAST. Observation of the one person who scored 6/15 on the receptive FAST suggested that he did not understand all of the questions. He looked puzzled with some items and yet did not ask for clarifications. Unlike the rest of the respondents he tended to point to a response too quickly after a question had been asked, without spending any time thinking about it. The rest of the respondents also showed variability in the way they responded to the questions, i.e., to some they gave prompt responses, to others they had to think longer, sometimes they asked for repetition of a question or they re-read an item. This respondent however responded to all the questions in a uniform manner. He also required two visits to complete the aphasia-adapted SS-QOL and therefore was not administered the measure twice. The interviewer's clinical opinion was that this respondent had such severe receptive aphasia that he could not give reliable responses to the instrument used.

Responses to questions 4, 5 and 6 supported further the accessibility of the instrument to the 17 people who scored more than 7 on the receptive scales of the FAST. They all agreed that the instrument was straightforward. Seven felt they had some or little difficulty understanding some of the items but were facilitated by the interviewer repeating the items and using gesture. The only items that were identified as hard by more than one participant (two) were md3 "did you have no interest in other people or activities" and fr5 "did you stay out of family activities that were just for fun?".

Two participants commented they had slight difficulty choosing their responses from the 'yes-no' format and in particular choosing between 'mostly yes' and 'mostly no'. Other comments included

'thinking about them (the items) was a bit hard'
'I like the way it's set out'.
'Quite clear'.
'90% of it is brilliant'.

4.2.2.2 Content validity and acceptability

All participants felt that the questionnaire overall covered the effects the stroke had on their lives. Eleven said 'no' when they were asked whether there is anything that is important to them that was not covered. Four participants made the following comments when asked whether they would add more questions.

- *'There should be more on feelings, e.g. frustration, embarrassment and worrying'*
- *'The speech. I need to talk'*
- *'Worrying about the future. How are you going to cope e.g., if your husband dies or if something happens to you?'*
- *'There was nothing on how the partner is affected. My husband gets fed up by not being able to have an intellectual conversation with me... I am terribly boring. He avoids talking to me.'*
- There was nothing on how difficult it is to find out information about things or access services.
- Attitudes to life before and after the stroke change: *'Before I used to worry about money...I had enough money but I always worried about it. Now I have no money ... but I don't worry at all.'*
- *'Let's say that before the stroke I was Anthony and now I am Tony. On the whole, Anthony did everything you know better, but on some things Tony is better'*

In terms of acceptability, there were no items that the interviewer misread. There were, however, two items that the respondents needed prompts in order to answer adequately. These were two items related to writing. Ue1 "how much trouble did you have with writing or typing?" is under the upper extremities domain as it reflects hand function. People with writing difficulties because of their aphasia (e.g., spelling difficulties) tended to report trouble in this area, because of their aphasia rather than hand function difficulty. Similarly, t4 "did

you have to write things down to remember them?" is under the thinking domain. People with aphasia and writing difficulties tended to say 'no' on this item because they could not write and not because they did not need to write things down to remember them.

Occasionally, respondents asked for clarification but no items failed the 30% criterion. The overall acceptability of the measure was supported by the following comments:

> *'it's interesting to have to think about how you feel'*
> *'very good'*
> *'it's good that somebody is doing something about it'*
> *'very interesting and helpful and it brought a lot of things to my head that I thought 'ahh...yes..'*
> *'if you had asked me these questions 4 years ago (just after the stroke) it would have been very difficult'.*

4.2.2.3 Further modifications

The two items that were identified as problematic in terms of accessibility (md3, fr5) were not removed at this stage. As only two participants found them difficult, it was decided to retain them and evaluate them further during the psychometric evaluation of the instrument (chapters 5 and 6).

Two strategies were used to facilitate respondents with the 'yes-no' response format of the second part. First, the administration was modified by spending more time and giving clear instructions during the practice item of this response format. The practice item was "Did you feel hopeless about your future?". After reading this item the interviewer should explain the response format by saying "Definitely yes, if you *really* felt hopeless about your future. Mostly yes, if *often* you felt hopeless about your future. Not sure, if you are not sure how you felt. Mostly no, if *occasionally* you felt hopeless, and definitely no if you did not feel hopeless at all about your future". Thereafter if the participant looks unsure on specific items the

interviewer should give similar prompts. Although this clarification introduces frequency in the response set, agreement is still the main underlying concept.

Second, it was thought that perhaps the 'yes-no' was confusing for some people because in the context of this questionnaire 'yes' which is a positive word suggested that the respondent had a problem or a negative feeling whereas 'no', a negative concept, suggested s/he was fine. To make more salient what the meaning of the responses was in the context of the questions, two anchor points were used at the extremes of the response format (a ✓ with 'definitely no' and a ✗ with 'definitely yes'). Using these anchor points had the secondary advantage of offering extra support to people with reading difficulties. For this reason, and for uniformity throughout the instrument the anchor points were used throughout the aphasia-adapted SS-QOL.

The majority of the respondents thought that the measure had good content validity. No items were added to increase the measure's content validity following the comments that people made, mostly because they were individual comments and no two respondents identified the same areas/items. Moreover, some of the comments (e.g., *'Before I used to worry about money...I had enough money but I always worried about it. Now I have no money ... but I don't worry at all.' 'Let's say that before the stroke I was Anthony and now I am Tony. On the whole, Anthony did everything you know better, but on some things Tony is better'*) identified areas related to the interpretation of HRQL outcomes, but not specific and concrete items or areas that could be added to the measure.

To increase the measure's acceptability the two problematic items on writing were presented differently. In the printed form they remained the same but the interviewer offered clarifications in what was said. Ue1 "how much trouble did you have with writing or typing?"

became "how much trouble did you have with writing or typing, that is using your hand to write or type?". "Did you have to write things down to remember them?" (t4) became "Did you have to write things down to remember them, or ask somebody else to write things down for you to remember?" The further testing of the acceptability of the measure with more rigorous psychometric methods is described in chapter 5.

4.3 Summary

This chapter described the development of an aphasia-adapted version of the SS-QOL. Firstly, the instrument was modified to make it communicatively accessible to people with aphasia and to increase its content validity with this population. This process involved consultation with professionals with related expertise (specialist SLTs, a linguist and a questionnaire development expert) and pilot testing with people with aphasia. The adapted version of the instrument was pre-tested with 18 people with aphasia. The pre-test suggested that the aphasia-adapted SS-QOL was accessible for self-report to people who scored at least 7 out of 15 in the receptive scales of the FAST. A few changes were made to increase further the accessibility and the acceptability of the instrument. The pre-test also suggested that the measure covered adequately the concept under investigation (content validity). The resulting instrument was named Stroke and Aphasia Quality of Life Scale (SAQOL). The presenter's form for the SAQOL is presented in appendix 4.3 and the scoring sheet in appendix 4.4. All prompts and examples that were given only orally were written in italics in the scoring sheet. The next two chapters describe the psychometric evaluation of the SAQOL, i.e., the further testing of its acceptability and the testing of its reliability and construct validity.

5 PSYCHOMETRIC EVALUATION OF THE SAQOL: METHODS

The preliminary psychometric evaluation of the SAQOL took place during the adaptation of the measure for use with people with aphasia and has been described in chapter 4. This included evaluating the accessibility, the content validity and the acceptability of the measure. This chapter will concentrate on what methods were used to test further the acceptability of the measure and to test its reliability and validity.

5.1 Acceptability

There are a number of factors that can be used as indicators of data quality and acceptability of a questionnaire to respondents. They include response rates, percentage of missing data and the distribution of scores across response categories (Mchorney et al., 1994). These were calculated to test the acceptability of the SAQOL. Missing data for each item should be below 10% (Fitzpatrick et al., 1998). Floor and ceiling effects were assessed by calculating the frequency of respondents endorsing the bottom and the top of the scale. The frequency of endorsement is simply the proportion of people who give each response alternative to an item (Streiner & Norman, 1995). For the items of a questionnaire to discriminate well between respondents, responses should be distributed across the response categories (Streiner & Norman, 1995). The WHO criterion of aggregate endorsement frequencies (AEF) <10% was followed (WHOQOL Group, 1998). This means that items should not have <10% of responses on two adjacent response alternatives. Maximum endorsement frequencies (MEF) (percentage of respondents endorsing one response alternative to an

item) were also calculated. MEF should be <80% (Streiner & Norman, 1995). It is generally accepted that skewness values should be in the range of –1 to +1. The data in this study, however, are derived from people who are long-term post stroke and thus some negative skewness should be expected and acceptable. The criterion that was set stated that the percentage of variables with skewness values of greater than an absolute value of 1 should not exceed 25%.

5.2 Reliability

Reliability is concerned with the stability and consistency of a measure. It refers to its homogeneity and the extent to which it is free from random error. If a measure is consistently yielding the same results time after time then it is free of random error. Essentially, reliability assessment is a matter of checking for such consistency (Singleton & Straits, 1999). There are four types of reliability: internal consistency, test-retest reliability, inter-rater reliability and parallel forms reliability. The two latter do not apply to the SAQOL. Inter-rater reliability refers to the level of agreement between two or more independent raters or observers of an individual. It is therefore not relevant to a self-report questionnaire where no raters are involved. Parallel forms reliability refers to the level of agreement between two or more versions of an instrument that are designed to measure the same concept using different items. Parallel forms reliability is not commonly assessed in HRQL measures and there were no alternative versions of the SAQOL to be used. The internal consistency and the test-retest reliability of the SAQOL were tested using the following methods.

5.2.1 Internal consistency

Internal consistency involves testing for homogeneity. It refers to the extent to which items in a scale measure the same concept, and the extent to which the items relating to a particular domain in a scale tap only this domain and no other.

Cronbach's alpha was used as a measure of the internal consistency of the SAQOL whole scale and subdomains. Cronbach's alpha is based on the average correlation among the items and the number of items in each scale. It should be above .70 for group comparisons and above .90 for individual assessments (Nunnally & Bernstein, 1994; Hays et al., 1993). A coefficient alpha of .70 was used as the minimally acceptable level for internal consistency reliability of the SAQOL scale and of its subdomains. Internal consistency is also evidenced by moderately high item-total correlations, which indicate that the items can be combined into a single scale. A low item-total correlation indicates that an item may be measuring something different from the other items in a scale. Item-total correlations should exceed .30 (Nunnally & Bernstein, 1994).

5.2.2 *Test-retest reliability*

The test-retest procedure involves checking the same individuals on the scale on two separate occasions and correlating the two sets of scores. The correlation in test-retest reliability testing tends to be high, provided no real change has occurred (due for example to time or intervention). However, determining what is an acceptable level of reliability is not a simple matter as it depends on the nature of the variable, the situation and the intended use of the measure (Rosenthal & Rosnow, 1991; Streiner & Norman, 1995). Rosenthal & Rosnow (1991) indicate that for clinical testing reliability coefficients of .85 or higher are acceptable, whereas in experimental research lower coefficients may be accepted as satisfactory. Streiner & Norman (1995) see a reliability of .75 as a minimal requirement for a useful instrument.

Test-retest reliability is usually expressed in Pearson's product correlation coefficients or intra-class correlation coefficients (ICCs). The ICC is the proportion of total variability accounted for by the variability among individuals. ICCs are sensitive to systematic changes in the mean level of responding (e.g., every individual's score decreasing by a constant)

whereas Pearson coefficients are not. ICCs between two administrations of the SAQOL were therefore calculated, to determine its test-retest reliability.

The testing and the re-testing should not be so close to one another that the participants can remember their responses and yet they should not be so distant from one another that a true difference has occurred in the measured variable. The test-retest period was 2 to 14 days as recommended by Streiner & Norman (1995).

5.3 Validity

Validity refers to the extent to which an instrument measures what it purports to measure. The validation of a measure includes assessing its face validity, content validity, criterion related validity and construct validity. The face and content validation of the SAQOL were undertaken during the adaptation of the measure for use with people with aphasia and they have been covered in Chapter 4. Here, the methods used to further assess the validity of the SAQOL and the rationale behind these methods are described.

Validation of the SAQOL included validation of the whole scale as a HRQL measure and validation of its subdomains.

5.3.1 Criterion related validity

Criterion related validity applies to measures that have been developed for some practical purpose, and where the investigator is interested in the usefulness of the measure as an indicator of a specific trait or behaviour (Singleton & Straits, 1999). Criterion related validity includes concurrent and predictive validity. In testing for criterion validity the most sensitive and meaningful criterion in the past, present or future should be selected (Rosenthal & Rosnow, 1991). In this case, this criterion may be either an established measure of HRQL that is seen as a gold standard or a professional judgment. It is well known, however, that

absolute gold standard measures do not exist for HRQL (e.g., Hays et al., 1993; Williams et al., 1999b). Similarly, given the multifaceted nature of the concept, no professionals are specifically trained to assess HRQL. HRQL measures are thus commonly evaluated with construct validity rather than criterion related validity.

5.3.2 Construct validity

Singleton & Straits (1999) point out that construct validity emphasizes the meaning of the responses to the instrument. Is the instrument measuring the underlying construct or could it be measuring something else? Construct validity is evaluated with within-scale analyses and comparisons with external criteria. Within scale analyses include assessing the scale's internal consistency, assessing the intercorrelations between its subdomains and factor analysis. Comparisons with external criteria represent an accumulation of evidence which may include correlations with measures measuring the same construct (convergent validity), correlations with measures measuring similar constructs, differences with measures measuring different constructs (discriminant validity) and differences among groups that should differ on the measure of the construct (known groups approach).

The whole scale construct validation of the SAQOL comprised within scale analyses and comparisons with external criteria. The subdomains validation of the measure consisted of comparisons with external criteria.

5.3.2.1 Whole scale validation: within scale analyses

The following within scale analyses were used to evaluate the construct validity of the SAQOL: internal consistency, intercorrelations between subdomains and factor analysis.

5.3.2.1.1 Internal consistency

Internal consistency provides evidence not only for the reliability but also for the construct validity of a scale. If a scale is measuring a single underlying construct then it should be homogenous i.e., have good internal consistency. The criterion of a Cronbach's alpha coefficient of more or equal to .70 was used, as recommended by Nunnally & Bernstein (1994), to evaluate the homogeneity of the SAQOL. Internal consistency is also evidenced by moderately high item-total correlations, which indicate that the items measure aspects of the same underlying construct and that they can be combined into a single scale. Item-total correlations should exceed .30 (Nunnally & Bernstein, 1994).

5.3.2.1.2 Intercorrelations between subdomains

If the SAQOL subdomains are part of a single underlying construct then they should be moderately correlated with the total mean score less the subdomain (corrected mean). A criterion was set for these moderate correlations of .30 to .80. Moderately high correlations (.50-.80) were expected between subdomains measuring physical abilities (e.g., Self Care, Mobility, Upper extremities). Moderately high correlations (.50-.80) were also expected between subdomains measuring psychosocial aspects (e.g., Mood and Personality). Conversely, subdomains measuring physical abilities and subdomains measuring psychosocial aspects should have lower correlations with one another (<.50). So, for example, self-care, mobility and upper extremities function should have higher correlations with each other than they do with mood, personality and family or social roles.

5.3.2.1.3 Factor analysis

Principal components analysis (PCA) and factor analysis (FA) are statistical techniques applied to a single set of variables[6] where the researcher is interested in discovering which variables form coherent subsets that are relatively independent of one another (Tabachnick & Fidell, 2001). These techniques can be used to confirm that items are correctly grouped together, that items in the same subdomain measure the same construct, that items in different subdomains measure different constructs, and to identify items that contribute little to their intended subdomain.

The choice between PCA and FA largely depends on the goals of the research. PCA merely decomposes the original variables into a set of linear variates whereas FA is used to test hypotheses, i.e., it is used when researchers believe there is a smaller set of 'factors' that cause or in some way influence the observed variables (Dancy & Reidy, 1999). Tabachnick & Fidell (2001) suggest using PCA when one wants an empirical summary of the data set and using FA when one is interested in a theoretical solution uncontaminated by unique and error variability. They still recommend using PCA as the first step in FA to check the factorability of the correlation matrices.

Both PCA and FA were used with the SAQOL. Before describing why and how these methods were used with the SAQOL data, the criteria used during these analyses are presented. These include pre-analyses tests, specifying the methods of extraction and determining acceptable factor loadings and number of items per factor.

To perform FA, the correlation matrix of the data needs to meet certain psychometric requirements. These minimally involve the Keiser-Meyer-Olkin (KMO) test of sampling

[6] In PCA and FA the term variable is commonly used to indicate an item in a scale.

adequacy and the Bartlett's test of sphericity (Ferguson & Cox, 1993). The KMO test of sampling adequacy indicates whether the associations between the variables in the correlation matrix can be accounted for by a smaller set of factors. It should be at least .5. The Bartlett's test of sphericity tests the null hypothesis that no relationships exist between any of the variables. A significant test statistic (based on chi-square) indicates that there are discoverable relationships in the data (Ferguson & Cox, 1993).

With regard to the method of factor extraction, the most common method used is to extract as many factors as there are eigenvalues greater than one. Ferguson & Cox (1993) point out that this method is only applicable when the initial communalities are all at unity and that in general it extracts more factors than generally required. Another method commonly used is the scree test. This involves plotting the eigenvalues for each factor against the number of factors (scree plot) and looking for a break in the scree plot or for the point where a line drawn through the points changes slope (Tabachnick & Fidell, 2001). Both these methods were used with the SAQOL data, as described in the results in chapter 6.

With regard to the acceptable magnitude of the loadings of variables on factors the criterion of loadings equal or greater than .40 was followed, which ensures good factor saturation (Ferguson & Cox, 1993). Still, some items may load on 2 or more factors with loadings $\geq .40$ (crossload), making difficult to judge which factor they belong to. The Ferguson & Cox (1993) criterion was followed, which states that if the difference between the crossloadings is $\geq .20$, the item is assumed to load on the factor where it has the highest loading. If however the difference is less than .20 then removal of the item is warranted as it is difficult to say which factor the item represents.

As far as number of variables per factor is concerned, Tabachnick & Fidell (2001) point out that "interpretation of factors defined by only one or two variables is hazardous under even the most exploratory factor analysis". Thus, at least 3 variables per factor is an essential criterion to ensure some factor stability and for an adequate interpretation of factors.

As has been indicated, both PCA and FA were used with the SAQOL. Initially, unrotated PCA[7] was performed to check that all variables loaded on the first component (i.e., that there was a single underlying construct that they all measured) and to check the factorability of the data set. FA (the Principal Axis Factoring (PAF) method) was then used to identify the model that best described how the variables grouped in underlying factors. Orthogonal varimax rotation was used to improve the interpretability and scientific utility of the solution (Tabachnick & Fidell, 2001).

Two parallel strategies were used in the FA of the scale. In strategy 1, the FA commenced with all the items of the SAQOL in the analysis. In strategy 2, item reduction was first performed and then FA was carried out on the item reduced version of the SAQOL.

5.3.2.1.3.1 Strategy 1: FA commencing with all item SAQOL

The rationale behind this strategy was to check whether the original conceptual model of the SS-QOL held up in this sample, i.e., whether indeed the variables grouped into 12 subdomains. A top-down and a bottom-up approach were followed. In the top-down approach, a PCA and a PAF were carried out within each subdomain to check that all the items measured one underlying construct and to identify those contributing little to the underlying subdomain construct. A second order PAF was then undertaken using the 12

[7] All PCA analyses performed were unrotated. Wherever PCA is mentioned hereafter it is unrotated PCA.

subdomains' means, to test whether the 12-factor model would be derived and if not to see whether the subdomains grouped into conceptually sound factors.

In the bottom-up approach, all items were entered into a PCA (to confirm whether there was a single underlying construct) and then into a PAF to derive a sound factor model, according to the criteria set above. If the original 12-subdomain model did not hold up, further PAF analyses would be performed in order to derive the factor model that best described the data. Following this approach could lead to a different version of the SAQOL.

5.3.2.1.3.2 Strategy 2: FA commencing with item reduced SAQOL

In strategy 2, items that did not meet set criteria in pre-analyses checks were removed, and PCA and PAF were carried out in the resulting item reduced version of the SAQOL. The rationale behind this strategy was to start the FA with a matrix that was derived from properly scaled variables (Ferguson & Cox, 1993). Standard psychometric methods for item reduction were applied including: missing data, MEF, AEF, item redundancy and item-total correlations (Streiner & Norman, 1995; WHOQOL Group, 1998; Nunnally & Bernstein, 1994). The following criteria were used. For missing data, items with greater than 5% missing data were removed. Items with MEF >80% and items with AEF <10% were removed. Items with item-total correlations <.30 were removed. In terms of items' redundancy, pairs of items with correlations greater than .75 were identified. If both items came from the same subdomain then the less specific item was eliminated. If items came from different subdomains, then the item from the subdomain with more items was eliminated. Following this strategy would result in a different version of the SAQOL, which would be shorter reducing respondent burden.

In summary, PCA and FA (PAF) were performed on the SAQOL to assess whether the original 12-subdomain model of the SS-QOL held up in this sample of people with chronic aphasia and if not to derive the factor model that best described this data. Different strategies were used in this process, which would lead to different versions of the SAQOL emerging. The psychometric properties of these new versions would need to be assessed in ways similar to the ones described in this chapter for the SAQOL. These versions would need to be compared to identify the one that was psychometrically and conceptually more sound.

The next section describes how the whole SAQOL scale was validated against external criteria.

5.3.2.2 Whole scale validation: comparisons with external criteria

Comparisons with external measures were used to test the convergent validity of the SAQOL, its correlations with measures measuring similar constructs and its discriminant validity.

5.3.2.2.1 Convergent validity

The principal underlying convergent validity is that different measures of the same concept should be correlated. Convergent validity is one of the most convincing pieces of evidence of construct validity (Singleton & Straits, 1999).

The lack of scales measuring HRQL in people with aphasia following stroke, which necessitated this research, made the testing of the convergent validity of the SAQOL a challenge. The developers used the Short-form 36 (SF-36) (Ware et al., 1993) in the original validation process (Williams et al., 1999a), which is a well regarded generic measure of QOL. The SF-36 could not be used with people with aphasia due to its linguistic complexity.

Another way of measuring people's HRQL is to ask them to rate it. In the original validation process of the SS-QOL respondents were asked to rate their overall quality of life as the same, a little worse or a lot worse than before the stroke. The authors then compared the patients overall SS-QOL scores with their ratings of overall HRQL. The rationale was that if the SS-QOL is a good indicator of HRQL, then people who rate their overall HRQL as worse than before the stroke should have lower SS-QOL scores whereas those who rate it as the same should have higher SAQOL scores. The use of a single question to measure a complex subjective concept like HRQL has serious implications on the measurement's validity and reliability. Still, in the absence of other suitable measures it was decided to replicate this approach and view the results as just one indication of convergent validity. ANOVA was used to compare the SAQOL scores of people who rated their quality of life as a lot worse, a little worse and the same/better than before the stroke.

5.3.2.2.2 Correlations with related variables

If a measure is valid, it should correlate with measures of other theoretically related variables or measures measuring similar constructs. In stroke outcomes research, depression has been repeatedly associated with reduced HRQL (Ahlsio et al., 1984; Niemi et al., 1988; King, 1996; Duncan et al., 1997; Jonkman et al., 1998; Neau et al., 1998; Clarke et al., 1999; Lofgren et al., 1999). Other factors reflecting affected mood have also been associated with reduced QOL, like anxiety (Ahlsio et al., 1984; Astrom et al., 1992; Astrom, Asplund & Astrom 1992) overall distress (Ebrahim et al., 1986) and loneliness and sleep problems (Wyller et al., 1998). It was thus anticipated that the SAQOL scores would correlate moderately with the GHQ-12, which is a measure of distress commonly used as a screening tool for psychiatric disorders.

Stroke outcome studies have also reported a strong association between reduced ADL/functional status and diminished HRQL (Ahlsio et al., 1984; Ebrahim et al., 1986; Niemi et al., 1988; Viitanen et al., 1988; Astrom et al., 1992; Astrom, Asplund & Astrom 1992; Kwa et al., 1996; King, 1996; Wilkinson et al., 1997; Jonkman et al., 1998; Neau et al., 1998; Clarke et al., 1999; Lofgren et al., 1999). Moderate correlations were expected between the SAQOL scores and the Frenchay Activities Index (FAI), which is a measure of functional activities after stroke (including ADL, hobbies and recreation and going out). Moderate correlations were also expected with the American Speech and Hearing Association Functional Assessment of Communication Skills (ASHA-FACS). This was because, although the ASHA-FACS is a measure of communication skills, it concentrates on functions/activities in which people use their communication skills (e.g., participating in conversations, using the phone, making money transactions).

5.3.2.2.3 Discriminant validity

A valid operational definition should separate the concept being measured from other concepts from which it is intended to differ (Singleton & Straits, 1999) or from dissimilar, unrelated concepts (Streiner & Norman, 1995).

A HRQL measure should be different from a language impairment measure, a cognitive decline measure and a social support measure. Although these 3 variables may contribute to HRQL after stroke, they are still distinct concepts and any correlations with the SAQOL should be low. It was thus predicted that the Frenchay Aphasia Screening Test (FAST), the Raven Coloured Progressive Matrices (RCPM) and the Social Support Scale (SSS) would have low correlations with the average SAQOL scores.

Table 5.1 summarises the whole scale construct validation of the SAQOL against external criteria.

Table 5-1: Construct validation of the whole SAQOL scale against external criteria.

Validity	SAQOL mean score
Convergent	Respondents' rating of their HRQL
Correlated Measures	GHQ-12, FAI, ASHA-FACS
Discriminant	FAST, RCPM, SSS

In summary, within scale analyses and comparisons with external criteria were used to test the construct validity of the overall SAQOL. Within scale analyses comprised internal consistency (Cronbach's alpha ≥.70, item-total correlations ≥.30), intercorrelations between subdomains (ranging from .50-.80 for closely related ones) and the corrected mean (.30-.80), and factor analysis. To test the convergent validity of the scale, ANOVA was used to compare the SAQOL scores of three groups of people who rated their quality of life as a lot worse, a little worse and the same/better than before the stroke. A number of comparisons with other measures were undertaken to test the SAQOL correlated measures and discriminant validity. No criteria were set for the absolute size of these correlations, but the correlations between the SAQOL and measures against which its discriminant validity was tested should be lower than the ones with measures against which its correlated measures validity was tested. The same criteria would be used to assess the construct validity of alternative versions of the instrument emerging from the FA.

5.3.2.3 Validation of the subdomains

According to the original conceptual model of the SS-QOL, the SAQOL has 12 subdomains: self-care, mobility, upper extremities function, work, vision, language, thinking, personality, energy, family roles and social roles. Aphasia has a considerable impact on psychological and social aspects of life (e.g., LeDorze & Brassard, 1995; Hemsley & Code, 1996; Sarno, 1997; Parr et al., 1997). It was, thus, of greater interest to explore the validity of the psychosocial domains of the SAQOL with people with aphasia. Four of the SAQOL domains reflect physical disability (i.e., mobility, self-care, vision, upper extremity function). Aphasia per se does not affect these domains. Validation of these physical domains would require administering to the subjects a number of extra external measures sensitive enough to e.g., differentiate between self-care and upper extremities functions. This would considerably increase respondent burden. It was therefore decided to concentrate the subdomains' validation on the non-physical domains of the SAQOL.

5.3.2.3.1 Comparisons with external criteria

The construct validation of the language, thinking, personality, mood, energy, family roles, social roles and work domains of the SAQOL included at least convergent and discriminant or correlated measures and discriminant validation for each, as described below.

5.3.2.3.1.1 Convergent validity

The convergent validity of the language subdomain was tested by correlating the language subdomain mean score with two measures: the FAST and the ASHA-FACS. The FAST is a measure of language impairment covering all domains of language that may be affected by aphasia, i.e., comprehension, expression, reading and writing. The ASHA-FACS is a measure of communication disability. These two instruments are established measures used in the

assessment of aphasia and both have been standardised with people with aphasia. The language domain of the SAQOL should correlate highly with these two measures.

The thinking domain of the SAQOL was tested against the RCPM. The RCPM assess a person's ability to reason by analogy, which underlines the ability to draw inferences. This ability appears to be one of the earliest to decline as a result of organic dysfunction (Raven et al., 1995). The RCPM is a valid instrument for the assessment of cognition in individuals with language impairments as its stimuli are pictorial symbols, no verbal responses are required and only minimal verbal instruction is necessary. It has been used to explore cognitive decline in brain damage and aphasia (e.g., Villardita, 1985). If the thinking domain of the SAQOL taps on cognitive processes then it should correlate highly with the RCPM.

The convergent validity of the mood domain was assessed against the GHQ-12. The GHQ is a psychometrically sound measure of distress. It has been used in stroke studies and it was shown to have superior properties to other measures, in identifying anxiety and depression after stroke (studies reviewed in chapter 3, under 'measures'). The 12-item version was chosen because it is brief thus reducing respondent burden; and it excludes all somatic items that could result in false-positives in people with chronic disabilities. The GHQ-12 was expected to have high correlations with the mood domain of the SAQOL.

The convergent validity of the work domain was assessed against the FAI. The FAI is a measure of general activities of stroke patients. It includes in and outside the home activities, social and leisure activities and an item on work. Similarly, the work domain of the SAQOL covers 'doing daily work around the house', 'finishing jobs', and 'doing the work you used to do'. High correlations were anticipated between the FAI and the work subdomain of the SAQOL.

5.3.2.3.1.2 Correlations with related variables

A number of measures of related variables were used to assess the validity of the SAQOL subdomains.

Social support is a theoretically related variable to social and family roles. Social roles and social support seem to reflect different aspects of the concept of social health. McDowell & Newell (1996) point out that different measures have been developed measuring social support, social adjustment or the ability to perform normal roles in society, and they see these as contrasting ways of defining social health. In the present investigation, the correlations of the family and social roles domains of the SAQOL with the Social Support Scale (SSS) were explored. The SSS assesses the perceived availability of four types of support (tangible, emotional/informational, social companionship and affectionate support). It was hypothesized that if a person performs his/her family and social roles adequately then s/he is more likely to receive social support from his/her family and friends. Thus, moderate correlations were expected between the family and social roles domains of the SAQOL and the SSS.

Studies on stroke and social support have shown that, following stroke, people experience a reduction in social support and social interaction (e.g., Labi et al., 1980; Angeleri et al., 1993; Neau et al., 1998; Trigg et al, 1999; Friedland & McColl, 1987; Fukunishi et al., 1997). This reduction in social support has serious implications for the well-being of stroke survivors. Some studies have reported an association between poor social support and psychosocial dysfunction including depression after stroke (Friedland & McColl, 1987; Fukunishi et al., 1997; and Morris et al., 1991 although the association reported in this study was not significant). Conversely, a number of studies have provided evidence that increased social support after stroke has a beneficial effect on outcome. Wyller et al. (1998) reported that a

firm social network was associated with increased subjective well-being and Friedland & McColl (1987) found that community support had a strong protective effect against psychosocial dysfunction after stroke. Taking all this into account at least moderate correlations would be expected between the mood domain of the SAQOL and the SSS.

The personality domain of the SAQOL includes three items: a broad item with unclear underlying concept ("my personality has changed"), an item on irritability and an item on impatience. No personality scale that primarily reflects these concepts was found and thus the convergent validity of the personality domain has not been tested. These concepts however seem to be conceptually close to mood items. The validity of the personality domain was explored by correlating it with the GHQ-12, expecting moderately high correlations. The GHQ-12 should correlate higher with the mood domain of the SAQOL than with the personality domain, if indeed the personality domain is conceptually distinct from the mood domain.

In addition, if indeed the personality domain of the SAQOL reflects personality traits then it would be expected to correlate at least moderately with the SSS. Although the relationship between personality factors and perceptions of social support is complex, there is substantial evidence that amount of social support and types of perceived social support may, to some extent, be explained in terms of personality (Sarason & Sarason 1982; Connell & D' Augelli 1990; Fukunishi & Rahe, 1995; Kitamura et al., 1999; Zellars & Perrewe, 2001).

The GHQ-12 should also have moderate correlations with the thinking, social and family roles domains of the SAQOL. This is based on research, which has indicated that depression after stroke is related –among other things- to functional disability (e.g., Ebrahim et al, 1987; Burvill et al., 1997; Herrmann et al., 1998; Kotila et al., 1998; Dennis et al., 2000); social

inactivity (e.g., Angeleri et al., 1993; Anderson et al., 1995); and cognitive impairment (e.g., Anderson et al., 1995; Kauhanen et al., 1999).

Moderate correlations were also anticipated between the GHQ-12 and the energy subdomain of the SAQOL. Other HRQL measures that include an energy subdomain were reviewed and it was found that in the SIP the 'alertness behaviour' scale was under the psychosocial categories. In the SF-36 the 'vitality' scale measured both physical and mental health components (McHorney et al., 1993). Still, in stroke survivors the vitality scale of the SF-36 correlated highly with the psychological functioning domain of the EuroQol (Dorman et al., 1999). In the NHP the 'energy level' section was under the 'non-physical' categories. These 'non-physical' categories correlated highly with the GHQ in stroke survivors (Ebrahim et al., 1986). It was, thus, anticipated that in this sample the energy subdomain would correlate moderately with the GHQ-12.

The validity of the family and social roles domains of the SAQOL was also tested against the FAI. As has been already indicated the FAI is a measure of activities after stroke, which includes in and outside the home activities and social and leisure activities. The family roles subdomain of the SAQOL includes items on participation in family activities and effect of physical health on family life. The social roles subdomain includes items on social and leisure activities e.g., going out, doing hobbies, seeing friends. Moderate correlations were expected between the FAI and the family and social roles subdomains.

Finally, the work domain of the SAQOL should correlate with the FAST and the GHQ-12. Research on the predictors of return to work after stroke has indicated that no aphasia (Black-Schaffer et al., 1990), no language understanding problems (Angeleri et al., 1993) and no major depression (Neau et al., 1998) were good predictors for return to work. The work

domain was, thus, anticipated to correlate at least moderately with severity of aphasia as measured by the FAST and severity of psychological distress as measured by the GHQ-12.

5.3.2.3.1.3 Discriminant validity

Each domain of the SAQOL is aimed at reflecting one aspect of HRQL. However as they reflect different aspects of the same underlying concept it is reasonable to expect that they will be inter-correlated. Still, each domain should have a higher correlation with an external measure that measures the same or a similar underlying concept than with external measures that measure other aspects of HRQL. For example, the language domain of the SAQOL should have higher correlations with the FAST and the ASHA-FACS than with the RCPM, the FAI and the SSS. Similarly the thinking domain should have higher correlations with the RCPM than with the FAST, the FAI and the SSS. The discriminant validity of the mood and personality domains was tested against the FAST, the ASHA-FACS, the RCPM and the FAI, with which they should have low correlations. The energy subdomain was anticipated to have low correlations with the FAST, the ASHA-FACS, the RCPM and the SSS. It was also expected that the family and social roles domains would correlate lower with measures of language ability and cognition (i.e., the FAST, the ASHA-FACS and the RCPM) than with measures of more related concepts such as social support, mood and activities (the SSS, GHQ-12 and FAI respectively). Finally, the work domain should correlate higher with measures of activities, mood and language abilities than with the social support measure.

In summary, the construct validation of the SAQOL subdomains involved comparisons with external measures to test their convergent, correlated measures and discriminant validity (table 5.2). No criteria were set for the absolute size of these correlations, but the following direction should be observed. Each subdomain should have higher correlations with the

measures against which its convergent validity was tested than with the measures against which its correlated measures validity was tested. It should also have higher correlations with the measures against which its correlated measures validity was tested than with the measures against which its discriminant validity was tested. Similar techniques would be used to test the subdomains validity of any alternative versions of the SAQOL emerging from the FA.

Table 5-2: Construct validation of the SAQOL subdomains against external criteria.

Subdomains	Validity		
	Convergent	**Correlated measures**	**Discriminant**
Language	FAST, ASHA-FACS		RCPM, FAI, SSS
Thinking	RCPM	GHQ-12	FAST, ASHA-FACS, FAI, SSS
Personality		GHQ-12, SSS	FAST, ASHA-FACS, RCPM, FAI
Energy		GHQ-12	FAST, ASHA-FACS, RCPM, SSS
Mood	GHQ-12	SSS	FAST, ASHA-FACS, RCPM, FAI
Family Roles		FAI, SSS, GHQ-12	FAST, ASHA-FACS, RCPM

Subdomains	Validity		
	Convergent	**Correlated measures**	**Discriminant**
Social Roles		FAI, SSS, GHQ-12	FAST, ASHA-FACS, RCPM
Work	FAI	FAST, GHQ-12	SSS

5.4 Summary

This chapter described the methods used in the psychometric evaluation of the SAQOL. The acceptability of the measure was further tested and its reliability and construct validity were evaluated. Response rates, percentage of missing data and the distribution of scores across response categories were calculated to test the acceptability of the SAQOL. Reliability testing comprised the assessment of the internal consistency and test-retest reliability of the scale. Within scale analyses and comparisons with external measures were used in the construct validation of the scale. The within scale analyses were the assessment of the internal consistency of the scale, the inspection of the intercorrelations between the scale's subdomains and the subdomains and the scale's corrected mean, and factor analysis. FA was used to test whether the original 12-subdomain conceptual model of the SS-QOL held up in the SAQOL data and if not to derive the best factor model to describe the data. This process was anticipated to result in alternative versions of the SAQOL that would need to be assessed for their psychometric properties in the same way as the SAQOL. Validation of the SAQOL subdomains consisted of comparisons with external measures to test their convergent validity, their correlations with related variables and their discriminant validity.

The results of the psychometric evaluation of the SAQOL (and its versions) are presented in the next chapter.

6 PSYCHOMETRIC EVALUATION OF THE SAQOL: RESULTS

6.1 Respondents

6.1.1 Recruitment

One hundred and sixteen eligible participants were identified during the recruitment period and were asked to take part in the study. Ninety-five people (82%) agreed to take part in the study. No further information is available on the 21 people who did not take part as they did not give their consent for their records to be reviewed. Table 6.1 summarises the numbers of respondents from the different sites and the reasons why some people were unable to participate.

Table 6-1: Response rates in the psychometric evaluation of the SAQOL.

	Southwark	Lambeth	Queen Mary's	Connect	Total (%)
Eligible	22	29	36	29	116 (100%)
Took part	17	21	31	26	95 (82%)
Not interested	4	3	1	0	7(6%)
Unable to consent	0	2	1	0	3 (2.6%)
Unable due to health problems	0	2	3	0	5 (4.3%)

	Southwark	Lambeth	Queen Mary's	Connect	Total (%)
Unable to establish contact	1	1	0	0	2 (1.7%)
Unable due to holiday plans	0	0	0	3	3 (2.6%)

Of the ninety-five people who took part to the study, 12 had such severe language problems (< 7/15 on the receptive domains of the FAST) that they were unable to self-report on the questionnaires that were used. For those participants proxy respondents were used (spouse/partner or other close relative or friend or main care giver). All results reported here are from the remaining 83 subjects.

6.1.2 Respondents' characteristics

Table 6.2 details the respondents' characteristics. The majority were male (62.7%) and they ranged in age from 21 to 92 (mean 61.67±15.47). About 43% were over 66 years old and 15.7% were between 21 and 45. The majority of the sample was white (78.3%) and married/had a partner (62.6%). More details on the respondents are given in chapter 7, where their characteristics are discussed in relation to their HRQL outcomes.

Table 6-2: Respondents' characteristics

Characteristics	N=83	Percent
Gender		
Female	31	37.3%
Male	52	62.7%
Age		
Mean (SD)	61.67 (15.47)	
Range	21-92	
21-45	13	15.7%

Characteristics	N=83	Percent
46-65	34	41%
66+	36	43.4%
Stroke type		
Ischaemic	36	43.4%
Haemorrhagic	16	19.3%
Unknown	31	37.3%
Time post stroke		
Mean in years (SD)	3.5 (3.09)	
Range	1y 1m-20y 10m	
1-2 years post onset	26	31.3%
2.1-4 years post onset	31	37.3%
4.1+ years post onset	26	31.3%
Comorbidity		
None or one comorbid condition	34	41%
Two or more comorbid conditions	49	59%
Ethnic group		
Asian	7	8.4%
Black	11	13.3%
White	65	78.3%
Marital status		
Married	42	50.6%
Has partner	10	12%
Single	14	16.9%
Divorced or spouse died	17	20.5%
Socioeconomic status (revised SEC)		
Professionals/senior managers	23	27.7%
Ass. Professional/ junior managers	6	7.2%
Other admin. and clerical workers	13	15.7%
Own account non-professional	5	6%
Supervisors, technicians and related workers	11	13.3%
Intermediate workers	9	10.8%
Other workers	12	14.5%
Never worked/other inactive	4	4.8%
Employment status		
Retired before the stroke	31	37.3%
Inactive because of the stroke	47	56.6%
Some p/t or voluntary work	3	3.6%
Students	2	2.4%

6.2 Acceptability

Before presenting any results on the SAQOL data, a list of the SAQOL items has been reproduced here (table 6.3) for ease of reference.

6.2.1 Response rates

All participants able to self-report (83) were administered the SAQOL in an interview format and all of them completed the whole scale.

Table 6-3: List of SAQOL items

Part 1

SC1.	How much trouble did you have preparing food?
SC2.	How much trouble did you have eating, for example, cutting food or swallowing?
SC4.	How much trouble did you have getting dressed?
SC5.	How much trouble did you have taking a bath or shower?
SC8.	How much trouble did you have using the toilet?
M1.	How much trouble did you have walking?
M4.	How much trouble did you have keeping your balance when bending over or reaching?
M6.	How much trouble did you have climbing stairs?
M7.	How much trouble did you have walking without stopping to rest or using a wheelchair without stopping to rest?
M8.	How much trouble did you have standing?
M9.	How much trouble did you have getting out of a chair?
W1.	How much trouble did you have doing daily work around the house?
W2.	How much trouble did you have finishing jobs that you started?
W3.	How much trouble did you have doing the work you used to do?
UE1.	How much trouble did you have writing or typing, i.e. using your hand to write or type?
UE2.	How much trouble did you have putting on socks?
UE4.	How much trouble did you have doing buttons?
UE5.	How much trouble did you have doing a zip?
UE6.	How much trouble did you have opening a jar?
V1.	How much trouble did you have seeing the TV well enough to enjoy it?
V2.	How much trouble did you have seeing things you wanted to reach?
V3.	How much trouble did you have seeing things off to one side?
L2.	How much trouble did you have speaking?
L3	How much trouble did you have speaking clearly enough to use the phone?
L5.	How much trouble did you have getting other people to understand you?
L6.	How much trouble did you have finding the word you wanted to say?
L7.	How much trouble did you have getting other people to understand you even when you repeated yourself?
L4.	How much trouble did you have understanding what other people say?

Part 2

T2.	Did you find it hard to concentrate?
T3.	Did you find it hard to remember things?
T4.	Did you have to write things down to remember them, (or ask somebody else to write things down for you to remember)?
T5.	Did you find it hard to make decisions?
P1.	Did you feel irritable?
P2.	Did you feel impatient with others?
P3.	Did you feel that your personality has changed?
MD2.	Did you feel discouraged about your future?
MD3.	Did you have no interest in other people or activities?
MD6.	Did you feel withdrawn from other people?
MD7.	Did you have little confidence in yourself?
MD8.	Did you have no interest in food?
E2.	Did you feel tired most of the time?
E3.	Did you have to stop and rest often during the day?
E4.	Did you feel too tired to do what you wanted to do?
FR5.	Did you stay out of family activities that were just for fun?
FR7.	Did you feel that you were a burden to your family?
FR8.	Did you feel that your physical condition interfered with your family life?
FR9.	Did you feel that your language problems interfered with your family life?
SR1.	Did you go out less often than you would like?
SR4.	Did you do your hobbies and recreation less often than you would like?
SR5.	Did you see your friends less often than you would like?
SR6.	Did you have sex less often than you would like?
SR7.	Did you feel that your phys. condition interfered with your social life?
SR8.	Did you feel that your language problems interfered with your social life?

6.2.2 Missing data

The proportion of missing data was very low ranging from 0-2.4%, with only 5 items out of 53 having any missing data. The 3 vision items (v1, v2, v3) were not applicable to a congenitally blind participant (1.2%). The "did you have to write things down to remember them" item (t4) (with verbal prompt for people with difficulty writing: "or ask somebody to write things down for you to remember them") was not applicable to one person who could not read or write (1.2%). Finally the item "did you stay out of family activities which were just for fun" (fr5) was not applicable to 2 participants who had no family in the country (2.4%). Scale mean scores could be calculated for 100% of participants.

6.2.3 *Distribution of scores*

Analysis of item endorsement frequencies showed that responses were distributed across response categories for most of the items. Nine items were affected by aggregate endorsement frequencies (AEF) at the low end of the response categories: sc2, sc8, m8, m9, v1, v2, v3, l4, md8. This meant that, on these variables, less than 10% of the respondents chose the 2 lowest responses i.e., the responses that indicate trouble in that variable. One of these items, v1, was also affected by maximum endorsement frequencies (MEF), which shows that most people had "no trouble" on this variable. These findings were somewhat expected as the participants had their strokes a long time ago and they have either recovered to a certain extent in some areas or adapted to some of their disabilities.

One item w3 "did you have trouble doing the work you used to do" was affected by AEF at the top end of the response categories indicating that despite their recovery, very few people after a stroke could go back to the work they used to do.

In summary, 10 of the 53 items of the SAQOL did not discriminate well between respondents.

6.2.4 *Skewness*

As expected there were some negatively skewed items. Nine items were negatively skewed with values greater than –1: sc8, m9, ue5, v1, v2, v3, l4, md3, md8. There were also two items, which were positively skewed, with values greater than +1: w3 and sr8. The overall proportion of skewed items was 20.7%, which was acceptable as it was below the set criterion of 25%. Table 6.4 gives a summary of the items that did not meet the acceptability criteria.

Table 6-4: SAQOL items not meeting the acceptability criteria

Criterion	Items
Missing data	None
MEF	v1 (ceiling effect)
AEF	w3 (floor effect) and sc2, sc8, m8, m9, v1, v2, v3, l4, md8
Skewness	sc8, m9, ue5, v1, v2, v3, l4, md3, md8, w3, sr8

6.3 Reliability

6.3.1 Internal consistency

Internal consistency of the whole scale (appendix 6.1): the SAQOL had an overall Cronbach's alpha coefficient of .93, indicating high internal consistency. Item-total correlations were also used to evaluate the homogeneity of the scale. They ranged from .07 to .67 with 11 items having item-total correlations below the criterion of .30. These items were: w3, v1, v2, v3, l4, l6, t4, p2, md8, sr5, sr6.

Internal consistency of the subdomains (appendix 6.2): eight of the twelve subdomains had an alpha of >.70. The work, vision, personality and family roles subdomains had alpha coefficients ranging from .58 to .69. Removal of certain items improved the subdomains internal consistency. In particular:

- Removal of w3 improved the work subdomain internal consistency from .58 to .67
- Removal of v2 improved the vision internal consistency from .68 to .76
- Removal of p3 improved the personality internal consistency from .61 to .71
- Removal of fr9 improved the family roles internal consistency from .69 to .72.

Item-total correlations within the subdomains were moderately high, with a few items failing the .30 criterion: w3, l4, p3, sr6, sr8.

6.3.2 Test-retest reliability

Test-retest reliability data were collected from 17 participants from the first recruitment site (Southwark and Lambeth). Appendix 6.3 presents the test-retest reliability respondents' characteristics. This sample of participants was similar to the overall sample in terms of age [mean (SD): 59.3 (16.6) as opposed to 61.6 (15.4)] and marital status (59% were in a relationship as opposed to 62%). There were however more male respondents (71% compared to 63% in the overall sample), more from ethnic minorities (35.2% compared to 21.7%) and they tended to be within 4 years after the stroke (100% compared to 69%). They were also more likely to be involved in some type of activity (12% in p/t, or voluntary work or students, compared to 6%). Their SAQOL mean scores were similar to the overall group [mean (SD): 3.67 (.8) compared to 3.38 (.6)].

The SAQOL had excellent test-retest reliability as indicated by the results presented in appendix 6.4 (1). All subdomains had intraclass correlation coefficients (ICCs) greater than the criterion of >.70 (range .84-.99). The overall scale mean score ICC was excellent at .98.

6.4 Validity

6.4.1 Construct validity

The construct validation of the SAQOL included testing first the validity of the overall scale and then testing the validity of the SAQOL subdomains.

6.4.1.1 Whole scale validation: Within scale analyses

6.4.1.1.1 Internal consistency

The overall scale alpha was .93 indicating high internal consistency. A number of items however had low item-total correlations (<.30) which suggested they did not fit well the underlying construct. These items were w3, v1, v2, v3, l6, l4, t4, p2, md8, sr5, sr6.

6.4.1.1.2 Intercorrelations between subdomains

Appendix 6.5 presents the intercorrelations between the SAQOL subdomains and between the subdomains and the corrected total mean. All subdomains had moderate to high correlations (ranging from .39 to .73) with the corrected total mean, except for the vision subdomain (.26). Vision also had very low correlations with most of the other subdomains (nine out of eleven were <.20). These values indicated that vision was measuring something different and minimally related to the other subdomains and the overall underlying construct.

With regard to the intercorrelations of the subdomains, as anticipated the ones that measured physical abilities (self-care, upper extremities, mobility, work) had high correlations with one another (.72-.84) and lower correlations (.09-.32) with other subdomains (mood and personality). All intercorrelations were below the set criterion of .80 except for the one of self-care with upper extremities (.84). This value indicated that these two subdomains measured very similar constructs and their existence as separate subdomains was questionable. Work also correlated highly with self-care (.79). Other than that, there was some evidence of unique reliable variance indicated by reliability coefficients for most subdomains being greater than the subdomains' intercorrelations.

In summary, the correlations between the subdomains and the corrected total mean score indicated that all except for vision contributed at least moderately to the underlying construct. The intercorrelations between the subdomains suggested that some of them measure the same or very similar concepts (self care, upper extremities, work). Thus, it was questionable whether there was enough evidence to support all of them as individual subdomains.

6.4.1.1.3 Factor analysis

6.4.1.1.3.1 Strategy 1: Factor analysis commencing with all-item SAQOL

Both a top-down and a bottom-up approach were followed, as explained in chapter 5.

Top-down approach

Principal Components Analysis (PCA) was undertaken within each subdomain to see whether each one measured one underlying component and then Principal Axis Factoring (PAF) was undertaken to see whether there were any items that contributed little to the underlying factor of each subdomain[8].

In the PCA (appendix 6.6), the Kaiser-Meyer-Olkin measure of sampling adequacy was acceptable (ranging from .55 to .87) and the Bartlett's test of sphericity was significant ($p<.001$) for each subdomain. All items loaded on the first component within each subdomain with loadings $\geq .40$, except for items l4 in the language domain and sr6 in the social roles domain (which still had loadings > .30).

In the PAF (appendix 6.7), some items were identified that contributed little to their intended subdomain (i.e., loaded less than .40). These were w3 (in work), v2 (in vision), l4 (in language), p3 (in personality), fr9 (in family roles), sr6 and sr8 (in social roles). Most of these items (w3, v2, p3, fr9) had already been identified as items that adversely affected the internal consistency of their subdomain in our reliability (internal consistency) analyses.

A second order FA was then performed, using the 12 subdomains' mean scores, to test whether a 12-factor model could be derived and if not whether the subdomains grouped into conceptually sound factors. PAF with varimax rotation was performed on the 12

subdomains' mean scores. The Kaiser-Meyer-Olkin measure of sampling adequacy was acceptable (.78) and the Bartlett's test of sphericity was significant ($p<.001$). A 3-factor model was extracted, that explained 54.4% of the variance (appendix 6.8). The vision subdomain did not load on any factor with values ≥.40 and two further subdomains crossloaded on two factors. The social roles domain loaded on factors 1 and 2 and the personality domain loaded on factors 2 and 3 (table 6.5). With these 2 subdomains crossloading on different factors and the language subdomain together with the physical domains, it was not feasible to interpret and name the factors.

Table 6-5: Factors derived from PAF of subdomains' mean scores

Factors	Subdomains
Factor 1	self care, mobility, work, upper extremities, language, (social roles)
Factor 2	(personality), mood, family roles, (social roles)
Factor 3	thinking, (personality), energy

( ): crossloaders

The rogue items identified, i.e., w3, v2, p3, l4, fr9, sr6 and sr8 were removed and the subdomains' mean scores were re-calculated. PAF with varimax rotation was run on the new subdomains' mean scores to see whether this would improve the model. The Kaiser-Meyer-Olkin measure of sampling adequacy was acceptable (.82) and the Bartlett's test of sphericity was significant ($p<.001$). Again, a 3-factor model was extracted, that explained 53.5% of the variance (appendix 6.9). This model was marginally better than the previous one, but still the vision subdomain did not load on any factor with values ≥.4. The mood subdomain crossloaded on factors 2 and 3 (table 6.6). Interpretability of the factors was a problem as in the previous model.

[8] Unless otherwise specified, the method of extraction for the PCA and PAF analyses is eigenvalues ≥1.

Table 6-6: Factors derived from PAF of subdomain mean scores, after removing rogue items

Factors	Subdomains
Factor 1	self care, mobility, work, upper extremities, language,
Factor 2	thinking, personality, energy, (mood)
Factor 3	family roles, social roles, (mood)

( ): crossloaders

Overall, these results indicated that the factor model derived from using the existing subdomains did not reflect all the subdomains (vision contributed very little). It was also conceptually unclear with the language domain grouping together with the physical domains and at least one subdomain contributing to more than one underlying factor. Thus, the top-down all item FA offered limited support for a 12-subdomain model of the SAQOL.

Bottom-up approach

In this approach, FA commenced by carrying out a PCA of all the items to see whether they all loaded on the first component (appendix 6.10). The Kaiser-Meyer-Olkin measure of sampling adequacy was acceptable (.61) and the Bartlett's test of sphericity was significant ($p<.001$). Five items (v2, v3, t4, md8, sr6) loaded very low (<.20) on the general component, which indicated they contributed little to the overall underlying concept.

PAF with varimax rotation was then carried out. Fourteen factors were extracted that explained 65.3% of the variance (appendix 6.11). The derived model was unstable as there were 5 items with no loadings ≥.40 on any factor, 5 items cross loading and after factor 11 there were only 1-2 items loading per factor. PAF with varimax rotation was then modelled on 12 factors to see whether the original 12-subdomain model could be replicated. The resulting model did not resemble the original 12-subdomain model. Only 1 item loaded on factor 12. It also had 5 items with no loadings ≥.40, and 3 items cross loading. The scree plot

(appendix 6.12) was difficult to interpret, but there seemed to be a kink after factor 4 and one after factor 7. Tabachnick and Fidell (2001) recommend that if one is unsure of the number of factors, one should perform several factor analyses, each time specifying a different number of factors, repeating the scree test and examining the residual correlation matrix. A conservative step-by-step approach was followed of extracting one less factor (11 and then 10, etc.) in each PAF (with varimax rotation) until the solution began to stabilise.

The Kaiser-Meyer-Olkin measure of sampling adequacy was acceptable (.61) and the Bartlett's test of sphericity was significant ($p<.001$). The 11, 10, 9, 8 factor models were all unstable with no or 1-2 items loading highest at the last factor and many items either crossloading or with no loadings ≥.40.

The model began to stabilise on the 7-factor solution (appendix 6.13). This model was conceptually acceptable. Factor 1 seemed to reflect physical abilities, factor 2 mood, factor 3 language/communication, factor 4 energy, factor 5 vision, factor 6 social roles and factor 7 thinking. Still, the last factor only had 2 items (loading ≥.40), and there were 2 crossloaders (fr5, fr8) in factors 1 and 2, and 5 items with no loadings ≥.40 (sc2, sc8, w3, l4, p2). These problems did not resolve with extracting even less factors and in order to improve and stabilise the model further, rogue items needed to be removed at this stage.

The items that crossloaded and the items that had no loadings ≥.40 (fr5, fr8, sc2, sc8, w3, l4, p2) were removed and PAF with varimax rotation was repeated. The resulting 7-factor model still had one factor with only 2 items and rogue items (t3 was crossloading and md8 had no loadings ≥.40).

A series of further FA were undertaken until a stable factor structure was derived. The process followed is described in appendix 6.14. The resulting model was derived from a 39-

item version of the SAQOL. The Kaiser-Meyer-Olkin measure of sampling adequacy was acceptable (.76) and the Bartlett's test of sphericity was significant ($p<.001$). Four factors were extracted (table 6.7, appendix 6.15) which explained 48% of the variance. This model had no crossloaders and no items only loading ≤.40.

Table 6-7: Factor structure of the 39-item SAQOL (SAQOL-39)

Factors	Items
Factor 1 (physical)	sc1, sc4, sc5, m1, m4, m6, m7, m8, m9, w1, w2, ue1, ue2, ue4, ue5, ue6, sr7
Factor 2 (psychosocial)	t5, p1, p3, md2, md3, md6, md7, fr7, sr1, sr4, sr5
Factor 3 (communication)	l2, l3, l5, l6, l7, fr9, sr8
Factor 4 (energy)	t4, e2, e3, e4

In this model, items from the self-care, the mobility, the work and the upper-extremities domains of the SS-QOL grouped together to form an overall physical domain. The items from the language domain grouped with the two items that reflected the impact of language problems on family life (fr9) and social roles (sr8) to form a domain that was given the broader name of communication. The rest of the social roles domain items grouped together with the mood and personality items and an item on difficulty in making decisions (t5) that was originally under the thinking domain in the SS-QOL. This domain was seen as reflecting psychosocial aspects of stroke and aphasia. Lastly, the energy items grouped with the item on having to write things down to remember them (t4). For people with aphasia, who have difficulties with reading and writing, this activity may well require more energy than it would

for people with no such difficulties. The overall domain was therefore interpreted as reflecting drive and energy.

In summary, the bottom-up all item FA of the SAQOL did not support the 12-subdomain model of the original SS-QOL. Instead a number of items that did not contribute well to the underlying structure of the instrument needed to be removed in order to derive a stable factor model. This resulted in a 39-item version of the SAQOL (SAQOL-39) (appendix 6.16) with a conceptually clear and psychometrically sound 4-factor structure.

6.4.1.1.3.2 Strategy 2: Factor analysis commencing with item reduced SAQOL

Table 6.8 summarises the criteria for item reduction and the items that failed these criteria and were removed.

Table 6-8: Criteria for item reduction, items failing these criteria and items eliminated from SAQOL

Criterion	Items failed	Items eliminated
Missing data (<5%)	None	None
MEF (≤80%)	v1	v1
AEF (≥10%)	sc2, sc8, m8, m9, w3, v1, v2, v3, l4, md8	sc2, sc8, m8, m9, w3, v1, v2, v3, l4, md8
Redundancy (≤75%)	m1 with m7, sc4 with ue2	m1, ue2
Item-total correlations (≥.3)	w3, v1, v2, v3, l4, l6, t4, p2, md8, sr5, sr6	w3, v1, v2, v3, l4, l6, t4, p2, md8, sr5, sr6

This process resulted in a 36-item version of the SAQOL. PCA was carried out to see whether all the items loaded on the first component (appendix 6.17). The Kaiser-Meyer-Olkin measure of sampling adequacy was acceptable (.80) and the Bartlett's test of sphericity

was significant (p<.001). All items loaded with values >.20 on the general component, which indicates that they all contribute to the overall underlying concept.

PAF with varimax rotation was carried out to explore the underlying factor structure. Eight factors with eigenvalues ≥ 1 were extracted in a model that explained 59.5% of the variance. In this model, there were only 1 or 2 items per factor after factor 5 and there were 3 crossloaders (m4, t3, fr8) and one item that loaded <.40 (p1). The scree plot changed direction after factor 4 and thus PAF with varimax rotation was repeated asking for 4 factors to be extracted. The resulting model (appendix 6.18) explained 48% of the variance. There were no crossloaders but there were 2 items with loadings <.40 (sr1 and sr4) that also had very similar loadings on factors 1 and 2. To increase the interpretability of the model, these two items were removed and the analysis was rerun. The Kaiser-Meyer-Olkin measure of sampling adequacy was acceptable (.80) and the Bartlett's test of sphericity was significant (p<.001). A 4-factor model was extracted from this 34-item version of the SAQOL (SAQOL-34, appendix 6.19) which explained 49% of the variance (appendix 6.20). This model was conceptually very similar to the SAQOL-39 as its factor structure indicates (table 6.9).

Table 6-9: Factor structure of the 34-item SAQOL (SAQOL-34)

Factors	Items
Factor 1 (physical)	sc1, sc4, sc5, m4, m6, m7, w1, w2, ue1, ue4, ue5, ue6, fr8, sr7
Factor 2 (psychosocial)	t2, t5, p1, p3, md2, md3, md6, md7, fr5, fr7
Factor 3 (communication)	l2, l3, l5, l7, fr9, sr8
Factor 4 (energy)	t3, e2, e3, e4

In summary, the psychometric evaluation of the SAQOL presented so far did not support a 12-subdomain structure. The internal consistency analyses of the scale indicated that some items did not contribute enough to the overall scale and that some subdomains measured almost the same concept. In the factor analyses there were items that needed to be removed in order to derive a stable and interpretable factor structure. Two versions of the SAQOL were derived following factor analyses, the SAQOL-39 and the SAQOL-34, with similar factor structure but some different items.

6.4.1.2 Whole scale validation: Comparisons with external criteria

The planned comparisons to assess the construct validity of the whole SAQOL scale were also appropriate for the 2 new versions of the SAQOL. The SAQOL-39 and the SAQOL-34 were just shorter versions of the instrument. No items were added that would considerably change the overall underlying concept of HRQL. The results for the 3 instruments will be presented in parallel.

6.4.1.2.1 Convergent validity

An indication of the overall measure's convergent validity was derived by comparing the SAQOL/SAQOL-39/SAQOL-34 scores of three groups of respondents: those who rated their HRQL as a lot worse than before the stroke (group 1), those who rated it as a little worse than before (group 2), and those who rated it as the same or better than before (group 3). It was hypothesized that group 1 would have lower SAQOL/SAQOL-39/SAQOL-34 scores than group 2, which in turn would have lower SAQOL/SAQOL-39/SAQOL-34 scores than group 3.

Univariate analysis of variance (ANOVA) was used to test this hypothesis. The results of the ANOVA for the SAQOL, the SAQOL-39 and the SAQOL-34 are reported in appendices

6.21, 6.22 and 6.23 respectively. The assumptions of normality and homogeneity of variance were met for all three versions of the measure (Kolmogorov-Smirnof tests of normality and Levene's tests of equality of variance not significant, appendices 6.21, 6.22 and 6.23). As anticipated there was a significant effect of quality of life group for SAQOL score ($F(2, 80)=11.340, p<.001$) and also for SAQOL-39 score ($F(2, 80)=10.609, p<.001$) and SAQOL-34 score ($F(2, 80)=11.939, p<.001$).

These results show that the mean SAQOL/SAQOL-39/SAQOL-34 scores were significantly different between groups 1, 2 and 3 but they do not indicate whether each group was significantly different from both the other two. Pairwise comparisons were therefore undertaken (e.g., group 1 versus group 2 and group 3, group 2 versus group 1 and group 3, etc.). All comparisons were significant ($p<.05$ level) for SAQOL (appendix 6.21) and SAQOL-39 (appendix 6.22). For SAQOL-34 the mean score of group 2 (people who scored their HRQL as a little worse than before the stroke) was not significantly different from group 3 (people who scored their HRQL as the same or better than before the stroke) (appendix 6.23).

Overall, these ANOVA results offer support for the convergent validity of the SAQOL, SAQOL-39 and SAQOL-34 mean scores. The SAQOL-34 however is not as strong as the other two versions of the instrument in picking up HRQL differences between people who feel their HRQL is little affected by the stroke and those who feel their HRQL is not affected by the stroke.

6.4.1.2.2 Correlations with related variables and disriminant validity

It was anticipated that the SAQOL/SAQOL-39/SAQOL-34 total mean scores would have moderate correlations with instruments measuring related variables (GHQ-12, FAI, and

ASHA-FACS) and lower correlations with instruments measuring less related or unrelated variables (FAST, RCPM and SSS).

Table 6.10 shows the results of these comparisons. All three measures had moderate correlations with the related variables (.43-.59) and lower correlations with the less related or unrelated variables (.19-.31). These results offer support for the correlated measures and discriminant validity of all three versions of the SAQOL.

Table 6-10: Correlations with related variables and discriminant validity of mean scores of SAQOL, SAQOL-39, SAQOL-34.

N=83	Correlations with related variables			Discriminant validity		
	GHQ-12	FAI	ASHA-FACS	FAST	RCPM	SSS
SAQOL	.58**	.59**	.44**	.29**	.29**	.26*
SAQOL-39	.53**	.58**	.46**	.31**	.27*	.19
SAQOL-34	.55**	.55**	.43**	.29**	.29**	.22*

**: significant at .01 level *: significant at .05 level

In summary, the above comparisons with external criteria supported the whole scale construct validity of the SAQOL, the SAQOL-39 and the SAQOL-34.

6.4.1.3 Validation of the subdomains

The results for the SAQOL subdomains' validation will be presented first. Then the process of assessing the subdomains' validity of the 2 versions of the instrument, the SAQOL-39 and

the SAQOL-34, that were derived from the factor analysis will be briefly presented before concentrating on the results for these 2 versions.

6.4.1.3.1 Validation of the SAQOL subdomains

Table 6.11 presents the results of the SAQOL subdomains' construct validation. No absolute criteria were set for the size of the correlations, but for each subdomain (i.e., within each row) the correlations in the second column (convergent validity) should be higher than those in the third column (correlated measures), which in turn should be higher than those in the fourth column (discriminant validity). The results supported the validity of the language, the work, and the energy subdomains.

Table 6-11: Results of the construct validity of the SAQOL subdomains against external criteria.

Subdomains	**Validity**		
	Convergent	**Correlated measures**	**Discriminant**
Language	FAST: .63** ASHA-FACS: .61**		RCPM: .11 FAI: .31** SSS: .10
Thinking	RCPM: .06	GHQ-12: .41**	FAST: .03 ASHA-FACS: .09 FAI: .18 SSS: .20
Personality		GHQ-12: .57** SSS: .40**	FAST: .03 ASHA-FACS: -.04 RCPM: .18 FAI: .18
Energy		GHQ-12: .32**	FAST: -.09 ASHA-FACS: .02 RCPM: .19 SSS: .13

Subdomains	Validity		
	Convergent	**Correlated measures**	**Discriminant**
Mood	GHQ-12: .57**	SSS: .24*	FAST: .11 ASHA-FACS: .18 RCPM: .27* FAI: .20
Family Roles		FAI: .29** SSS: .24* GHQ-12: .41**	FAST: .12 ASHA-FACS: .21 RCPM: .31**
Social Roles		FAI: .37** SSS: .18 GHQ-12: .41**	FAST: .24* ASHA-FACS: .34** RCPM: .31**
Work	FAI: .61**	FAST: 32** GHQ-12: .34**	SSS: .07

**: probability significant at the .01 level *: probability significant at the .05 level
__ : values not supporting construct validity of subdomain in relation to other values in the same row.

The personality subdomain correlated as highly with the GHQ-12 as the mood subdomain. This high correlation with a measure of emotional well-being cast doubts on whether indeed the personality subdomain measured something different from well-being/mood. The social roles subdomain had a very low correlation with the Social Support Scale and higher correlations with the activities measures (the FAI and the communication activities measure, the ASHA-FACS). It, thus, seemed to reflect social activities rather than social roles. The mood, the social roles and the family roles subdomains had higher than anticipated correlations with the cognitive measure, the RCPM. No explanation was found for these higher than anticipated correlations. On the contrary, the thinking subdomain had a very low correlation with the RCPM, which seems to indicate it measures something different from cognitive functioning.

Overall, the results of the construct validation of the SAQOL subdomains against external criteria did not support all of them as individual entities. In addition, for some of them, it is unclear what their underlying construct was (personality, thinking, social roles).

6.4.1.3.2 Validation of the SAQOL-39 and SAQOL-34 subdomains

These two instruments had a 4-subdomain structure. As a result, the construct validity testing of their 4 subdomains was different from the originally planned validation of the SAQOL subdomains (8 tested out of 12), which has been presented in chapter 5. Thus, the methods of testing the constuct validity of the SAQOL-39 and SAQOL-34 subdomains will be presented first and then the results.

6.4.1.3.2.1 Methods

The constuct validity of the SAQOL-39 and the SAQOL-34 subdomains was tested by exploring their correlations with external measures. No criteria were set for the absolute size of these correlations, but the following direction should be observed. Each subdomain should have higher correlations with the measures against which its convergent validity was tested than with the measures against which its correlated measures validity was tested. It should also have higher correlations with the measures against which its correlated measures validity was tested than with the measures against which its discriminant validity was tested.

The construct validation of the SAQOL-39 and the SAQOL-34 subdomains is presented in table 6.12. In the validation of the SAQOL subdomains the validity of the physical subdomains was not tested. One of the reasons behind this decision was that no external measures sensitive enough to differentiate between self-care and upper extremities functions, for example, could be found. In the SAQOL-39 and SAQOL-34, however, physical items on self-care, mobility, hand function were grouped together into one subdomain. This allowed

for the testing of the construct validity of this physical domain against external measures that had been used already. It was hypothesized that the physical subdomain should correlate highly with the FAI, the measure of activities after stroke. It should also correlate moderately with the ASHA-FACS, the measure of activities that require communication; and the GHQ-12 since functional disability after stroke is related to emotional distress (e.g., Ebrahim et al, 1987; Burvill et al., 1997; Herrmann et al., 1998; Kotila et al., 1998; Dennis et al., 2000). The physical subdomain should have lower correlations with unrelated measures, i.e., the FAST, the RCPM and the SSS.

Table 6-12: Construct validation of the SAQOL-39 and SAQOL-34 subdomains

Validity	**SAQOL-39 and SAQOL-34 subdomains**			
	Physical	**Psychosocial**	**Communication**	**Energy**
Convergent	FAI	GHQ-12	FAST ASHA-FACS	
Correlated measures	GHQ-12 ASHA-FACS	SSS FAI		GHQ-12
Discriminant	FAST RCPM SSS	FAST ASHA-FACS RCPM	RCPM FAI SSS	FAST ASHA-FACS RCPM SSS

The psychosocial subdomain of the SAQOL-39 and the SAQOL-34 was tested against the GHQ-12, with which it should correlate highly. Moderate correlations were expected with measures of related constructs such as the SSS. Moderate correlations were also expected with the FAI since, as has been already indicated, reduced activities after stroke affect well-being. Lower correlations were expected with measures that assessed difficulties in one

particular area of functioning, such as language (FAST and ASHA-FACS) and cognition (RCPM).

The communication and the energy subdomains were tested in the same way as the language and the energy subdomains of the SAQOL. The communication domain should have high correlations with the language measures (FAST and ASHA-FACS), and low correlations with the RCPM, the FAI and the SSS. The energy domain should have moderate correlations with the GHQ-12 and lower correlations with the FAST, the ASHA-FACS, the RCPM and the SSS.

6.4.1.3.2.2 Results

Table 6.13 shows the results of the construct validation of the SAQOL-39 and SAQOL-34 subdomains against external criteria.

The convergent validity of the physical, psychosocial and communication domains was supported by high correlations with measures of the same or very similar underlying constructs (ranging from .55 to .67 for the SAQOL-39 and .54-.66 for the SAQOL-34). The validity of the physical, psychosocial and energy domains was further supported by moderate correlations with measures of related constructs (.28-.42 for the SAQOL-39 and .26-.43 for the SAQOL-34). Both measures also had good discriminant validity as indicated by lower correlations with measures of unrelated constructs (.02-.27 for the SAQOL-39 and .02-.27 for the SAQOL-34). Although there appears to be a slight overlap of values, within each subdomain correlations with similar measures (convergent validity) were higher than correlations with related measures (correlated measures validity), which in turn were higher than correlations with less related measures (discriminant validity).

Table 6-13: Results of the construct validation of the SAQOL-39 and SAQOL-34 subdomains

Validity	SAQOL-39[1] and *SAQOL-34*[2] subdomains			
	Physical	**Psychosocial**	**Communication**	**Energy**
Convergent	FAI: .67**, *.66***	GHQ-12: .62**, *.64***	FAST: .55**, *.54*** ASHA-FACS: .55**, *.56***	
Correlated measures	GHQ-12: .39**, *.399*** ASHA-FACS: .42**, *.43***	SSS: .28*, *.33*** FAI: .31**, *.26**		GHQ-12: .32**, *.37***
Discriminant	FAST: .26*, *.27** RCPM: .20, *.22** SSS: .10, *.10*	FAST: .12, *.08* ASHA-FACS: .20, *.14* RCPM: .27*, *.23**	RCPM: .16, *.17* FAI: .21, *.20* SSS: .08, *.05*	FAST: -.10, *-.09* ASHA-FACS: .02, *.02* RCPM: .14, *.21* SSS: .12, *.17*

[1]: correlations of the SAQOL-39 presented in normal font
[2]: correlations of the SAQOL-34 presented in *grey italics*
*: probability significant at the .05 level
**: probability significant at the .01 level

Overall, our results offered good support for the construct validity of the SAQOL-39 and the SAQOL-34 subdomains.

6.5 Comparison of the three SAQOL versions

So far the results from the acceptability, reliability and validity testing of the SAQOL have been presented. During the factor analysis of the SAQOL two shorter versions of the questionnaire were derived: the SAQOL-39 and the SAQOL-34. The results of the comparisons of these two measures with external criteria (in their construct validity assessment) have also been presented. Here, the results of the acceptability, reliability and within scale validity analyses of the SAQOL-39 and the SAQOL-34 are presented.

Comparisons are drawn between the three versions of the SAQOL on all aspects of their acceptability, reliability and validity. Table 6.14 summarises their psychometric properties.

6.5.1 Acceptability

The SAQOL-39 and the SAQOL-34 had very few items affected by problematic endorsement frequencies (2 items of the SAQOL-39 only) and skewness (<11%). They compared favourably to the SAQOL, in which 21% of the items were affected by skewness and 10 items were affected by AEF/MEF.

6.5.2 Reliability

The SAQOL-39 and the SAQOL-34 had very good internal consistency (scale alpha ≥.92 and subdomains >.73) (appendices 6.24, 6.25, 6.26, 6.27) and test-retest reliability (ICCs ranging from .89-.99) (appendix 6.4 (2) and 6.4 (3)). The SAQOL also had good test-retest reliability (ranging from .85-.99) and scale internal consistency (.93), but 4 of its subdomains had poor internal consistency (<.70).

6.5.3 Validity

In the construct validity testing of the instruments within scale analyses were undertaken, which included checking the intercorrelations between the subdomains and between the subdomains and the corrected total mean for too high or too low correlations and factor analysis. The intercorrelations between the subdomains of the SAQOL did not support the 12-subdomain structure of the instrument. One of them (vision) had very low correlations with most of the subdomains (<.20 for 9 out of 11) and the corrected total mean (.26). Two other (self-care and upper extremities) seemed to measure the same or a very similar underlying construct (.84). The factor analysis did not support the 12-subdomain structure of

the SAQOL either. The models that included all the SAQOL items were hard to interpret and unstable (with items loading <.4, items crossloading and factors with less than 3 items).

Table 6-14: Psychometric properties of the SAQOL versions.

	SAQOL	SAQOL-39	SAQOL-34
Acceptability			
Missing data	0-2.4%	0-1.2%	0-2.4%
AEF, MEF	10 items affected*	2 items affected	No items affected
Skewness (>±1)	11 items affected (21%)	4 items affected (10.2%)	3 items affected (8.8%)
Reliability			
Internal consistency - scale	Good (.93)	Good (.93)	Good (.92)
Internal consistency - subdomains	4 domains < .7	All good (.74-.94)	All good (.77-.93)
Test-retest reliability - scale	Good (.98)	Good (.98)	Good (.99)
Test-retest reliability – subdomains	All good (.84-.99)	All good (.89-.98)	All good (.90-.98)
Validity (within scale analyses)			
Intercorrelations between subdomains	Self-care with upper ext.:.84, too high	All acceptable (.09-.47)	All acceptable (.06-.47)
Correlations of sub-domains and mean	Vision=.26 too low	All acceptable (.38-.58)	All acceptable (.36-.60)
Factor analysis	Reduced support for original 12 subdomain model	4 factors (with > 3 items per factor, no crossloaders, no items loading <.4)	4 factors (with > 3 items per factor, no crossloaders, no items loading <.4)
Validity (comparisons with external criteria)			
Scale - convergent	Good	Good	Adequate
Scale – correlations with related measures	Good (.44-.57)	Good (.45-.58)	Good (.43-.55)
Scale - discriminant	Good (.25-.29)	Good (.19-.31)	Good (.22-.29)
[9]Subdomains - convergent	1 of the 4 tested poor (.06)	All good (.55-.67)	All good (.54-.66)

[9] Absolute size of correlation not critical in these, as the set criterion is that, for each subdomain, correlations with similar measures (convergent validity) should be higher than correlations with just related measures (correlated measures validity), which should in turn be higher than correlations with unrelated measures (discriminant validity).

	SAQOL	SAQOL-39	SAQOL-34
Subdomains - correlations with related measures	2 of the 7 tested poor (.57, .18)	All good (.28-.42)	All good (26-.43)
Subdomains - discriminant	3 of the 8 tested poor (.27-.34)	All good (.02-.27)	All good (.02-.27)

*: Problematic areas shaded in dark grey.

In contrast, in the SAQOL-39 all 4 of the subdomains correlated moderately with the corrected total mean (.38-.58). They also had low to moderate correlations with one another as anticipated (.09-.47) (appendix 6.28 (2)). This was also the case for the SAQOL-34 (.06-.47). All of the SAQOL-34 subdomains also correlated moderately with the corrected total mean (.40-.60) (appendix 6.28 (1)). Overall, the intercorrelations between the subdomains and between the subdomains and the corrected total mean supported the 4-subdomain structure of the SAQOL-39 and the SAQOL-34. The factor analysis offered extra support for the 4-subdomain structure of the SAQOL-39 and the SAQOL-34. All the items loaded considerably on one factor ($\geq .40$) and there were no items crossloading to other factors. All the factors were conceptually clear and had > 3 items.

The construct validity of the three versions of the instrument was also assessed through comparisons with external criteria. The results indicated that all three versions had good overall scale construct validity (appendices 6.21, 6.22, 6.23; table 6.10).

In terms of the validity of their subdomains, some of the SAQOL subdomains (thinking, personality, family and social roles) had poor construct validity. The thinking subdomain did not seem to reflect any cognitive processes as it correlated very low with the RCPM (.06) but it had a moderate correlation with the GHQ-12 (.41). The personality subdomain did not seem to reflect anything different from the mood subdomain (similar pattern of correlations with all external measures and identical correlations with the GHQ-12: .57). The family roles

and social roles subdomains were more closely related to mood (correlations with the GHQ-12 for both: .41) and cognition (correlations with the RCPM for both: .31) than to the related concept of social support (correlations with the SSS: .24 and .18 respectively). These results suggested that there was not enough evidence to support these subdomains as individual entities, measuring the intended distinct concepts.

In the SAQOL-39 and the SAQOL-34, items from the personality, mood, family roles and social roles SS-QOL domains were all in one subdomain, which was seen as reflecting psychosocial aspects. Some of the thinking items were also there (t2: difficulty concentrating and t5:difficulty making decisions in the SAQOL-34 and t5 in the SAQOL-39), suggesting that these items were more related to low mood than cognitive functioning. The thinking items that were about effort in remembering things (t3 and t4) grouped together with the energy items. Overall, the SAQOL-39 and the SAQOL-34 subdomains had very good construct validity, as their comparisons with external criteria suggested (table 6.13).

The psychometric properties of the SAQOL-39 and the SAQOL-34 were very similar. Both instruments were also considerably shorter than the SAQOL, which would reduce respondent burden during administration. In terms of their conceptual cover, the SAQOL-34 had one item in the physical domain (fr8) on the effects of physical problems on family life that the SAQOL-39 did not have. Still, the SAQOL-39 had four items more than the SAQOL-34 in the physical domain (m1, m8, m9, ue2). The SAQOL-39 had one more item in the communication domain (l6) on difficulties with word finding, which is commonly affected in people with aphasia. The SAQOL-34 had two items in the psychosocial domain that the SAQOL-39 did not have (t2, on difficulty concentrating and fr5, on avoiding family activities). Still, it was noted that fr5 was an item that 2 of the participant in the pre-test study had identified as difficult. Moreover, the SAQOL-39 had more items in this domain. In

particular, it had more items on social activities (sr1, sr4 and sr5 on going out, doing hobbies and seeing friends less), an area often considerably affected in people with aphasia (e.g., Parr et al., 1997). Overall, these results suggested that the SAQOL-39 had a greater conceptual depth than the SAQOL-34, without being considerably longer.

6.6 Summary and implications of findings

This chapter presented the results of the psychometric evaluation of the SAQOL, the SAQOL-39 and the SAQOL-34 on a sample of 83 people with long-term aphasia (> 1 year) following stroke. All 83 participants were able to complete the SAQOL in an interview format and missing data were minimal. The acceptability of the SAQOL was weakened by an uneven distribution of scores: 10 items were affected by AEF (<10%), one item was affected by MEF (>80%) and 21% of the items were affected by skewness. These items would not discriminate well between respondents. Most of these items were removed in the shorter versions of the SAQOL and as a result the SAQOL-39 and the SAQOL-34 had higher acceptability.

The results supported the reliability and the validity of the overall SAQOL scale, but not of its subdomain structure. Four of the 12 subdomains (work, vision, personality and family roles) had poor internal consistency and three had poor overall external measures validity (thinking, personality, social roles). Factor analysis indicated that not all of the items contributed considerably to the overall score and that no stable, conceptually clear factor structure could be derived from all of the 53 items of the questionnaire. Two shorter versions of the SAQOL were derived, the SAQOL-39 and the SAQOL-34. Both these instruments had a stable, conceptually clear 4-factor structure and high scale and subdomain internal consistency, test-retest reliability and construct validity.

Overall, the results suggested that the SAQOL did not have adequate acceptability, reliability and validity with our sample of people with long-term aphasia following stroke. In contrast, the SAQOL-39 and the SAQOL-34 had good psychometric properties with this group of people. The SAQOL-39 was also broader conceptually than the SAQOL-34, without being considerably longer. As such, it stood out as the preferred version of the SAQOL and it was the version that was used in further analyses.

7 PREDICTORS OF HRQL: METHODS AND RESULTS

The second research question of this project explored the predictors of HRQL in people with long-term aphasia following stroke. A cross sectional study was undertaken to address this question and the overall design of the study and the measures used have been described in chapter 3. This chapter concentrates on the methods used to analyse the data collected and the results of the analysis.

7.1 Methods

Multiple regression analysis was undertaken to assess the relationship between HRQL (dependent variable, DV) and several potential predictors (independent variables, IVs). The potential IVs included demographic, stroke and health related variables, variables that have been identified as predictors of HRQL in people with stroke and variables of theoretical interest. They were: age, gender, ethnic background, marital status, socioeconomic status, employment status (demographic variables); type of stroke, time post onset of stroke, and comorbidity (stroke and health related variables); and psychological distress, level of activity, communication disability, cognition, social support and satisfaction with services for stroke.

HRQL, the DV, was assessed with the SAQOL-39. As has been described in chapter 3 information on demographic, stroke and health related variables was collected from the participants SLT records. This information was confirmed and supplemented through a short case history interview with the participants. Psychological distress was measured with the GHQ-12, level of activity with the FAI, severity of aphasia/communication disability with the ASHA-FACS, cognition with the RCPM, social support with the SSS and satisfaction with services received for stroke with the PSI.

7.1.1 *Multiple regression assumptions*

Multiple regression analysis requires that certain assumptions are met, in order for the model derived to be *unbiased*[10] and applicable to the population of interest (Field, 2000). These include the ratio of cases to IVs; the absence of outliers among the IVs and on the DV; the absence of multicollinearity; the normality, linearity and homoscedasticity of residuals (the differences between obtained and predicted DV scores); the independence of errors; and the examination for outliers in the solution. Various procedures and diagnostic tests were undertaken to test these assumptions.

There was a large number of potential predictors and a relatively modest sample size. This could challenge the viability of the regression analysis by reducing the cases to variables ratio. Tabachnick & Fidell (2001) suggest that for testing multiple correlation the simplest rule of thumb is $n \geq 50 + 8m$ (where m is the number of IVs). To control the number of variables that would enter the regression model, univariate analyses were initially undertaken between each IV and HRQL. One-way factor ANOVA, independent t-tests and Pearson's product correlation coefficients were calculated depending on the nature of the IVs. The demographic, stroke and health variables that were not significantly associated ($p > .05$) with HRQL in univariate analyses were not entered in the regression model. It was decided that emotional distress/depression, reduced activities, cognitive level, aphasia, social support and satisfaction with stroke care would be included in the regression model, if they were correlated with HRQL at a $p < .1$. This decision was based on two main reasons. These variables are of theoretical interest as they have been implicated in previous research and their contribution to HRQL for people with aphasia needs to be assessed and better

[10] An unbiased model indicates that *on average* the regression model from a sample is the same as the population model (Field, 2000)

understood. They are also of greater interest to care providers as they may be addressed in rehabilitation and be subject to intervention.

Outliers among the IVs and on the DV were explored with Cook's distance, leverage values and Mahalanobis distances. Cook's distance is a measure of the influence of a case on the model and values greater than 1 may be outliers. Leverage gauges the influence of the observed value of the DV over the predicted values. If no cases exert undue influence on the model then all cases should be close to the average leverage value. The average leverage value is defined as $(m+1)/n$ (where m is the number of IVs). Stevens (1992) criterion was followed where no cases should have leverage values of more than 3 times the average. Mahalanobis distance measures the distance of cases from the means of the IVs. It is distributed as a chi-square (χ^2) variable, with degrees of freedom equal to the IVs. To determine if any cases were multivariate outliers the critical χ^2 at the desired alpha level was inspected (table C.4, p 933, Tabachnick & Fidell, 2001). If outliers were to be identified they should be deleted, rescored or the variables should be transformed (Tabachnick & Fidell, 2001).

Absence of multicollinearity means that IVs should not be strongly correlated with one another. The correlation matrix of the IVs was inspected for high correlations. For absence of multicollinearity correlation coefficients should be below .80 (Field, 2000). The tolerance statistic was also used to test for multicollinearity. Its values should not be below .2 (Menard, 1995).

The normality, linearity and homoscedasticity of residuals mean that residuals are normally distributed about the predicted DV values, that they have a straight-line relationship with the predicted DV values and that the variance of residuals about the predicted DV values is the

same for all predicted DV values (Tabachnick & Fidell, 2001). Normality was tested by inspecting the histogram of the standardised residual plots and their normal probability plots. The histogram should look like a normal distribution (bell-shaped). The normal probability plots include a straight line that represents the normal distribution and dots that represent the residuals. These dots should lie on or very close to the line. Homoscedasticity and linearity were tested by inspecting the scatterplot of the standardised residuals against the standardised predicted values of the DV. This should look like a random array of dots evenly dispersed around zero for homoscedasticity to be met. There should be no curve for the assumption of linearity to be met.

The assumption of independence of errors states that errors of prediction (residuals) are independent of one another. This was tested with the Durbin-Watson statistic, which tests for correlations between residuals. It can vary between 0 and 4 with a value of 2 meaning that the residuals are uncorrelated. Values greater than 1 and less than 3 were seen as acceptable (Field, 2000).

The examination for outliers in the solution identifies cases that are poorly fit by the regression model and lower the multiple correlation. To assess whether there were any outliers in the solution, cases with standardised residuals greater than an absolute value of 2 were identified. Ninety-five per cent of standardised residuals should lie within ± 2. If more than 5% of cases had standardised residuals with an absolute value greater than 2 then the model would be a poor representation of the actual data.

7.1.2 *Multiple regression method and analyses*

With regard to regression method, standard multiple regression[11] was used (Tabachnick & Fidell, 2001). In standard multiple regression all IVs are entered in the regression equation simultaneously. This way, each IV is evaluated in terms of what it adds to the prediction of the DV that is different from the predictability afforded by all other IVs (Tabachnick & Fidell, 2001). This method was preferred to sequential[12] and statistical regression as it suited best the research question: standard multiple regression is the method of choice when the relationship among variables is assessed and the multiple correlation among variables is explored (Tabachnick & Fidell, 2001).

The initial standard multiple regression analysis identified which IVs contributed significantly to the regression model. Following the initial analysis, the regression was repeated this time excluding any variables that were statistically redundant (Field, 2000). This approach allowed for a smaller number of IVs to be used in the final regression model, leading to an improved cases to variables ratio.

The following multiple regression analyses were undertaken. ANOVA was used to determine whether the *R* for regression was significantly different from zero. R^2 was calculated to get the amount of variance in the DV explained by the model. The adjusted R^2 was calculated to estimate the amount of variance in the DV explained by the model if the model had been derived from the population from which the sample was taken. The individual contribution of the IVs was assessed by inspecting their unstandardised coefficients (*B*). These represent the change in the DV associated with a unit change in the IV and *t*-statistics were used to assess whether their contribution was significant. The standardised (β) regression coefficients

[11] Standard multiple regression is also known as forced entry regression.

[12] Sequential multiple regression is also known as hierarchical regression.

(the change in the DV associated with a standard deviation change in the IV) were calculated to get a better insight into the relative contribution of each IV. Squared semipartial correlations (sr^2) were also calculated. In standard multiple regression, sr^2 for an IV is the amount by which R^2 is reduced if that IV is removed from the regression equation. It represents the unique contribution of the IV to R^2 (Tabachnick & Fidell, 2001). The difference between total variance (R^2) and unique variance (the sum of the sr^2 of the IVs) was also calculated as it represents variance that the IVs contribute jointly to the total variance explained by the model.

In summary, standard multiple regression analysis was used to assess the relative impact of a selected set of IVs on HRQL. Demographic, stroke and health related variables were entered in the regression model if they were significantly associated with HRQL ($p < .05$) in univariate analyses. Emotional distress, level of activities, communication disability, level of cognition, social support and satisfaction with stroke care were entered in the regression model if they were associated with HRQL ($p < .1$) in univariate analyses. The assumptions of multiple regression were tested. All data analyses were performed using SPSS 10.0 for Windows (SPSS Inc., 1999).

7.2 Results

7.2.1 Respondents characteristics

The respondents were described briefly in chapter 6 and their characteristics were detailed in table 6.2. This table is reproduced here for ease of reference (table 7.1).

The majority were male (62.7%) and they ranged in age from 21 to 92 (mean 61.67±15.47). About 43% were over 65 years old and 15.7% were between 21 and 45. The majority of the sample were white (78.3%) and married/had a partner (62.6%). Although almost 56% of the

sample were of working age (≤ 65) only 6% were involved in some type of work (part-time or voluntary work and students). No participants were in full-time work.

Table 7-1: Respondents' characteristics

Characteristics	N=83	Percent
Gender		
Female	31	37.3%
Male	52	62.7%
Age		
Mean (SD)	61.67 (15.47)	
Range	21-92	
21-45	13	15.7%
46-65	34	41%
66+	36	43.4%
Stroke type		
Ischaemic	36	43.4%
Haemorrhagic	16	19.3%
Unknown	31	37.3%
Time post stroke		
Mean in years (SD)	3.5 (3.09)	
Range	1y 1m-20y 10m	
1-2 years post onset	26	31.3%
2.1-4 years post onset	31	37.3%
4.1+ years post onset	26	31.3%
Comorbidity		
None or one comorbid condition	34	41%
Two or more comorbid conditions	49	59%
Ethnic group		
Asian	7	8.4%
Black	11	13.3%
White	65	78.3%
Marital status		
Married	42	50.6%
Has partner	10	12%
Single	14	16.9%
Divorced or spouse died	17	20.5%
Socioeconomic status (revised SEC)		
Professionals/senior managers	23	27.7%
Ass. Professional/ junior managers	6	7.2%
Other admin. and clerical workers	13	15.7%
Own account non-professional	5	6%
Supervisors, technicians and related workers	11	13.3%
Intermediate workers	9	10.8%
Other workers	12	14.5%
Never worked/other inactive	4	4.8%
Employment status		
Retired before the stroke	31	37.3%

Characteristics	N=83	Percent
Inactive because of the stroke	47	56.6%
Some p/t or voluntary work	3	3.6%
Students	2	2.4%

Participants' socioeconomic class was determined by their last occupation before the stroke, using the collapsed version of the socioeconomic classification (SEC). According to this, approximately 35% were professionals and managers, 35% were other administrative and clerical workers, or own account non-professional and supervisors, or technicians and related workers, 25% were intermediate or other workers and 5% had never worked.

Data on the type of stroke the respondents had suffered were available for about 63%. The majority of them had suffered an ischaemic stroke (43.4% of the whole sample) and the rest had suffered a haemorrhagic stroke (19.3% of the whole sample). In terms of time post stroke, 31.3% were 1-2 years post stroke, 37.3% were more than 2 years and 31.3% were more than 4 years after the stroke. In terms of other, apart from stroke, health problems, the respondents got a score for number of comorbid conditions ranging from 0 to 4. Of the 83 participants, 12 (14.5%) had no comorbid conditions, 22 (26.5%) had 1, 23 (27.7%) had 2, 18 (21.7%) had 3 and 8 (9.6%) had 4 or more comorbid conditions.

7.2.2 *Univariate analyses*

HRQL as measured by the SAQOL-39 was normally distributed (Kolmogorov-Smirnov test ns at $p \leq .2$) with a mean (SD) of 3.27(.7) and a median of 3.26 and scores ranging from 1.72 to 4.46. Univariate analyses were used to assess the relations between HRQL and demographic, stroke-related, health-related and other variables.

7.2.2.1 Demographic variables

Age was significantly correlated with the mean SAQOL-39 scores (r = -.27, p<.05) with increasing age resulting in poorer HRQL scores. There were no significant differences between men and women in their SAQOL-39 scores ($t(81)$ = .15, $p \leq .88$, ns). Univariate ANOVA was used to assess whether SAQOL-39 scores were significantly different between different ethnic groups (Asian, Black and White). The results were not significant ($F(2,80)$ = 1.46, $p \leq .24$, ns). The respondents were then divided into two ethnic groups: white versus all other. They were compared with independent t-test and again there was no significant difference between them ($t(81) = 1.52, p \leq .13$, ns).

A number of analyses were undertaken to explore whether marital status had an effect on HRQL in people with aphasia. First, 4 groups (married, has partner, single, divorced or spouse/partner died) were compared on their SAQOL-39 scores using ANOVA ($F(3,79)$ = .56, $p \leq .64$, ns). Then independent t-tests were used to compare those in a relationship (married, has partner) versus those not in a relationship (single, divorced or spouse/partner died) ($t(81)$ = -.18, $p \leq .85$, ns) and those married versus all other ($t(81)$ = .57, $p \leq .57$, ns). None of these comparisons yielded significant results.

The SAQOL-39 scores of the 8 SEC groups were compared using ANOVA and no significant differences were found ($F(7,75)$ = .64, $p \leq .72$, ns). As some of the groups included only a few cases, closely related groups (e.g., professionals/senior managers and associated professionals/junior managers) were combined to form 4 groups, which were compared using ANOVA. The results were not significant ($F(3,79)$ = .92, $p \leq .43$, ns). The effect of current employment status was explored by comparing the SAQOL-39 scores of 3 groups of people (retired, inactive because of the stroke and part-time/voluntary

work/students) with ANOVA. The results were not significant ($F(2,80) = 2.19, p \leq .12$, ns). The 'inactive because of the stroke' group included 11 people beyond retirement age (> 65). To assess whether the impact of inability to work was more significant for people of working age, these comparisons were repeated with only the participants of age ≤65 (n=47, retired=6, inactive because of the stroke=36, some part-time or voluntary work or students=5). The results were not significant ($F(2,44) = 1.17, p \leq .32$, ns).

In summary, the only demographic variable that was significantly associated with HRQL was age, with increased age associated with poorer HRQL. Gender, ethnic background, marital status, socioeconomic status and employment status were not significantly associated with HRQL in this group of people with aphasia. These variables were not included in further analyses.

7.2.2.2 Stroke-related and other health variables

Participants were divided in two groups according to the type of stroke they had suffered (ischaemic versus haemorrhagic). The SAQOL-39 scores of these two groups were compared with independent t-test and the results were not significant ($t(50) = -1.22, p \leq .23$, ns).

Time post onset was correlated with the SAQOL-39 scores and the results were not significant ($r = .10, p \leq .37$, ns). To explore further whether time post onset might be associated with different levels of HRQL, the participants were grouped according to their time post stroke in two different ways. First, they were divided into two groups (1-2 years post onset versus more than 2 years post onset) and their SAQOL-39 scores were compared with independent t-test ($t(81) = .80, p \leq .43$, ns). Then they were divided into 3 groups (1-2 years post onset, 2.1-4 years post onset and 4.1+ years post onset) and their SAQOL-39

scores were compared with ANOVA ($F(2,80) = .55, p \leq .58$, ns). None of these statistics was significant.

To explore whether the presence of other -apart from stroke- long-term health problems affected participants HRQL, participants were divided into two groups: those with no or 1 comorbid condition and those with 2 or more. The SAQOL-39 scores of these two groups were compared with independent *t*-tests. The results were significant ($t(81) = 2.78, p < .01$), indicating that people with more long-term health problems had poorer HRQL.

In summary, the stroke variables explored in this study (type of stroke and time post onset) were not significantly associated with the participants HRQL. Comorbidity was significantly and negatively correlated with HRQL. This variable was included in the subsequent multiple regression analysis.

7.2.2.3 Other variables

Participants' scores on the measures assessing depression/emotional distress (GHQ-12), level of activities (FAI), communication disability (ASHA-FACS), cognitive decline (RCPM), social support (SSS) and satisfaction with stroke care (PSI) were correlated with their HRQL (SAQOL-39) scores. For the GHQ-12, the scores were re-coded so that high scores were good scores (i.e., indicative of low emotional distress). The total scores were used for the FAI. There was one item in the FAI that asked about gardening and was not applicable to 30% or the respondents who did not have a garden. Only one other item had missing data in the FAI (1.2%). Missing data were imputed for each case, using the case's mean (missing data for each case ranged from 0-13.3%). The total scores were used for the PSI and the average scores for the ASHA-FACS and the SSS as recommended by the authors. The RCPM scores were converted to Standard Progressive Matrices (SPM) grades (Raven et al., 2000). The SPM

grades range from 1-5 and they represent percentile ranks. SPM grades were also re-coded so that 5 was 'intellectually superior', at or above the 95th percentile and 1 was 'intellectually impaired', at or below the 5th percentile. Descriptive statistics for these measures are presented in appendix 7.1 [13].

The results suggested that HRQL was significantly poorer in people with high emotional distress ($p<.01$), high communication disability ($p<.01$), low activity level ($p<.01$) and low cognitive level ($p<.05$) (table 7.2). High levels of social support were somewhat associated with better HRQL (the results approached significance with $p \leq .08$). All these variables were entered in the subsequent multiple regression analysis. Satisfaction with stroke care was not significantly associated with HRQL. This variable was not used in subsequent analyses.

Table 7-2: Correlations of SAQOL-39 with GHQ-12, FAI, ASHA-FACS, SPM grade, SSS and PSI.

		GHQ-12	FAI	ASHA-FACS	SPM grade	SSS	PSI
SAQOL-39	Pearson's correlation (r)	.53***	.58***	.46***	.27*	.19°	.10
	Sig. (two-tailed)	.000	.000	.000	.014	.08	.365
	N	83	83	83	82	83	82

***$p<.001$; *$p<.05$; ° $p<.1$(2-tailed)

7.2.3 Multiple regression analysis

Multiple regression analysis was performed to assess the relationship between the DV, HRQL, as expressed by the SAQOL-39 mean scores and selected IVs. The IVs were age, comorbidity, emotional distress (GHQ-12), activity level (FAI), communication disability (ASHA-FACS), cognitive level (SPM grade) and social support (SSS). Comorbidity was the only categorical variable and to enter the regression equation a dummy variable was created, where 0 was 'no or 1 comorbid condition' and 1 was '2 or more comorbid conditions'.

[13] Note that there are no distributional assumptions about IVs in multiple regression (Tabachnick & Fidell, 2001).

7.2.3.1 Multiple regression assumptions

The cases to variables ratio was tested, using the formula $n \geq 50 + 8m$. There were 7 IVs (m) and $n = 83$, which meant that the desirable cases to variables ratio was not met ($83 < 106$) in the first regression analysis.

Evaluation of the rest of the regression assumptions indicated that no transformation of variables was necessary. There were no outliers among IVs and on the DV: there were no particularly influential cases (maximum *Cook's distance* = .155, i.e., there were no values >1); the average leverage ($(m+1)/n$) was 0.09 and the maximum *centered leverage* was .275 which is below ($3(m+1)/n$) as recommended by Stevens (1992); using a $p<.001$ criterion for *Mahalanobis distance*, there were no multivariate outliers among the cases (max = 22.304 < critical χ^2 for 7df at 24.322).

There was absence of multicollinearity among IVs. In the correlation matrix of all the predictors, there were no values greater than .60 (table 7.3). All tolerance values met the set criterion and they were >.20 (appendix 7.2, under coefficients: collinearity statistics).

Inspection of the histogram of the standardised residual plots and their normal probability plots indicated that the residuals were normally distributed (appendix 7.3). Inspection of the scatterplot of the standardised residuals versus the standardised predicted values of the DV indicated that the assumptions of homoscedasticity and linearity were met (appendix 7.3). The errors of prediction were independent of one another (*Durbin-Watson* test of independence of errors = 2.09).

Cases with standardised residuals greater than an absolute value of 2 were identified to see whether there were any outliers in the solution. Only 3 cases were found with standardised

residuals greater than an absolute value of 2 (2.50, -2.25 and 2.35). The model was a good representation of the data as more than 95% (96.34%) of residuals were within ±2.

Table 7-3: Testing for multicollinearity: correlations between the IVs of the multiple regression model

		ASHA-FACS	FAI	GHQ-12	SPM grade	SSS	COMORB.	AGE
ASHA-FACS	Pearson Correlation	-	.596**	.102	.401**	-.076	.014	-.271*
	N		83	83	82	83	83	83
FAI	Pearson Correlation	.596**	-	.266*	.239*	.028	-.106	-.334**
	N	83		83	82	83	83	83
GHQ-12	Pearson Correlation	.102	.266*	-	.179	.425**	-.241*	-.269*
	N	83	83		82	83	83	83
SPM grade	Pearson Correlation	.401**	.239*	.179	-	.111	-.062	-.312**
	N	82	82	82		82	82	82
SSS	Pearson Correlation	-.076	.028	.425**	.111	-	.023	-.112
	N	83	83	83	82		83	83
COMORB.	Pearson Correlation	.014	-.106	-.241*	-.062	.023	-	.368**
	N	83	83	83	82	83		83
AGE	Pearson Correlation	-.271*	-.334**	-.269*	-.312**	-.112	.368**	-
	N	83	83	83	82	83	83	

**p<.01; *p<.05 (2-tailed)

In summary, the desirable cases to variables ratio was not met, in the first regression analysis. The assumptions of normality, linearity, homoscedasticity and independence of residuals were met. There were no outliers among the IVs and on the DV and acceptable levels of outliers in the solution. There was no multicollinearity of the IVs. Overall, no transformation of variables was needed for the multiple regression analysis.

7.2.3.2 Standard multiple regression results

Table 7.4 displays a summary of the regression model, including the adjusted R^2, the R^2 change, the unstandardised (B) and the standardised (β) regression coefficients, the t-statistics and the probability levels. The overall model accounted for 51% (adjusted R^2=.51) of the variance in the SAQOL-39 scores. R for regression was significantly different from zero, with $F(7,74) = 13.260, p < .001$.

Table 7-4: Summary of standard multiple regression analysis of the relation of HRQL with correlated predictors.

Predictors	Adjusted R^2	R^2 Change	B	β	t
(Constant)			.63		1.09
ASHA-FACS			.18	.22	2.15*
FAI			2.531E-02	.36	3.52**
GHQ-12	.51***	.56***	7.823E-02	.35	3.81**
SPM grade			3.430E-02	.04	.51
SSS			4.563E-02	.06	.71
Comorbidity			-.30	-.21	-2.48*
Age			4.869E-03	.11	1.17

Dependent Variable: SAQOL-39 mean
***p<.001; **p<.01; *p<.05

Inspection of the B coefficients showed that emotional distress (GHQ-12) ($t(74) = 3.81$, p<.001), activity level (FAI) ($t(74) = 3.52$, p=.001), communication disability (ASHA-FACS) ($t(74) = 2.147$, p<.05) and comorbidity ($t(74) = -2.48$, p<.05) were all significant predictors of HRQL (SAQOL-39). The β coefficients allow for the direct comparison of the predictors and they indicated that the most important predictors were activity level and distress, followed by communication disability and comorbidity. Three variables, cognition (SPM grade), social support (SSS) and age, were not significant predictors. Inspection of the 95% confidence intervals for the IVs showed that for these three variables the confidence intervals included zero (appendix 7.2, under coefficients: 95% confidence intervals for B).

This is further evidence that these three variables weaken the overall model, as in some samples they have a negative relationship with HRQL and in others they have a positive relationship. For example, low cognitive level was associated with good HRQL in some cases and poor HRQL in others.

A second regression analysis was run including only the significant predictors (i.e., emotional distress, activity level, communication disability and comorbidity). This model was stronger as all the assumptions were met including the recommended cases-to-variables ratio where $n \geq 50 + 8m$, $n \geq 50+(8 \cdot 4)$, $n \geq 82$ and here $n=83$. Details of the regression assumptions for the second regression analysis are presented in appendix 7.4.

This model accounted for 52% (adjusted $R^2=.52$) of the variance in the SAQOL-39 scores. R for regression was significantly different from zero, with $F(4,78) = 23.37$, $p < .001$. B coefficients showed that emotional distress ($t(78) = 4.62$, $p<.001$), activity level ($t(78) = 3.40$, $p=.001$), communication disability ($t(78) = 2.29$, $p<.05$) and comorbidity ($t(78) = -2.18$, $p<.05$) were all significant predictors of HRQL (table 7.5).

Table 7-5: Summary of 2nd standard multiple regression analysis of the relation of HRQL with correlated predictors.

Predictors	Adjusted R^2	R^2 Change	B	β	t
(Constant)			1.17		2.89**
ASHA-FACS			.17	.22	2.29*
FAI	.52***	.54***	2.38E-02	.34	3.40***
GHQ-12			8.30E-02	.38	4.62***
Comorbidity			-.24	-.17	-2.18*

Dependent Variable: SAQOL-39 mean
***$p \leq .001$; **$p<.01$; *$p<.05$

The semipartial correlations of the IVs (appendix 7.5, under coefficients: correlation) were squared ($.35^2+.26^2+.18^2+.17^2$) and deducted from the R^2 (.54) to estimate the unique and

shared variance of the IVs. The unique variance of the IVs was .12 for emotional distress, .07 for activity level, .03 for communication disability and .03 for comorbidity. For all four of them it was .25. All together they contributed another (.54 - .25) .29 in shared variability.

In summary, high emotional distress, low activity level, high communication disability and high comorbidity were significant predictors of poorer HRQL. These variables accounted for 52% (adjusted) of the variance of the SAQOL-39.

7.3 Summary

Multiple regression analysis was undertaken to explore what were the main predictors of HRQL in people with chronic aphasia following stroke. The potential IVs were: a) demographic variables: age, gender, ethnic background, marital status, socioeconomic status, employment status; b) stroke and health related variables: type of stroke, time post onset of stroke, and number of other long-term health conditions (comorbidity); and c) other variables: psychological distress, level of activity, severity of aphasia, cognition, social support and satisfaction with services received for stroke. Univariate analyses were undertaken to determine which IVs would enter the regression model. The assumptions of regression were tested and the standard multiple regression method was used. The regression analysis was run twice, the second run including only the significant predictors of the first run. In the second model all the set assumptions of multiple regression were met. The significant predictors of HRQL in people with chronic aphasia following stroke were emotional distress, activity level, communication disability and comorbidity. These predictors accounted for 52% (adjusted) of the variance of the HRQL scores.

8 DISCUSSION

This chapter begins with discussing the overall strengths and limitations of the study. Then the methodology and the results of the two research questions addressed in this study, namely a) the development of the SAQOL and the psychometric evaluation of the three versions of the instrument and b) the assessment of the predictors of HRQL in people with chronic aphasia after stroke, are discussed. Suggestions for future research are also discussed in the context of each research question and the chapter closes with the overall conclusions of this study.

8.1 Study strengths

8.1.1 Conceptual clarity

HRQL is a concept that has attracted a lot of attention in recent years. Still, an overall accepted definition is elusive. For the purposes of health care evaluation, HRQL can be seen as reflecting *the impact of a health state on a person's ability to lead a fulfilling life* (Bullinger et al., 1993) and as incorporating *the person's subjective evaluation of his/her physical, mental/emotional, family and social functioning* (Berzon et al., 1993; Hays et al. 1993; de Haan et al. 1993). A main strength of this definition is that it incorporates the areas identified by people with long-term disabilities as most affected by their illnesses, namely physical/functional and social aspects (including performing roles) of life (Bowling, 1995b). It is endorsed by the International Society for Quality of Life Research (ISOQOL). This definition was followed in the current study. In reviewing the relevant literature, differences in the conceptualisation of HRQL were highlighted wherever possible.

Still, it is acknowledged that more research is needed in order to refine, define and understand the concept of HRQL better. Hunt (1997) argues for pure research in this area. She also argues for a second type of research, which focuses on eliciting the patients' views on their medical treatments. She suggests that a very clear distinction should be maintained between these two undertakings.

8.1.2 Recruitment

A potential source of error in survey research is non-response bias: if the respondents differ from the non-respondents in an important or systematic way, then the sample may not be representative of the population targeted (Singleton & Straits, 1999). Several steps were undertaken, as recommended (Fowler, 1993; Singleton & Straits, 1999), to increase response rates and avoid this bias.

During recruiting and giving information on the project to eligible participants, personal contact allowed the interviewer to enlist the participants' cooperation more effectively. Initial contact involved sending personalised letters to people and following-up with as many phone calls as necessary, in order to give preliminary information on the project and arrange a time to meet. In the first face-to-face contact the interviewer had the chance to check participants' understanding and readily answer their questions or concerns. Every effort was made to make it clear to participants that their help was important and to explain how it would be useful; to ensure confidentiality; and to clarify that the participants would not be threatened by the tasks or the uses to which the data would be put (Fowler, 1993). To reduce non-response resulting from lack of availability, the interviewer had a flexible schedule and visited participants at times that were convenient to them.

All this resulted in a high response rate (82%). There is no agreed-upon standard for a minimum acceptable response rate. Fowler (1993) acknowledges that even well conducted surveys often have a response rate of 60-75% and quotes a standard set by the Office of Management and Budget (which reviews surveys done under contract to the USA government) of a minimum of 75%. Singleton & Straits (1999) recommend a 70% response rate as the minimally acceptable for interview surveys. This study's high response rate was well above these figures and should have at least partly offset the potential for non-response bias.

8.1.3 Mode of administration

All data collection was done through face-to-face interviewing. This had several advantages in particular in relation to the population under study. Face-to-face interviewing is particularly recommended with respondents who may have difficulty understanding the items of the questionnaires used (Streiner & Norman, 1995), as people with aphasia may have. The clinical training of the interviewer as a SLT ensured that she could facilitate the respondents' understanding of the items and expressing of their responses.

This mode of administration also eliminated another source of systematic non-response bias during data collection. In surveys, and postal surveys in particular, highly educated respondents are more likely to cooperate than poorly educated ones and the case is similar for people who have strong views about the issue under study (Singleton & Straits, 1999). In this study all participants who agreed to take part and were able to self-report were interviewed and completed all the assessments in a uniform manner.

Face-to-face interviewing also eliminated withdrawal bias. This arises when participants who withdraw from a study differ systematically from those who remain (Sackett, 1979). It can

occur in postal surveys, for example, if some people fail to return a follow-up questionnaire. In this study, all respondents were visited as required to complete all the assessments and no respondents withdrew.

The interviewer could also make sure that the respondents did not miss answering some items. This may happen by mistake or because the respondent is not sure what the items require or for other reasons, like boredom and tiredness. The interviewer made sure all the items were asked and clarified what was required wherever necessary. She also made sure that the process was not too long or too tiring for the respondents and that they got breaks when they needed them or an extra visit.

Other practical benefits of face-to-face interviewing included the following. The interviewer made sure that the person who actually responded was the targeted participant. She skipped questions that did not need to be asked to all participants with ease (e.g., items m4 and m5 in the SAQOL or some questions in the PSI) and without the confusion that 'skip patterns' often cause in self-administered questionnaires (Streiner & Norman, 1995). She also made sure that all participants did all the questionnaires and all the questions within them in the same order, which ensured a uniform administration. Thus, it could be assumed that differences in answers could be attributed to differences in the respondents rather than to differences in the stimuli (e.g., the order of the questions, misreading the questions etc.) (Fowler, 1993).

8.1.4 Accessibility of materials

One of the main problems in applying the results of stroke studies to the subgroup of stroke people with aphasia is that people with aphasia were treated and assessed differently from the other respondents in these studies. As it has been described in chapter 2, in some studies

some people with aphasia were excluded (those with severe aphasia), or all of them were excluded or proxy respondents were used for some of them (those with severe aphasia). From a different perspective, HRQL is a highly subjective concept and as such, the people most qualified to report on it are the people themselves who are being assessed.

These two reasons necessitate assessing the HRQL of people with aphasia with measures that are accessible to them. This approach was followed in this study. The chosen measures were as far as possible linguistically straightforward and could be interview administered. In addition, they did not have a standardised administration. This way their presentation could be modified, without changing their content, in order to make them more communicatively accessible to people with aphasia.

Making all material used accessible to people with aphasia in combination with interviewing them in order to facilitate them further resulted in the collection of a rich pool of data from people with aphasia. This study, to the best of our knowledge, is the largest study of the HRQL of people with aphasia based on self-report data.

8.1.5 Dealing with sensitive issues

During the interviews with the respondents some sensitive issues were raised. People discussed their current concerns and fears, e.g., fears of having another stroke or concerns about getting worse after completing therapies. Some people indicated they were feeling depressed or increasingly worried and anxious or socially isolated. Some felt they had nobody to talk to or to help them with these feelings. Some complained about lack of treatment or follow-up for their disabilities.

It is essential for researchers who interview people with long-term disabilities to have experience in dealing with sensitive issues. In this study, the interviewer's background as a

therapist facilitated her in dealing with sensitivity and compassion with the issues raised by the respondents rather than treating them as just a source of information for the study. Information on settings that could provide potentially useful services for them were made available to participants. They were offered the option of either the researcher referring them to a selected service or contacting the service they wished directly themselves. Wherever necessary, the respondents were asked whether they wished their GP or other professionals involved in their care to be informed of their concerns. This approach resulted in about 10 participants being put in contact with people/services/schemes (e.g., self-help groups, counselling for people with aphasia, conversation partners, GP's and other health professionals) that could help them with a previously unidentified problem.

8.2 Study limitations

8.2.1 Selection bias

Selection bias refers to error due to systematic differences in characteristics between those who are selected for study and those who are not (Sackett, 1979). In the pilot and pre-test parts of this study, during the development of the SAQOL, convenience sampling was undertaken rather than random sampling. In convenience sampling, the researcher selects a requisite number of participants from cases that are readily available (Singleton & Straits, 1999). This method of sampling limits the generalisability of the findings to the group studied. This method was adopted for practical reasons, i.e., because it was quick and easy and offered immediate access to respondents. Singleton & Straits (1999) suggest that if the research is at an early stage and generalisability is not an issue, then convenience sampling may be perfectly acceptable.

Probability sampling was used during the psychometric evaluation of the SAQOL versions and the assessment of the predictors of HRQL in people with aphasia. Probability sampling

is based on a process of random selection, which gives each case in the population an equal chance of being included in the sample. With this type of sampling, one knows to which population the sample may be generalised as well as the limits of generalisability (Singleton & Straits, 1999).

With regard to the specific type of probability sampling, cluster sampling was used, in which clusters of cases were initially identified (community SLT services and Connect) and then within these clusters all identified eligible participants were asked to take part. This type of sampling was dictated by the fact that there are no lists of all people with chronic aphasia available (in order to use simple or stratified random sampling). Even available stroke registers may not include accurate information for people with aphasia. An additional ethical consideration is that they tend to be over-researched[14]. Thus, the most feasible way to identify people in this population was to recruit from settings that provide services for people with chronic aphasia.

A selection bias here may arise from the fact that people with chronic aphasia who have never received or no longer receive SLT are not reached. To address this issue at least partly, people from the discharged lists (last 6 months) of the approached settings were also included in this study. Still it is acknowledged, that if this subgroup of people with aphasia have different characteristics from the people with aphasia who are referred to SLT, then selection bias would arise.

An alternative sampling technique to capture people with chronic aphasia in the population would be to recruit through GP's. This method has two main disadvantages. First, due to the low prevalence of aphasia it would have been necessary to recruit from numerous GP's in

[14] Personal communication with Charles Wolfe, developer of the South London research register

order to get the number of people with aphasia required for this study. Given the resources and the time scale of this study, this was not feasible. Second, even if it were feasible, GP's do not have specialist knowledge of aphasia and therefore they may not identify some people with aphasia, especially those with mild expressive problems. This method, therefore, is also susceptible to selection bias.

8.2.2 Generalisability of results

The population under study was people with chronic aphasia following stroke. To judge the generalisability of the findings of this study, the respondents' characteristics are compared to available information on the characteristics of the stroke population in England, as there is no available information in particular for people with aphasia.

Stroke in England has a higher prevalence in people from manual social classes (Department of Health, 1998). Participants in this study were recruited from sites in the Southeast of England. The South of England is generally considered to be more affluent than the North (Office of National Statistics, 2000). In this study, there was a tendency for respondents to come from the higher socio-economic classes (57% from non-manual social classes and 35% managers or professionals). This probably reflects the geographical area from which they were drawn. Still, it limits the applicability of the results to the overall population of people with chronic aphasia following stroke in England.

Stroke is more common in men (Department of Health, 1998). In this sample about 63% of the participants were male. In South London, 24% of the population is Black or Asian (Stewart et al., 1999) and in this sample about 22% were Black or Asian. Stroke is also more common in older people and in this sample almost 44% of the participants were over 65. The incidence of stroke is much lower in younger people with only 6.3% of all strokes being

below the age of 45 in South London (Stewart et al., 1999). In this study, it was aimed to recruit more young people in order to have enough numbers to explore the effect of age as a predictor of HRQL. As a result about 15% were at or below 45.

In summary, the participants were varied in terms of age, gender and ethnic background. The results of this study should generalise well to the population of people with chronic aphasia living in the Southeast of England. The high representation of higher socio-economic classes and the potential selection bias of missing the hard to reach subjects mean that the generalisability of the results to the overall population of people with chronic aphasia in England can be questioned.

8.2.3 *Interviewer bias, social desirability and 'faking good'*

Despite its numerous advantages, interview administration has the potential to lead to interviewer bias. Interviewer bias refers to systematic differences in soliciting, recording and interpreting information from subjects (Hennekens & Buring, 1987). Systematic differences in recording and interpreting information are likely to occur due to differences between different interviewers (inter-interviewer bias). In this study, one interviewer carried out all the interviews and therefore inter-interviewer bias was eliminated. However, a potential source of interviewer bias may have been the interviewer's characteristics, which could have lead to systematic differences in the soliciting of information from the respondents.

Streiner & Norman (1995) refer to a number of studies that have shown that race differences between the interviewer and the interviewee can affect the latter's responses. The quoted studies however addressed political issues in which race was a factor (e.g., preference for a black or white political candidate). In this study, race was not an issue in relation to the

concept under study, so it is not known whether and how it could have affected the interviewees.

The interviewer's gender could also have had an effect. The responses women elicit may be different from those given to male interviewers, especially when sexual issues are being discussed (Hyman et al., 1954). Only one question in the current study addressed a sexual issue (item sr6: 'did you have sex less often than you would like?', in the SAQOL). Age could also be a factor although age differences between the interviewer and the interviewee have not been extensively studied (Streiner & Norman, 1995).

To minimise the effects of interviewer bias, the interviewer tried to develop a rapport with participants to make them feel relaxed and able to speak freely. She also acknowledged responses without making any judgements and tried to show an open and reflective mind. This approach may have helped in getting the respondents' trust and may have led to honest responses.

Social desirability concerns the unintentional tendency to report positive answers, and 'faking good' concerns the intentional creating of a false positive impression (Streiner & Norman, 1995). A standard way of measuring whether social desirability has affected responses to a questionnaire is the simultaneous administration of a social desirability scale (e.g., the Crowne-Marlowe scale, Crowne & Marlowe, 1960). Such a scale was not administered in this study to avoid increasing respondent burden, as a number of scales were already being administered.

Social desirability and 'faking good' may have influenced responses particularly as all participants knew that the researcher was a SLT. This information was given to them at the beginning of the project to reassure them that the researcher had the skills to communicate

with people with communication difficulties and would try to facilitate them if they had difficulties understanding or expressing something. It is possible that the participants, knowing the researcher's background as a therapist, focused (intentionally or unintentionally) on positive aspects in general and of the care they had received in particular (in the PSI).

To minimise the potential effect of this bias, the researcher assured the participants that all information provided was confidential and that taking part in the study would not influence the care they received. She emphasised the importance of getting people's true opinions about their care and about how they felt they were doing. Moreover, the participants took part in this research willingly. The sincerity in wanting to help was often shown by people indicating that they were doing this project to help other people with similar problems (e.g., by pointing with emphasis to the relevant picture in the project information booklet). This willingness to participate suggests that information provided was of good quality and not intentionally biased.

8.3 Development of the SAQOL and psychometric evaluation of the SAQOL versions

A stroke specific scale, the SS-QOL, rather than a generic measure was chosen for adaptation for people with aphasia. This decision was based on the reported higher validity of disease-specific scales (de Haan et al., 1993) and higher responsiveness to change (Patrick & Deyo, 1989) compared to generic measures.

Recent research has offered extra evidence to support this choice. Hamedani et al. (2001) developed a measure (HSQuale) for the assessment of HRQL in young people with haemorrhagic strokes. They found it had better content validity and discriminated better between patients than a generic measure (the SF-36, which had ceiling effects in 5 of its 8

domains). Williams et al. (1999b) asked people after ischaemic stroke to rate their HRQL as better or worse than before the stroke. They then compared the scores of the two groups on the SS-QOL and the SF-36 and found that the SS-QOL was significantly associated with the patient reported HRQL whereas the SF-36 was not. These results suggested that generic measures may be less sensitive to meaningful changes in HRQL after stroke than stroke-specific measures. This may suggest that their domains or their items lack relevance to people with stroke.

The next sections discuss the strengths and limitations of the methods used in developing the aphasia-adapted version of the SS-QOL and in testing its psychometric properties; and the strengths, limitations and implications of the findings of this process.

8.3.1 Methodology: development of the SAQOL

The development of the aphasia adapted version of the SS-QOL, i.e., the SAQOL, involved two stages. In the first stage, the instrument was modified to make it communicatively accessible to people with mild to moderate receptive aphasia and to increase its content validity with this population group. This stage involved consultations with professionals with expertise in aphasia and questionnaire development and a pilot testing with people with aphasia. In the second stage, the modified version was pre-tested in a group of people with aphasia to evaluate its content validity, accessibility and acceptability and to identify whether any revisions were necessary before testing its psychometric properties further in a large sample.

The main strength of the methodology used in the development of the SAQOL was the involvement of people with aphasia. During the pilot testing the opinions of the people with aphasia who took part determined, to a large extent, the response format that replaced the

'strongly agree-strongly disagree' response format. They also assisted in identifying ways of making the scale more user-friendly for them, by e.g., converting statements into questions, rephrasing negative items and dropping the grid format of the responses.

People with aphasia were also interviewed in the pre-testing of the instrument. Although the involvement of users in the pre-testing of new instruments is crucial to determine the acceptability of the questionnaire in terms of content and phrasing (Fowler, 1993), many researchers either do not pre-test or do not report their pre-test findings. Pre-testing the SAQOL revealed what difficulties people with aphasia had and how the administration and the presentation of the instrument had to be modified to facilitate them (e.g., give more clear instructions for the 'yes-no' response format, add anchor points ✓ with 'definitely no' and a ✗ with 'definitely yes' on the presenter's form).

Another strength was the involvement of SLT's in the process of identifying what aspects of the SS-QOL would need to be modified and how. SLT's were primarily involved due to their specialist knowledge of people with aphasia and what they may find hard and how they can be facilitated. Their opinions of the instrument, however, were also important for another reason. SLT's may be among the primary users of the instrument for assessing HRQL in people with aphasia in clinical practice. Their involvement in this study gave the opportunity to gauge what their first reactions to the instrument were. Most of them expressed a need for measures to assess HRQL in people with aphasia and felt the SS-QOL had the potential to become an accessible and useful measure for people with aphasia.

A limitation of the methodology used to develop the SAQOL was the sampling used. As has been already indicated, convenience sampling was used, which limits the applicability of the results beyond the sample used. Another limitation with the same outcome is the small

number of respondents both in the pilot and in the pre-test studies (12 and 18 people respectively). Still, it is generally thought that small numbers are acceptable in pilot and pre-test studies as long as the participants come from the population under study (Fowler, 1993; Singleton & Straits, 1999).

8.3.2 Methodology: Psychometric evaluation of the SAQOL versions

The psychometric evaluation of the SAQOL involved further testing of its acceptability and testing of its reliability and construct validity. Response rates, percentage of missing data and the distribution of scores across response categories were calculated to test the acceptability of the SAQOL. Reliability testing comprised the assessment of the internal consistency and test-retest reliability of the scale. Within scale analyses and comparisons with external measures were used in the construct validation of the scale. The within scale analyses were the assessment of the internal consistency of the scale, the inspection of the intercorrelations between the scale's subdomains and the subdomains and the scale's corrected mean, and factor analysis (FA). FA was used to test whether the original 12-subdomain conceptual model of the SS-QOL held up in the SAQOL data and to derive the best factor model to describe the data. This process as anticipated resulted in alternative versions of the SAQOL that were assessed with all the methods described here for their psychometric properties. Validation of the SAQOL subdomains consisted of comparisons with external measures to test their convergent validity, their correlations with related variables and their discriminant validity.

The main strength of this methodology is that psychometrically sound techniques were used and the properties of the instrument were tested against rigorous scientific criteria. In terms of acceptability, the recommended criteria of missing data < 10%, AEF<10% and MEF<80% were followed. In reliability, the criteria of Cronbach's alpha >.70, item-total

correlation >.30 and ICC's>.75 were used to ensure the good internal consistency and test-retest reliability of the measure.

In terms of validity, the within scales analyses used on the whole scale validation followed the criteria of internal consistency (Cronbach's alpha) >.70 and intercorrelations between subdomains and the corrected total mean in the range of .30-.80. Meeting these criteria is commonly seen as evidence of the homogeneity of the scale (Nunnally & Bernstein, 1994).

In the factor analyses that were performed the factorability of the data set was tested with Keiser-Meyer-Olkin (KMO) test of sampling adequacy and the Bartlett's test of sphericity. Commonly the studies using FA do not report on these statistics and therefore it is not known whether the recommended criteria were met. The KMO test, when more than at least .50, suggests that the associations between the variables in the correlation matrix can be accounted for by a smaller set of factors. The Bartlett's test of sphericity suggests, when significant, that there are discoverable relationships in the data (Ferguson & Cox, 1993). If these criteria are not met then the derived factor model will not be psychometrically sound. A factor model is also more robust when the factor loadings are ≥.40, when there are at least 3 items per factor and when the crossloaders are eliminated (Ferguson & Cox, 1993; Tabachnick & Fidell, 2001). All these criteria were followed in this study.

The FA performed also had another advantage. Different strategies were used in order to derive the best factor model that would describe the data in our sample of people with aphasia. The first strategy was more bound to theory and the operational definition of HRQL followed and, therefore, it commenced with all the items. One approach (top-down) particularly sought to test whether the original conceptual model of 12-subdomains would hold up. Factor analyses were therefore performed within the subdomains. The other

approach (bottom-up) was more data driven and, therefore, factor analyses were performed with the items and not within the subdomains. The other strategy was the least conceptually driven in that it commenced with only the best, in terms of their psychometric properties, items. Following these alternative approaches allowed for simultaneous comparisons of different factor models both in terms of conceptual basis and psychometric properties. This would allow the best factor model in terms of conceptual basis, psychometric properties and describing the data to be derived.

In the testing of the validity of the measure against external measures, this study was particularly strong in setting clear hypotheses about the expected correlations of the SAQOL scale and its subdomains and external measures. These hypotheses were based both on related theory and published research in the area. Then evidence was gathered to test these hypotheses. The more evidence that supports the hypothesized relationships, the greater one's confidence that a particular operational definition (as reflected in an instrument) is a valid measure of the concept (Singleton & Straits, 1999).

A challenge in the validation of the SAQOL against external criteria was related to the external measures used. An assumption operating is that external measures are valid with the population under study. Apart from the FAST, the ASHA-FACS and the FAI, the rest of the measures used have not been validated specifically with people with aphasia. This is seen as an unavoidable challenge, as there are no validated measures for people with aphasia in the targeted domains.

Every effort was made to choose appropriate measures. The RCPM and the GHQ-12 have been extensively used in stroke research and the SSS was standardised on a group of chronic disease outpatients. In addition, none of these measures required expressive language (the

respondents could point to the response of their choice) thus expressive aphasia was not expected to considerably affect administration of the measures. Lastly, as has been indicated, all the measures were administered by a SLT who made every effort to facilitate the understanding of people with aphasia. These measures were, therefore, seen as valid as is possible currently with people with aphasia.

A limitation in the psychometric evaluation of the SAQOL was the sample size. The sample size was adequate for the psychometric analyses performed. Even for the FA, Tabachnick & Fidell (2001) point out that "if there are strong correlations and a few distinct factors, a sample size of 50 may be adequate, as long as there are notably more cases than factors". In the data of this study, there were 83 cases for 4 factors. Still, a sample of more than 300 cases would have allowed more state-of-the-art techniques to be used such as tests of scaling assumptions (using the multi-trait/ multi-item analysis program-revised, MAP-R) or exploration of whether weighting of the items would improve the scale or not (Ware et al., 1997). Some instrument developers have devised weights for each item relative to their contribution to the total score. However, Streiner & Norman (1995) indicate that differential weighting of items is rarely worthwhile and suggest that "when there are at least 40 items in a scale, differential weighting contributes relatively little, except added complexity for the scorer. With fewer than 40 items (20, according to Nunnally, 1970), weighting *may* have some effect". Still, these are identified as areas that could be addressed in further research when testing of the SAQOL versions with larger samples.

8.3.3 Main findings: Development of the SAQOL and psychometric evaluation of the SAQOL versions

The findings are discussed in terms of the properties tested i.e., accessibility, acceptability, reliability and validity.

8.3.3.1 Accessibility

The pre-test of the study indicated that the SAQOL was highly accessible to people with any severity of expressive aphasia and moderate or mild receptive aphasia. Accessibility was judged by taking into account the opinion of people with aphasia, in a semi-structured interview, and by determining the FAST receptive score beyond which the respondents were able to reliably respond to two consecutive administrations of the instrument (2-7 days apart). All participants who found the instrument straightforward (17 out of the 18) had a score of 7/15 in the FAST and responded reliably to the two administrations of the SAQOL. The accessibility of the SAQOL was further supported in the main survey study, where all participants with a score of 7/15 in the FAST self-reported on the SAQOL.

Determining the level of aphasia severity beyond which people with aphasia may not be able to self-report is seen as an important outcome of this study. None of the studies reviewed reported any clear criteria on how they judged which people with aphasia could self-report on their HRQL and which could not. This may have led to the exclusion of people with perfectly adequate communication skills to self-report on their HRQL. The FAST is a practical and very quick to administer measure (the receptive domains take 2-5 minutes) that can prove a valuable screening tool for clinicians and researchers prior to testing the HRQL of people with aphasia with any version of the SAQOL.

The two items of the SAQOL that were identified as hard by more than one participant were md3 and fr5. These items were retained, firstly, because only two people had difficulty with them and, secondly, because it was thought that the altered administration of the instrument (e.g., explaining better the relevant response format) could facilitate the respondents in answering these items more easily. During the administration of the SAQOL in the survey study, the items were further tested for their psychometric properties in terms of how they

worked as part of their intended subdomain and the overall scale. This resulted in the fr5 being dropped from the SAQOL-39 as in the factor analysis it was found to contribute considerably to more than one underlying subdomain (crossloaded on two factors).

8.3.3.2 Acceptability

In the pre-test study, people with aphasia found the SAQOL highly acceptable. Their behaviour when considering the items indicated they needed prompting with answering two items (ue1 and t4). The administration of these items was modified to facilitate them. In the survey study, the acceptability of the SAQOL was tested psychometrically and it was found that 21% of the items were affected by skewness and 10 items by AEF/MEF. These items would not discriminate well between respondents, thus affecting the potential usefulness of the measure. Most of these items were removed in the shorter versions of the SAQOL and as a result the SAQOL-39 and the SAQOL-34 had high acceptability.

8.3.3.3 Reliability

The SAQOL and its subdomains had good test-retest reliability (ICCs ranging from .85-.99). The whole scale had good internal consistency (Cronbach's alpha:.93), but 4 of its subdomains had poor internal consistency (Cronbach's alpha:<.70). These results suggested that some items measured something different from the intended concept of their subdomain and questioned the 12-subdomain structure of the SAQOL

The SAQOL-39 and the SAQOL-34 had very good test-retest reliability (ICCs ranging from .89-.99). They also had very good scale (Cronbach's alpha:≥.92) and subdomains (Cronbach's alpha in all: >.73) internal consistency. These results suggested that these two scales and their subdomains were homogenous. All their items tapped different aspects of the same attribute and not different parts of different traits (Streiner & Norman, 1995).

8.3.3.4 Validity

In terms of content validity, four items were added during the adaptation of the measure for people with aphasia. These items were tested in the survey study and proved to be useful additions as 3 of them remained in the shorter and psychometrically more sound versions of the SAQOL. In the pre-test, the majority of the respondents thought the measure covered the effects of stroke and aphasia on their lives. When asked, a few respondents made suggestions about items that could be added but each one emphasised a different area. No further items were therefore added and the measure was thought to have good overall content validity.

The construct validity of the measure was tested in the survey study with a variety of methods. Overall, the results supported the construct validity of the overall SAQOL scale, but not of its subdomain structure. Three of the 12 subdomains (thinking, personality, social roles) had poor overall external measures validity. FA indicated that not all of the items contributed considerably to the overall score and that no stable, conceptually clear factor structure could be derived from all of the 53 items of the SAQOL. The two shorter versions derived through FA, the SAQOL-39 and the SAQOL-34, had a stable, conceptually clear 4-factor structure and high scale and subdomain construct validity. These results have been largely discussed in chapter 6 in order to explain the choice of the SAQOL-39 in the further analyses. This discussion explicates this choice a bit further and then concentrates on the underlying domain structure of the two instruments.

As has been already indicated, the SAQOL-39 was identified as the best overall measure for the assessment of HRQL in this group of people with chronic aphasia. It was preferred to the SAQOL as it had a clear subdomain structure and the extra advantage of being considerably shorter (14 items shorter). The main disdvantage of the SAQOL was that,

although it worked as an overall scale, its items could not be grouped into meaningful, homogenous subdomains, distinct from one another.

The SAQOL-39 was preferred to the SAQOL-34 as it was somewhat broader conceptually without being considerably longer. This could be due to the fact that the SAQOL-39 was more conceptually driven than the SAQOL-34. It retained as many items as were possible without compromising its psychometric quality. Moreover, it retained more items on social activities (sr1, sr4 and sr5 on going out, doing hobbies and seeing friends less), an area identified as important for HRQL by people with aphasia (Cruice et al., 2001) and considerably affected in people with aphasia (Parr et al., 1997; LeDorze & Brassard, 1995) The SAQOL-34 might have been the preferred version if the instrument were tested in a very large (>300 participants) and with no doubt representative sample of the targeted population. If these were the conditions, then there would be increased confidence that the items removed in the item reduction were indeed items that were not relevant to, or did not discriminate well in the whole population. As these were not the conditions of this study the SAQOL-39 was the preferred version.

The subdomains of the SAQOL-39 were the following: physical, psychosocial, communication and energy. These domains have been repeatedly identified by people after stroke as among the most affected by their stroke (Williams et al., 1999a; Duncan et al., 1999; Buck et al., 2000; Buck et al., 2001; Cruice et al., 2001). Although all the domains of the accepted definition of HRQL are reflected in the SAQOL-39, not all of them stood out as individual domains (social and emotional combined to psychosocial). Moreover the SAQOL-39 had the distinct domains of communication and energy. It is common for disease-specific measures to include domains that are particular to the disease under question, for example the Arthritis Impact Measurement Scales (Meenan, 1982) have domains on dexterity and

pain; the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire (Aaronson et al., 1993) has symptom scales on nausea and vomiting and pain; the Parkinson's Disease Questionnaire (Andreu et al., 2000) has domains on stigma and bodily discomfort. This is seen as reflecting the increased relevance of the measure to the population under study.

In summary, the development and psychometric evaluation of the SAQOL showed the overall measure to be an accessible, acceptable, reliable (in terms of internal consistency and test-retest reliability) and valid measure for the assessment of HRQL in people with chronic aphasia after stroke. The SAQOL, however, did not have a clear subdomain structure. HRQL is a multifaceted concept and the instruments measuring it should reflect its different underlying domains. A clear subdomain structure increases the usability of a scale as it provides more specific information. It gives the flexibility to the users to look both at overall HRQL and at specific aspects of it, for example in this case, psychosocial aspects or communication. The SAQOL-39 is a shorter version of the SAQOL derived through FA. It has a clear 4-subdomain structure and high accessibility, acceptability, internal consistency, test-retest reliability and construct validity with people with chronic aphasia after stroke.

8.3.4 Future research

In this research, the SAQOL-39 was the preferred measure for the assessment of HRQL in people with chronic aphasia. People with chronic aphasia are a subgroup of people with aphasia, which is a subgroup of people with stroke. The SAQOL is an adaptation of a stroke specific scale that can be used also with people with aphasia. As such it may be a useful measure to be used with an overall stroke population. Further research in different populations of stroke survivors including both people with and without aphasia would evaluate the psychometric properties of the SAQOL in a broader population. Already, there

has been interest on the aphasia-adapted version of the SS-QOL from researchers around the world, who are looking for a stroke specific scale that can be used with people with aphasia. For people working with varied stroke populations the SAQOL is recommended for language and cultural adaptation and further psychometric testing. The SAQOL-39 can also be tested further in such populations. It is particularly recommended for the subgroup of people with chronic aphasia. Only further psychometric testing will show whether indeed existence of both versions of the scale is justified.

A limitation in the psychometric evaluation of the SAQOL-39 relates to its administration. What was actually administered to the participants in this study was the SAQOL and the SAQOL-39 was derived in the FA of the SAQOL. This means that the item reduction and the psychometric testing of the instrument were done in one stage. Whilst this is a common way of developing outcome measures it is not the preferred way. Ideally, a measure should be subjected to one field test for item reduction and preliminary psychometric evaluation and then to a second field test for extended psychometric evaluation of the item-reduced version. This is because results generated from modified instruments are not directly comparable to results generated from the source instrument (Fletcher, 1995; Erickson, 2000). For scientific integrity, the psychometric properties of the SAQOL-39 should be tested again in an independent sample.

Further psychometric evaluation of the SAQOL/SAQOL-39 in a large independent sample could allow standard scores to be developed that would allow comparisons between the different SAQOL/SAQOL-39 subdomains and also comparisons with other standardised scales. Standard scores also show how each person is doing compared to everybody else (Streiner & Norman, 1995). Further psychometric evaluation should also incorporate responsiveness to change. This is the degree to which an instrument is able to detect clinically

significant change over time. A health outcome measure must be able to detect small but clinically important differences in outcome which clinicians and patients regard as important (Deyo et al., 1991). In terms of clinical trials, highly responsive scales are preferred because they allow clinical trials to be performed with smaller samples (Wright & Young, 1997). Preliminary evidence from this study on the responsiveness of the SAQOL-39 include the absence of floor and ceiling effects, the exclusion of items that did not differentiate well between respondents and the exclusion of a whole area (vision) that probably remained static in the population under study (Fitzpatrick et al., 1992). Still, further testing is needed if the instrument is to be used in clinical trials or in the evaluation in routine clinical practice.

So far the measure has been used in research as a research tool and not as a clinical outcome measure. Further research with the SAQOL-39 should evaluate its appropriateness as an outcome measure in people with chronic aphasia undergoing rehabilitation programmes. At present, the SAQOL-39 is being used in a study to evaluate the services provided to people with aphasia at Connect-the communication disability network. This study may provide useful information on the appropriateness, the applicability and responsiveness of the measure, as an outcome measure.

In this study, data were collected from proxy respondents for people with severe receptive aphasia. These data are from a limited sample (12 people) but their analysis will provide useful information and help generate further ideas on how to assess the HRQL in people with severe receptive aphasia. Descriptive statistics from these data can be compared with the descriptive statistics of the data presented here, to see whether similar patterns emerge or not. Further research can also look at proxy and self-report agreement on the SAQOL/SAQOL-39 in people with less severe aphasia. This could help explore what disagreement there is between proxy and self-report, whether there is a clear direction in the

disagreement (e.g., proxy always overrates) and whether there is any justification in adjusting the proxy data so that they are closer with the self-report data.

A broader issue in the area of HRQL assessment that applies to the SAQOL/SAQOL-39 is the interpretation of the findings. Whilst the scores of a HRQL instrument may be useful in research, their meaning in a clinical setting is less obvious. Statistical significance does not imply clinical significance (Fayers & Machin, 2000) and only increased use and familiarity with specific instruments can begin to unravel clinically meaningful differences (Lydick, 2000).

A recent expert panel symposium in the Mayo Clinic focused on exploring ways of interpreting if a HRQL change is a clinically significant change in oncology, with the view of providing consensus papers from 30 HRQL experts. These would serve as a resource for researchers conducting HRQL research and clinicians that wish to incorporate HRQL measures in clinical practice for clinical decision making (Sloan, 2000; Sloan et al., 2002). Guyatt et al. (2002) discuss the main ways of establishing the interpretability of HRQL measures. One way is to relate HRQL changes to an independent standard or anchor that is itself interpretable (anchor-based approaches). This 'anchor' could be another measure. Or it could be patient-driven: firstly the smallest change in score that patients consider, on average, important is established (the minimum important difference- MID); then the proportion of patients who have achieved the MID is estimated. The other way is the distribution-based approaches, where an effect is expressed in terms of the underlying distribution of results. So effects may be expressed in terms of between-person SD units, within-person SD units, and the standard error of measurement. Guyatt et al. (2002) argue that use of multiple strategies is likely to enhance the interpretability of any particular instrument.

Research on the interpretation of HRQL results also needs to explore further different stakeholders' perspectives. For example, a MID can have a different meaning for a patient (e.g., the smallest change seen as important) and a different meaning for a clinician (e.g., the smallest change that would lead to a change in treatment plan) (Frost et al., 2002).

Methods of score interpretation need and are attracting increasing attention in the HRQL literature as they are of crucial importance for widespread adoption of HRQL measures into clinical practice. Further research into the SAQOL/SAQOL-39 could use it alongside other commonly used measures in stroke and aphasia and explore how to calibrate the meaning of changes in the SAQOL/SAQOL-39 in relation to the other measures in order to provide meaningful information to clinicians and consequently to patients. The MID with people with stroke and aphasia can also be explored using the SAQOL/SAQOL-39.

Lastly, further research is needed to evaluate the SAQOL/SAQOL-39 as a measure for making decisions at the individual-patient level. So far the instruments have been explored for their use for group level application. Measures that are used to make treatment decisions for individual patients need to be evaluated using different criteria from those used in this study (McHorney & Tarlov, 1995).

8.4 Predictors of HRQL in people with aphasia

8.4.1 Methodology

A number of studies have explored the factors associated with HRQL in people with stroke using univariate and bivariate statistics (e.g., Alshio et al., 1984; Astrom et al., 1992; Astrom, Asplund & Astrom, 1992). Such techniques allow for exploration of associations between two variables but they are not appropriate for use when the relative importance of different variables on another variable is explored. Multiple regression is the method of choice when

one wants to assess the relationship between a dependant variable and a number of potential predictors (Tabachnick & Fidell, 2001) and this method was used in this study.

In terms of the method of multiple regression, standard multiple regression was preferred to sequential and statistical regression as it suited best the research question: standard multiple regression is the method of choice when the relationship among variables is assessed and the multiple correlation among variables is explored (Tabachnick & Fidell, 2001).

In sequential regression, the order of entry of the IVs in the regression equation is determined by the researcher. Importance of IVs in the prediction equation is determined according to logic or theory. Sequential regression is recommended for specific hypothesis testing (Tabachnick & Fidell, 2001). The proportion of variance attributable to some IVs after variance due to IVs already in the equation is accounted for is tested. In this study's area of research, HRQL in people with aphasia, there is not enough strong evidence from the existing literature to justify entering variables in the equation in order of importance. Using sequential regression could be misleading as common variance between two IVs would be attributed to the one that entered the equation first.

In statistical stepwise regression again the IVs enter the regression equation in order, but the order is based on statistical criteria. This method is recommended when model building is the purpose of the research. The main disadvantage of statistical regression is that decisions about which variables enter and which are omitted from the equation are based solely on statistics computed from the particular sample (Tabachnick & Fidell, 2001). It therefore requires large and representative samples otherwise its results can be misleading. A cases to IV ratio of 40 to 1 is recommended, and an even larger sample is recommended if cross-validation (deriving the solution with some of the cases and testing it on the others) is to be

used to test the generalizability of the solution (Tabachnick & Fidell, 2001). A number of the studies reviewed on the predictors of HRQL in people after stroke used stepwise regression without an adequate cases to variables ratio (e.g., King, 1996; Kwa et al., 1996; Neau et al., 1998; Jonkman et al., 1998). This has obvious implications for the generalisability of their results.

In this study, univariate analyses were undertaken prior to the regression analysis to identify the significant variables to enter the regression model. This improved the cases to IVs ratio, ensuring that the derived solution was meaningful rather than an artefact of a low ratio (Tabachnick & Fidell, 2001).

With regard to the rest of the regression assumptions the absence of outliers ensured that there were no cases with undue influence on the regression coefficients and that the solution could generalise well to the population under study. The absence of multicollinearity ensured that there were no redundant variables in the model, that the importance of the predictors could be calculated and that the predictor equation was stable (Field, 2000; Berry, 1993). The normality, linearity, homoscedasticity and independence of residuals ensured that the statistics calculated were accurate and that the model could generalise to the population under study.

In short, using the most appropriate multiple regression method on the basis of the research question and making sure the assumptions of regression were met prior to carrying out the analysis ensured that a sound, meaningful model of predictors of HRQL for people with chronic aphasia after stroke was derived. This in combination with the 82% response rate also strengthened the generalisability of the model to the population under study.

8.4.2 *Main findings*

In this study the main predictors of poorer HRQL in people with chronic aphasia after stroke were reduced participation in activities, emotional distress, communication disability and comorbidity (adjusted R^2=.52). Cognition, social support, satisfaction with stroke services, stroke type (infarct versus haemorrhage) and demographic variables (age, gender, ethnicity, marital status, employment status and socioeconomic status) were not significant predictors of HRQL in these participants. These findings are discussed in relation to existing literature and their strengths, limitations and implications are raised.

The number of comorbid conditions was a significant predictor of HRQL in multivariate analysis whereas age was not. Existing evidence on the relationship between age and HRQL after stroke is not conclusive. In this study, there was a tendency for older people to have more comorbid conditions ($r = .37, p < .001$). This seems to indicate that it is not age itself that leads to reduced HRQL in people with chronic aphasia but rather the increased health problems that age may bring with it.

The other demographic variables explored in this study, namely gender, ethnic background, marital status, socioeconomic status and employment status, were also not significantly associated with HRQL. This finding strengthens the argument for an equitable service provision to all patients regardless of their demographic characteristics.

The stroke variables explored in this study were also not significantly associated with HRQL. With regard to the stroke type (infarct versus haemorrhage), it is acknowledged that no meaningful conclusions can be drawn from this study due to the high percentage of missing data (37%). This was due to the fact that information on the participants' strokes was obtained through their patients' notes and there was not enough available information. For

the same reason, factors like severity of stroke (as determined, for example, by level of consciousness at onset) or site and extent of lesion could not be explored. With regard to time post onset, the findings of this study suggest that once people with aphasia have reached the chronic stage after their stroke (i.e., more than 1 year post onset) time is not a significant predictor of HRQL and other factors become pertinent.

Physical disabilities and reduced level of activities have been repeatedly identified as among the main predictors of HRQL after stroke (Ahlsio et al., 1984; Ebrahim et al., 1986; Niemi et al., 1988; Viitanen et al., 1988; Astrom et al., 1992; Astrom, Asplund & Astrom 1992; Kwa et al., 1996; King, 1996; Wilkinson et al., 1997; Jonkman et al., 1998; Neau et al., 1998; Clarke et al., 1999; Lofgren et al., 1999; Carod-Artal et al., 2000). The most commonly used physical disabilities/ADL measure in stroke research is the Barthel Index. Although it is seen as an excellent measure of physical/functional disability, it lacks any assessment of the ability to carry out leisure activities or tasks in the community (e.g., Wilkinson et al., 1997). In this study the FAI was used, which is a valid measure of physical/functional disability (correlated with the Barthel with $r=.83$, $p<.001$) (Wilkinson et al., 1997) and also includes social function (e.g., leisure activities, work, travelling).

The fact that physical disability/reduced participation in activities continues to impact on HRQL in the long-term after the stroke has implications for rehabilitation. Traditionally, rehabilitation and recovery of function after stroke occur within the first 6 months. Still, there is some evidence that functional improvements occur a long time after rehabilitation has been completed and they are mostly due not to improvement in existing impairments but due to better compensation (Ferrucci et al., 1993). Facilitating and maximizing this process may be an eligible goal of long-term rehabilitation. Kelly-Hayes & Paige (1995) point out that rehabilitation efforts after 1 year are more likely to focus on factors related to socialisation

rather than recovery of function. An example of a service like this is reported by Drummond & Walker (1995). They found that, in a randomised control trial, five OT visits on leisure rehabilitation, delivered more than 1 year after stroke, led to increased number of, and time spent in, leisure activities. The difference was significant compared to standard OT and controls.

Related to the issue of participation in activities is the issue of emotional distress and depression. Depression after stroke affects functional recovery and improvement in the depression leads to improved functional recovery (Chemerinski et al., 2001; Gainotti et al., 2001). More generally, high emotional distress and depression have been associated with increased mortality 12 and 24 months after a stroke (House et al., 2001) and repeatedly associated with reduced HRQL (Ahlsio et al., 1984; Ebrahim et al., 1986; Niemi et al., 1988; King, 1996; Duncan et al., 1997; Jonkman et al., 1998; Neau et al., 1998; Wyller et al., 1998; Clarke et al., 1999; Lofgren et al., 1999; Carod-Artal et al., 2000). In terms of HRQL, this study's findings show a similar pattern with the subgroup of people living with chronic aphasia after stroke and emphasise the potential importance of these aspects for effective service provision.

In particular, these results highlight the importance of both identifying and then providing services to people experiencing emotional distress after the stroke. A caveat here is that identifying that emotional distress contributes to functional recovery and is a significant predictor for HRQL does not necessarily mean that service providers should add assessing emotional distress to their battery of assessments. Asking people to reveal these kinds of problems is probably unethical unless something is going to be done with the information obtained, such as offering appropriate services or timely onward referral.

In terms of appropriate interventions, a number of double-blind controlled trials have documented the efficacy of drug treatments, like tricyclic antidepressants (Lipsey et al., 1984), trazodone (Reding et al., 1986), and selective serotonin reuptake inhibitors (Andersen et al., 1994; Dam et al., 1996) in treating depression post-stroke. However, what is still unclear is whether improvements in depressive symptoms will also improve functional status and HRQL, or whether this will require further rehabilitation therapies (Herrmann et al., 1998).

Knapp et al. (2000) reviewed randomised trials of non-drug strategies to resolve psychosocial difficulties after stroke. These included educational and informational programmes, leisure therapy, support workers and counselling. On the whole, the results were disappointing, with only weak positive findings. The authors, however, concluded that their review did not establish that psychosocial interventions after stroke were ineffective. In the reviewed trials the therapists were not trained in delivering psychological interventions; they were rarely supervised; compliance to therapy was not measured; and the therapists' work was not based on an explicit psychological theory (Knapp et al., 2000).

Services addressing the emotional distress that people with aphasia are dealing with are often not available routinely. The clear link with HRQL demonstrated here suggests that it should have a higher priority in service provision. The evidence presented so far suggests, however, that this need not necessarily be through implementing full-blown psychological therapies. For one thing, single and simple interventions are rarely effective in rehabilitation (Sinclaire & Dickinson, 1998). Moreover, theories of coping (Lazarus & Folkman, 1984) suggest that providing only support or information would not result in changes to patient or carer mood. Tackling internal resources by addressing individual coping skills would be more likely to affect a change (Knapp et al., 2000). Ways of addressing this may include, for example, incorporation of work on self-esteem and confidence building alongside other therapies (e.g.

Pound et al., 2000), or modification of attitude and behaviour by health care staff and carers, which can affect patients' motivation for and response to rehabilitation (Maclean et al., 2000; Parr et al., 1997).

The majority of stroke studies that included people with aphasia concluded that the HRQL of people with aphasia was not significantly different from that of people living with the effects of stroke without aphasia. If presence of aphasia were not a significant predictor of HRQL then it seemed likely that severity of aphasia, within an all aphasic population, would be an even weaker predictor.

This was not the case in these findings. Communication disability was measured with the ASHA-FACS. The ASHA-FACS correlate highly with measures of aphasia language impairment, such as the Western Aphasia Battery (Kertesz, 1982) ($r = .76, p<.05$) (Frattali et al., 1995) and in the current study the FAST ($r = .79, p<.01$). Severity of communication disability (as measured by the ASHA-FACS) was a significant predictor of HRQL with higher communication disability resulting in poorer quality of life. This was despite the fact that the majority of our participants had high scores on the ASHA-FACS, i.e., indicative of mild communication disability (67.5% scored at or above 6, with scores ranging from 1 to 7). These findings are similar to the Kwa et al. (1996) study where 38% of the subjects had aphasia (measured with the BDAE). Severity of aphasia was a significant predictor of quality of life despite the fact that 25% of their subjects could not complete the quality of life assessment due to communication problems.

A number of methodological issues may explain why aphasia was not a significant predictor of HRQL in other stroke studies. As has been already indicated (chapter 2), in some studies aphasia resulted in missed assessments (Ebrahim et al., 1986; Angeleri et al., 1993; Wilkinson

et al., 1997). In other studies, proxy respondents were used instead of the person with aphasia (Astrom et al., 1992; Astrom, Asplund & Astrom, 1992; de Haan et al., 1995; Tuomilehto et al., 1995; Neau et al., 1998) and the results were analysed alongside the self-report data. This was despite the documented disagreement between proxy and self-report in rating functional abilities (Knapp & Hewison, 1999) and quality of life (Sneeuw et al., 1997) after stroke. Lastly, in the remaining reviewed studies that included people with aphasia (Niemi et al., 1988; King, 1996; Lofgren et al., 1999; Bethoux et al., 1999) quite complex instruments were used to measure HRQL. None of these studies give any information on how people with aphasia managed these complex instruments. The validity of these assessments is questioned as people with aphasia may have not understood at least some of the items or may have not been able to express their responses with precision.

Cognitive level was not a significant predictor of HRQL in this sample. The findings agree with those of one study that specifically investigated the role of cognitive decline on HRQL after stroke (Kwa et al., 1996). These authors assessed both aphasia and cognition with valid measures. They used the BDAE for aphasia and the CAMCOG to measure cognition, which is part of the Cambridge Examination for Mental Disorders of the Elderly (CAMDEX, Roth et al., 1986). They indicated that, during the CAMCOG, people with aphasia were helped if needed with gestures and pointing to ensure that it was not their language skills that were assessed but their cognition. They concluded that cognitive decline was not a significant predictor of HRQL after stroke.

A few studies have associated cognitive decline with reduced HRQL after stroke (Niemi et al., 1988; Jonkman et al., 1998; and Clarke et al., 1999). In the first two of these studies cognition was assessed with the WAIS and the WMS, which rely heavily on language. For people with aphasia, it is unclear whether such instruments measure language or cognitive

skills. The third study (Clarke et al., 1999) did not attempt to differentiate between aphasia and cognitive decline. Rather the authors measured "cognitive disability" with the communication and cognition sub-scales of the FIM. Such assessments will tend to identify people with aphasia as also having cognitive decline. The conclusion, therefore, in these studies that cognitive decline affects HRQL may well mask the effect of aphasia on HRQL.

A number of studies have found that aspects of social support seem to affect quality of life after stroke (Osberg et al., 1988; Viitanen et al., 1988; Astrom et al., 1992; Astrom, Asplund & Astrom 1992; Wyller et al., 1998). In this sample of people with chronic aphasia, social support was not a significant predictor of HRQL. This lack of agreement with previous studies could be attributed to a number of factors. Firstly, these studies did not assess HRQL as defined in the present study. Rather they assessed life satisfaction (Osberg et al., 1988; Viitanen et al., 1988; Astrom et al., 1992; Astrom, Asplund & Astrom, 1992), quality of life including financial well-being (King, 1996) or subjective well-being (Wyller et al., 1998).

Secondly, their conceptualisation of social support was different. Some of them actually looked at social networks (e.g., marital status, number of friends and relatives, number of people in the household) (Osberg et al., 1988; Astrom et al., 1992; Astrom, Asplund & Astrom, 1992) or social integration/social functioning (e.g., participation in social activities) (Viitanen et al., 1988; Osberg et al., 1988). Wyller et al. (1998) used a self-developed questionnaire and it is not known what aspects of social support they looked at. In this study, the SSS was used to measure social support, as it is a measure very closely reflecting the most widely accepted definition of social support. Social support is generally defined "in terms of the availability of people whom the individual trusts, on whom he can rely, and who make him feel cared for and valued as a person" (McDowell & Newell, 1996). This definition agrees with those proposed by others in that it emphasises the subjective experience, i.e. the

importance of the recipient's *perception* of support (e.g., Cobb, 1976; House, 1981; Turner & Noh, 1983; Friedland & McColl, 1992).

Thirdly, the lack of association between social support and HRQL, in this sample of people with chronic aphasia after stroke, may be related to the distribution of the social support scores. The SSS scores range from 1 to 5 with high scores indicating high social support and in our sample the median was 3.9 and the mean 3.7. Only 12% of the participants scored 1 or 2 in this scale compared to 66.3% who scored 4 or 5. The fact that this sample had high levels of support may account, at least partly, for the lack of a significant association between social support and HRQL.

Lastly, this lack of association may indeed be a true finding. In a related area, Robinson et al. (1999) found that during the first few weeks after stroke perceived social support was highly associated with depression whereas during the chronic period (12- or 24-month follow up) this association was not significant and other factors like financial security, living arrangements and work experience were more pertinent.

8.4.3 Future research

This study suggested that extent of communication disability has a significant impact on the HRQL of people with aphasia following stroke. Future studies could use the SAQOL/SAQOL-39 with stroke survivors with and without aphasia. This would allow for direct comparisons between different stroke groups. It would enable us to understand better the impact of aphasia on the HRQL of people after stroke.

More research is needed in the area of HRQL outcomes in severe aphasia using a range of methodologies. As has been already indicated, the findings of this study on HRQL in people with severe aphasia using proxy respondents will be explored. Alternative methodologies

include qualitative techniques like participant and non-participant observation. All of these approaches however are methodologically challenging. HRQL is generally defined as a subjective concept. This makes it hard to observe without making value judgements that link the observed behaviour to the assumed subjective perception. This is problematic for measurement.

Emotional distress was one of the main predictors of HRQL in this group of people with aphasia. Drug treatments have been documented to be effective in treating depression post-stroke but further research is needed to explore in systematic ways whether and how non-drug treatments could be used to address psychosocial outcomes. In their review of RCT's in this area, Knapp et al. (2000) make some suggestions for further research. The evaluated interventions should have a sound theoretical basis; they should be plausible in terms of timing, intensity and duration of treatment; and they should also be plausible in terms of the training and supervision of those providing them. Appropriate measures should be used to measure change and studies evaluating the effectiveness of such interventions should use rigorous scientific criteria. Such research would shed some light on whether non-drug treatments for emotional distress and other psychosocial problems after stroke have indeed limited effectiveness or whether they have been ineffectively delivered or evaluated.

Further work is also needed to investigate the inter-relationship between communication disability, emotional distress and activity level and how they interact to affect HRQL. In this area, longitudinal cohort studies could be used to start unraveling cause and effect relationships.

Future studies could also investigate the influence of social support on quality of life in aphasia. Using a combination of different support indicators such as social network (e.g.,

number of friends and relatives, contact with friends and relatives, group membership) and perceived support (e.g., the SSS) may help explore whether there are any effects that were not identified in the current investigation.

Lastly further studies could replicate this study with different groups of people with aphasia. This would allow greater confidence to be placed to the findings of this study.

8.5 Conclusions

This study explored the impact of stroke and aphasia on the ability of people with chronic aphasia to lead a fulfilling life. There are various ways of assessing HRQL. This study sought to assess this concept in a way that could be replicated in clinical practice, so that, primarily, it could give health professionals working with people with aphasia useful insights on their clients HRQL. The use of a single questionnaire was thought to be the most feasible way of achieving this.

In this study the SAQOL was developed and a shorter version of it, the SAQOL-39, was found to be a useful measure for the assessment of HRQL in people with aphasia after stroke. Firstly, it was accessible to the population under study and despite their communication disability 87% of the aphasic participants were able to self-report in an interview format. Secondly, it was acceptable to the participants. Thirdly, it had high internal consistency, test-retest reliability and construct validity. Thus, the initial stages of the psychometric evaluation of the instrument showed it had good properties.

Its good psychometric properties mean that the SAQOL-39 can be used by health professionals to assess and understand better the HRQL of people with chronic aphasia. To avoid conceptual confusion and to increase the appropriateness of the measure, the users should explain to their clients what the instrument covers and why they are using it prior to

its administration. They should ensure their clients are willing to have the domains of HRQL that the SAQOL-39 covers assessed and that they understand the potential uses of the provided information. These may include to understand better the impact of aphasia and stroke on different aspects of people's lives, to identify areas they are particularly unhappy about or areas where their needs are not met with the rehabilitation they receive.

In relation to this, this study also indicated that the SAQOL-39 can be used as a tool in discussing difficult issues with clients. Because of their language difficulties, people with aphasia have increased difficulty in discussing their feelings or explaining what causes them distress. Their answers on some of the items of the questionnaire can facilitate an exploration of their needs and any further input or services they may need.

The absence of an aphasia-friendly HRQL measure so far has meant that health professionals who wished to assess the HRQL of their aphasic clients had to rely on either qualitative interviewing or using a battery of tests to assess each domain of the concept[15]. Both these approaches require a lot of time to administer. Data derived from qualitative interviewing also require a lot of time to code and interpret. Scores from different tests in a battery of assessments are not directly comparable. As a result, administering, scoring and interpreting such findings can be a great burden to clinicians. The SAQOL-39 is a quick to administer measure and a clear manual is being developed on its administration and scoring. It can thus reduce the burden of assessing HRQL in clinical practice.

As is common with new measures, further research is needed on the psychometric properties of the SAQOL-39 and on its appropriateness as a clinical outcome measure. Pending this research it may prove to be a useful outcome measure with potential uses in treatment

[15] Anecdotal evidence from SLT's and stroke physicians

evaluation, service evaluation, clinical audit and individual client assessment and treatment prioritisation.

This study also explored what the main predictors of HRQL, as measured by the SAQOL-39, were. The HRQL of people living with long term aphasia after stroke is significantly affected by emotional distress, reduced participation in activities, severity of communication disability and increased comorbidity. Service providers need to take these factors into account when planning and implementing interventions aimed at improving people's HRQL. More research is needed in order to evaluate the effectiveness of such interventions. Still, the factors identified here as important in predicting HRQL suggest that long-term services to people with aphasia need to consider certain areas in particular. These comprise facilitating emotional health, enabling participation in someone's immediate social context and in the community and society more generally (Byng et al 2000, Pound et al 2000) and engaging with the factors which contribute to communication disability.

REFERENCES

1. Aaronson N.K., Ahmedzai S., Bergman B., Bullinger M., Cull A., Duez N.J., Filiberti A., Flechtner H., Fleishman S.B., de Haes J.C. (1993) The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl.Cancer Inst.* **85,** 365-376.

2. Adams R.J., Meador K.J., Sethi K.D., Grotta J.C., & Thomson D.S. (1987) Graded neurologic scale for use in acute hemispheric stroke treatment protocols. *Stroke* **18,** 665-669.

3. Ahlsio B., Britton M., & Murray V. (1984) Disablement and Quality of Life After Stroke. *Stroke* **15,** 886-890.

4. Andersen G., Vestergaard K., & Lauritzen L. (1994) Effective treatment of poststroke depression with the selective serotonin reuptake inhibitor citalopram. *Stroke* **25,** 1099-1104.

5. Anderson C.S., Linto J., & Stewart-Wynne E.G. (1995) A Population Based Assessment of the Impact and Burden of Caregiving for Long-term Stroke Survivors. *Stroke* **26,** 843-849.

6. Andreu N., Collet J.-P., Hanley J., Montastruc J.-L., Rascol O., & Wood-Dauphinee S. (2000) Quality of Life in Parkinson's Disease: Measurement Study of the Parkinson's Disease Questionnaire (PDQ-39) French Version. *Quality of Life Research* **9,** 283.

7. Angeleri F., Angeleri V.A., Foschi N., Giaquinto S., & Nolfe G. (1993) The Influence of Depression, Social Activity, and Family Stress on Functional Outcome After Stroke. *Stroke* **24,** 1478-1483.

8. Antaki C. & Rapley M. (1996) Questions and Answers to Psychological Assessment Schedules: Hidden Troubles In 'Quality of Life' Interviews. *Journal of Intellectual Disability Research* **40,** 421-437.

9. Arber S. (1997) Insights about the Non-employed Class and Health:Evidence from the General Household Survey. In: *Constructing Classes. Towards a New Social Classification for the UK* (eds Rose D. & O'Reilly K.), pp. 78-92. ESRC/ONS, Swindon.

10. Astrom M., Asplund K., & Astrom T. (1992) Psychosocial Function and Life Satisfaction After Stroke. *Stroke* **23,** 527-531.

11. Astrom M., Adolfsson R., Asplund K., & Astrom T. (1992) Life Before and After Stroke. Living Conditions and Life Satisfaction in Relation to a General Elderly Population. *Cerebrovasc Dis* **2,** 28-34.

12. Bamford J., Sandercock P., Warlow C., & Slattery J. (1989) Interobserver Agreement for the Assessment of Handicap in Stroke Patients. *Stroke* **20,** 828.

13. Bamford J., Sandercock P., Dennis M., Burn J., & Warlow C. (1990) A prospective study of acute cerebrovascular disease in the community: the Oxfordshire Community Stroke Project--1981-86. 2. Incidence, case fatality rates and overall outcome at one year of cerebral infarction, primary intracerebral and subarachnoid haemorrhage. *J Neurol Neurosurg.Psychiatry* **53,** 16-22.

14. Baumann B.O. (1961) Diversities in Conceptions of Health and Physical Fitness. *Journal of Health and Human Behavior* **3,** 39-46.

15. Bech P. (1990) Measurement of Psychological Distress and Well-Being. *Psychotherapy and Psychosomatics* **54,** 77-89.

16. Bech P. (1993) Quality of Life Measurement in Chronic Disorders. *Psychotherapy and Psychosomatics* **59,** 1-10.

17. Beck A.T., Ward C.H., Mendelson M., & et al. (1961) An Inventory for Measuring Depression. *Archives of General Psychiatry* **4,** 561-571.

18. Bergner M., Bobbitt R., Carter W., & Gilson B. (1981) The Sickness Impact Profile. *Medical Care* **19,** 787-805.

19. Bergner M. & Rothman M.L. (1987) Health Status Measures: An Overview and Guide for Selection. *American Review of Public Health* **8,** 191-210.

20. Berry W.D. (1993) Understanding Regression Assumptions. Sage Publications, Newbury Park, California.

21. Berzon R., Hays R.D., & Shumaker S.A. (1993) International Use, Application and Performance of Health-Related Quality of Life Instruments. *Quality of Life Research* **2,** 367-368.

22. Bethoux F., Calmels P., & Gautheron V. (1999) Changes in the Quality of Life of Hemiplegic Stroke Patients With Time: A Preliminary Report. *American Journal of Physical Medicine and Rehabilitation* **78,** 19-23.

23. Black-Schaffer R.M. & Osberg J.S. (1990) Return to work after stroke: development of a predictive model. *Arch.Phys Med Rehabil* **71,** 285-290.

24. Bond S. & Bond J. (1990) Outcomes of Care Within a Multiple Case Study In the Evaluation of the Experimental NHS Nursing Homes. *Age Ageing* **19,** 11.

25. Bowling A. & Browne P.D. (1991) Social Networks, Health, and Emotional Well-being Among the Oldest Old in London. *Journal of Gerontology* **46,** S20-S32.

26. Bowling A. (1995) What Things Are Important in People's Lifes? A Survey of the Public's Judgements to Inform Scales of Health-Related Quality of Life. *Social Science and Medicine* **41,** 1447-1462.

27. Bowling A. (1995) Measuring Disease. Open University Press, Buckingham.

28. Brock D. (1993) Quality of Life Measures in Health Care and Medical Ethics. In: *The Quality of Life* (eds Nussbaum M. & Sen A.) Clarendon Press, Oxford.

29. Buck D., Jacoby A., Massey A., & Ford G. (2000) Pre-testing Items In a Stroke-specific Quality of Life Measure: The Value of Cognitive Interviewing. *Journal of Neurolinguistics* **13,** 267-271.

30. Buck D., Jacoby A., & Ford G. (2001) Reliability and Validity of 'NEWSQOL': The Newcastle Stroke-specific Quality of Life Measure. *Quality of Life Research* **10,** 258.

31. Bullinger M., Anderson R., Cella D., & Aaronson N.K. (1993) Developing and Evaluating Cross Cultural Instruments: from Minimum Requirements to Optimal Models. *Quality of Life Research* **2,** 451-459.

32. Burvill P., Johnson G., Jamrozik K., Anderson C., & Stewart-Wynne E. (1997) Risk factors for post-stroke depression. *Int J Geriatr.Psychiatry* **12,** 219-226.

33. Byng S., Pound C. & Parr S. (2000) Living with Aphasia: Frameworks for Therapy Interventions. In: *Acquired Neurological Communication Disorders: A Clinical Perspective.* (ed Papathanasiou I.) Whurr Publishers, London.

34. Carlson J.S. & Jensen C.M. (1981) Reliability of the Raven Coloured Progressive Matrices Test: Age and Ethnic Group Comparisons. *Journal of Consulting and Clinical Psychology* **49,** 320-322.

35. Carod-Artal J., Egido J.A., Gonzalez J.L., & Varela d.S. (2000) Quality of life among stroke survivors evaluated 1 year after stroke: experience of a stroke unit. *Stroke* **31,** 2995-3000.

36. Chemerinski E., Robinson R.G., & Kosier J.T. (2001) Improved recovery in activities of daily living associated with remission of poststroke depression. *Stroke* **32,** 113-117.

37. Clark P. & Bowling A. (1990) Quality of Everyday Life in Long-stay Institutions for the Elderly. *Social Science and Medicine* **30,** 1201.

38. Clarke P.J., Black S.E., Badley E.M., Lawrence J.M., & Williams J.I. (1999) Handicap in Stroke Survivors. *Disability and Rehabilitation* **21,** 116-123.

39. Cobb S. (1976) Social Support As a Moderator of Life Stress. *Psychosomatic Medicine* **38,** 300-314.

40. Code C. & Muller D.J. (1992) The Code-Muller Protocols:Assessing Perceptions of Psychosocial Adjustment to Aphasia and Related Disorders. Whurr, London.

41. Code C., Muller D.J., Hogan A., & Herrmann M. (1999) Perceptions of Psychosocial Adjustment to Acquired Communication Disorders: Applications of the Code-Muller Protocols. *International Journal of Language and Communication Disorders* **34,** 193-207.

42. Code C., Muller D.J., & Herrmann M. (1999) Perceptions of Psychosocial Adjustment to Aphasia: Applications of the Code-Muller Protocols. *Seminars in Speech and Language* **20,** 51-63.

43. Cohen S. & Wills T.A. (1985) Social Support and the Buffering Hypothesis. *Psychological Bulletin* **98,** 310-357.

44. Cohen S. & Syme S.L. (1985) Social Support and Health. Academic, Orlando, Florida.

45. Connell C.M. & D'Augelli A.R. (1990) The contribution of personality characteristics to the relationship between social support and perceived physical health. *Health Psychology* **9,** 192-207.

46. Crowne D.P. & Marlowe D. (1960) A New Scale of Social Desirability Independent of Psychopathology. *Journal of Consulting Psychology* **24,** 354.

47. Cruice M., Worrall L., & Hickson L. (2000) Quality of Life for People with Aphasia: Performance and Usability of Quality of Life Assessments. *Asia Pacific Journal of Speech, Language and Hearing* **5,** 85-91.

48. Cruice M., Worrall L., & Hickson L. (2000) Quality of Life Measurement in Speech Pathology and Audiology. *Asia Pacific Journal of Speech, Language and Hearing* **5,** 1-20.

49. Cruice M., Worrall L., & Hickson L. (2001) "No Life in it" or "Really Wonderful Really": Just Exactly How Is the Quality of Life With Aphasia Described and Determined? Presentation at the British Aphasiology Society Biennial International Conference. Exeter, UK.

50. Dam M., Tonin P., De Boni A., Pizzolato G., Casson S., Ermani M., Freo U., Piron L., & Battistin L. (1996) Effects of fluoxetine and maprotiline on functional recovery in poststroke hemiplegic patients undergoing rehabilitation therapy. *Stroke* **27,** 1211-1214.

51. Dancey C.P. & Reidy J. (2002) Statistics without Maths for Psychology. Using SPSS for Windows. Prentice Hall, Harlow, UK.

52. de Haan R., Horn J., Limburg M., Van Der M.J., & Bossuyt P. (1993) A comparison of five stroke scales with measures of disability, handicap, and quality of life. *Stroke* **24,** 1178-1181.

53. de Haan R.J., Limburg M., Van der Meulen J.H.P., Jacobs H.M., & Aaronson N.K. (1995) Quality of Life After Stroke: Impact of Stroke Type and Lesion Location. *Stroke* **26,** 402-408.

54. Demeurisse G., Demol O., & Robaye E. (1980) Motor Evaluation In Vascular Hemiplegia. *European Neurology* **19,** 382-389.

55. Dennis M., O'Rourke S., Slattery J., Staniforth T., & Warlow C. (1997) Evaluation of a Stroke Family Care Worker: Results of a Randomised Controlled Trial. *British Medical Journal* **314,** 1071.

56. Dennis M., O'Rourke S., Lewis S., Sharpe M., & Warlow C. (2000) Emotional Outcomes After Stroke: Factors Associated with Poor Outcome. *Journal of Neurology, Neurosurgery and Psychiatry* **68,** 47-52.

57. Department of Health. Health Survey for England. Cardiovascular Disease. 1998. Accessed in, www.archive.official-documents.co.uk/document/doh/survey98/hse-00.htm.
Ref Type: Report

58. Department of Health. Good Practice in Consent Implementation Guide: Consent to Examination and Treatment. 2001. London, Department of Health Publications.
Ref Type: Report

59. Deyo R.A., Diehr P., & Patrick D.L. (1991) Reproducibility and Responsiveness of Health Status Measures. *Controlled Clinical Trials* **12,** 142-158.

60. Donaldson S.W., Wagner C.C., & Gresham G.E. (1973) Unified ADL Evaluation Form. *Archives of Physical Medicine and Rehabilitation* **54,** 175-179.

61. Dorman P., Dennis M., & Sandercock P. (1999) How Do Scores on the EuroQol Relate to Scores on the SF-36 After Stroke. *Stroke* **30,** 2146-2151.

62. Drummond A.E.R. & Walker M.F. (1995) A Randomised Control Trial of Leisure Rehabilitation After Stroke. *Clinical Rehabilitation* **9,** 283-290.

63. Duncan P.W., Samsa G.P., Weinberger M., Goldstein L.B., Bonito A., Witter D.M., Enarson C., & Matchar D. (1997) Health Status of Individuals with Mild Stroke. *Stroke* **28,** 740-745.

64. Duncan P.W., Wallace D., Lai S.-M., Johnson D., Embretson S., & Jacobs Laster L. (1999) The Stroke Impact Scale Version 2.0. Evaluation of Reliability, Validity, and Sensitivity to Change. *Stroke* **30,** 2131-2140.

65. Ebrahim S., Barer D., & Nouri F. (1986) Use of the Nottingham Health Profile with Patients After a Stroke. *Journal of Epidemiology and Community Health* **40,** 166-169.

66. Ebrahim S. (1995) Clinical and Public Health Perspectives and Applications of Health-Related Quality of Life Measurement. *Social Science and Medicine* **41,** 1383-1394.

67. Enderby P., Wood V., & Wade D. (1987) Frenchay Aphasia Screening Test. NFER-Nelson, Windsor.

68. Erickson P. (2000) Assessment of the Evaluative Properties of Health Status Instruments. *Medical Care* **38,** 95-99.

69. EuroQol Group (1990) EuroQol: A New Facility for the Measurement of Health Related Quality of Life. *Health Policy* **16,** 199-208.

70. Farquhar M. (1995) Elderly People Definitions of Quality of Life. *Social Science and Medicine* **41,** 1439-1446.

71. Fayers P.M. & Jones D.R. (1983) Measuring and Analysing Quality of Life in Cancer Clinical Trials: A Review. *Statistics In Medicine* **2,** 429-446.

72. Fayers P.M. & Machin D. (2000) Quality of Life: Assessment, Analysis and Interpretation. John Wiley & Sons Ltd, Chichester, UK.

73. Ferguson E. & Cox T. (1993) Exploratory Factor Analysis: A User's Guide. *International Journal of Selection and Assessment* **1,** 84-94.

74. Ferrans C. & Powers M. (1985) Quality of Life Index: Development and Psychometric Properties. *Advances in Nursing Science* **8,** 24.

75. Ferrucci L., Bandinelli S., Guralnik J.M., Lamponi M., Bertini C., Falchini M., & Baroni A. (1993) Recovery of functional status after stroke. A postrehabilitation follow- up study. *Stroke* **24,** 200-205.

76. Field A. (2000) Discovering Statistics Using SPSS for Windows. Sage Publications, London.

77. Fitzpatrick R., Fletcher A., Gore S., Jones D., Spiegelhalter D., & Cox D. (1992) Quality of life measures in health care. I: Applications and issues in assessment. *BMJ* **305,** 1074-1077.

78. Fitzpatrick R., Davey C., Buxton M.J., & Jones D.R. (1998) Evaluating Patient-based Outcome Measures for Use in Clinical Trials. *Health Technology Assessment* **2.**

79. Fletcher A., Gore S., Jones D., Fitzpatrick R., Spiegelhalter D., & Cox D. (1992) Quality of life measures in health care. II: Design, analysis, and interpretation. *BMJ* **305,** 1145-1148.

80. Fletcher A. (1995) Quality-of-life measurements in the evaluation of treatment: proposed guidelines. *Br.J Clin Pharmacol.* **39,** 217-222.

81. Folstein M., Folstein S., & McHugh P. (1975) "Mini-Mental State": A Practical Method for Grading the Cognitive State of Patients for the Clinician. *Journal of Psychiatric Research* **12,** 189-198.

82. Foster A. & Young J. (1996) Specialist Nurse Support for Patients with Stroke in the Community: A Randomised Control Trial. *British Medical Journal* **312,** 1642-1646.

83. Fowler F.J. (1993) Survey Research Methods, 2nd edn. Sage, Newbury Park, CA.

84. Frattali C.M., Thompson C.K., Holland A.L., Wohl C.B., & Ferketic M.M. (1995) Functional Assessment of Communication Skills for Adults. American Speech and Hearing Association, Rockville, MD.

85. Frey W.D. (1984) Functional Assessment in the 1980's: A Conceptual Enigma, A Technical Challenge. In: *Functional Assessment in Rehabilitation* (eds Halpern A.S. & Furher M.J.) Brookes, Baltimore.

86. Friedland J. & McColl M. (1987) Social Support and Psychosocial Dysfunction After Stroke: Buffering Effects in a Community Sample. *Archives of Physical Medicine and Rehabilitation* **68,** 475-480.

87. Friedland J.F. & McColl M. (1992) Social Support Intervention After Stroke: Results of a Randomised Trial. *Archives of Physical Medicine and Rehabilitation* **73,** 573-581.

88. Frost M.H., Bonomi A.E., Ferrans C.E., Wong G.Y., & Hays R.D. (2002) Patient, clinician, and population perspectives on determining the clinical significance of quality-of-life scores. *Mayo Clin Proc.* **77,** 488-494.

89. Fukunishi I. & Rahe R.H. (1995) Alexithymia and coping with stress in healthy persons: alexithymia as a personality trait is associated with low social support and poor responses to stress. *Psychological Reports* **76,** 1299-1304.

90. Fukunishi I., Aoki T., & Hosaka T. (1997) Correlations for Social Support with Depression in the Chronic Poststroke Period. *Perceptual and Motor Skills* **85,** 811-818.

91. Gainotti G. (1997) Emotional, Psychological and Psychosocial Problems of Aphasic Patients: An Introduction. *Aphasiology* **11,** 635-650.

92. Gainotti G., Antonucci G., Marra C., & Paolucci S. (2001) Relation between depression after stroke, antidepressant therapy, and functional recovery. *J Neurol Neurosurg.Psychiatry* **71,** 258-261.

93. Geddes J.M., Fear J., Tennant A., Pickering A., Hillman M., & Chamberlain M.A. (1996) Prevalence of self reported stroke in a population in northern England. *J Epidemiol.Community Health* **50,** 140-143.

94. Gibbs R.G., Todd J.C., Irvine C., Lawrenson R., Newson R., Greenhalgh R.M., & Davies A.H. (1998) Relationship between the regional and national incidence of transient ischaemic attack and stroke and performance of carotid endarterectomy. *Eur.J Vasc.Endovasc.Surg.* **16,** 47-52.

95. Gill T.M. & Feinstein A.R. (1994) A Critical Appraisal of the Quality of Quality of Life Measurements. *JAMA* **272,** 619-626.

96. Goldberg D.P. (1972) The Detection of Psychiatric Illness by Questionnaire. Oxford University Press, London.

97. Goldberg D.P. & Hillier V.F. (1979) A Scaled Version of the General Health Questionnaire. *Psychological Medicine* **9,** 139-145.

98. Goldberg D.P. & Huxley P. (1980) Mental Illness in the Community:The Pathway to Psychiatric Care. Tavistock, London.

99. Goldberg D.P. & Williams P. (1988) A User's Guide to the General Health Questionnaire (GHQ). NFER-Nelson, Oxford.

100. Goodglass H. & Kaplan E. (1983) The Assessment of Aphasia and Related Disorders. Lea & Febiger, Philadelphia.

101. Gottschalk L.A. & Lolas F. (1992) The Mearurement of Quality of Life through the Content Analysis of Verbal Behaviour. *Psychotherapy and Psychosomatics* **58,** 69-78.

102. Granger C.V., Cotter A.C., Hamilton B.B., & Fiedler R.C. (1993) Functional assessment scales: a study of persons after stroke. *Arch.Phys Med Rehabil* **74,** 133-138.

103. Gresham G.E., Phillips T.F., Wolf P.A., McNamara P.M., Kannel W.B., & Dawber T.R. (1979) Epidemiologic Profile of Long-term Stroke Disability: The Framingham Study. *Archives of Physical Medicine and Rehabilitation* **60,** 487-491.

104. Gurel L., Linn M.W., & Linn B.S. (1972) Physical and Mental Impairment-of-Function Evaluation in the Aged. *Journal of Gerontology* **27,** 83-90.

105. Guyatt G.H., Bombardier C., & Tugwell P. (1986) Measuring Disease Specific Quality of Life in Clinical Trials. *Canadian Medical Association Journal* **134,** 899.

106. Guyatt G.H. (2001) Methods Used to Date for Clinical Significance.Presentation at the ISOQOL Annual Meeting. Amsterdam, The Netherlands.

107. Guyatt G.H., Osoba D., Wu A.W., Wyrwich K.W., & Norman G.R. (2002) Methods to explain the clinical significance of health status measures. *Mayo Clin Proc.* **77,** 371-383.

108. Hackett M.L., Duncan J.R., Anderson C.S., Broad J.B., & Bonita R. (2000) Health-Related Quality of Life Among Long-Term Survivors of Stroke. Results From the Auckland Stroke Study, 1991-1992. *Stroke* **31,** 440-447.

109. Hackett M.L. & Anderson C.S. (2001) Differences in Health Related Quality of Life among Stroke Subtypes. *Quality of Life Newsletter* 22-23.

110. Hamedani A.G., Wells C.K., Brass L.M., Kernan W.N., Viscoli C.M., Maraire J.N., Awad I.A., & Horwitz R.I. (2001) A quality-of-life instrument for young hemorrhagic stroke patients. *Stroke* **32,** 687-695.

111. Hamilton M. (1960) A Rating Scale for Depression. *Journal of Neurology, Neurosurgery and Psychiatry* **23,** 56-62.

112. Harwood R., Gompertz P., & Ebrahim S. (1994) Handicap One Year After a Stroke: Validity of a New Scale. *Journal of Neurology, Neurosurgery and Psychiatry* **57,** 825-829.

113. Hawker G., Melfi C., Paul J., Green R., & Bombardier C. (1995) Comparison of a generic (SF-36) and a disease specific (WOMAC) (Western Ontario and McMaster Universities Osteoarthritis Index) instrument in the measurement of outcomes after knee replacement surgery. *J Rheumatol.* **22,** 1193-1196.

114. Hays R.D., Anderson R., & Revicki D. (1993) Psychometric Considerations in Evaluating Health-Related Quality of Life Measures. *Quality of Life Research* **2,** 441-449.

115. Hemsley G. & Code C. (1996) Interactions Between Recovery in Aphasia, Emotional and Psychosocial Factors in Subjects with Aphasia, Their Significant Others and Speech Pathologists. *Disability and Rehabilitation* **18,** 567-584.

116. Hennekens C.H. & Buring J.E. (1987) Epidemiology in Medicine. Little, Brown & Company, Boston.

117. Herrmann N., Black S.E., Lawrence J., Szekely C., & Szalai J.P. (1998) The Sunnybrook Stroke Study: a prospective study of depressive symptoms and functional outcome. *Stroke* **29,** 618-624.

118. Heyrman J. & van Hoeck K. Measuring Health Outcome: Shouldn't We First Define Health? Paper presented to the WONCA/SIMG Congress, Quality of Care in Family Medicine/General Practice.The Hague, Netherlands, June 13-17. 1993. Ref Type: Unpublished Work

119. Hilari K. Modification of the Stroke-Specific Quality of Life Scale for People with Aphasia. Quality of Life Research 9[3], 285. 2000. Ref Type: Abstract

120. Hilari K. & Byng S. (2001) Measuring quality of life in people with aphasia: the Stroke Specific Quality of Life Scale. *Int J Lang Commun.Disord.* **36 Suppl,** 86-91.

121. Hochstenbach J.B., Donders A.R., Mulder T., vanLimbeek J., & Schoonderwaldt H. (1996) Many Chronic Problems in CVA Patients at Home. *Ned Tijdschr Geneesk* **140,** 1182-1186.

122. Hodge J. (1994) The Quality of Life. A contrast between Utalitarian and Existentialist approaches. In: *Quality of Life. Perspectives and Policies.* Routledge, London.

123. Hoen B., Thelander M., & Worsley.J. (1997) Improvement in Psychological Well-Being of People with Aphasia and Their Families:Evaluation of a Community- Based Programme. *Aphasiology* **11,** 681-691.

124. Horn S.D., Chachich B., & Clopton C. (1983) Measuring Severity of Illness: A Reliability Study. *Medical Care* **21,** 705-714.

125. House A., Knapp P., Bamford J., & Vail A. (2001) Mortality at 12 and 24 months after stroke may be associated with depressive symptoms at 1 month. *Stroke* **32,** 696-701.

126. House J.S. (1981) Work Stress and Social Support. Addison-Wesley, Reading, MA.

127. Hunt S.M., McKenna S.P., McEwen J., Williams J., & Papp E. (1981) The Nottingham Health Profile: Subjective Health Status and Medical Consultations. *Social Science and Medicine* **15,** 221-229.

128. Hunt S.M. (1997) The Problem of Quality of Life. *Quality of Life Research* **6,** 205-212.

129. Hyman H.H., Cobb W.J., Feldman J.J., Hart C.W., & Stember C.H. (1954) Interviewing in Social Research. University of Chicago Press, Chicago.

130. Iversen I.A., Silberberg N.E., Stever R.C., & et al. (1973) The Revised Kenny Self-Care Evaluation: A Numerical Measure of Independence In Activities of Daily Living. Sister Kenny Institute, Mineapolis, Minesota.

131. Jenkinson C., Bardsley M. & Lawrence K. (1994) Measurement in Subjective Health Assessment. In: *Measuring Health and Medical Outcomes* (ed Jenkinson C.) UCL Press, London.

132. Jenkinson C. (1995) Evaluating the Efficacy of Medical Treatment: Possibilities and Limitations. *Social Science and Medicine* **41,** 1395-1401.

133. Jennet B. & Bond M. (1975) Assessment of Outcome After Severe Brain Damage. *Lancet* **i,** 480-484.

134. Johnson G., Burvill P.W., Anderson C.S., Jamrozik K., Stewart-Wynne E.G., & Chakera T.M. (1995) Screening instruments for depression and anxiety following stroke: experience in the Perth community stroke study. *Acta Psychiatr.Scand* **91,** 252-257.

135. Jonkman E.J., deWeerd A.W., & Vrijens N.L. (1998) Quality of Life After a First Ischemic Stroke. Long-term Developments and Correlations With Changes in Neurological Deficit, Mood and Cognitive Impairment. *Acta Neurol Scand* **98,** 169-175.

136. Kaplan R.M., Bush J.W., & Berry C.C. (1976) Health Status: Types of Validity and the Index of Well-being. *Health Services Research* **11,** 478-507.

137. Karnofsky D.A., Abelmann W.H., & Craver L.F. (1948) The Use of Nitrogen Mustards in the Palliative Treatment of Carcinoma. *Cancer* **1,** 634-656.

138. Katz S. & Akpom C.A. (1976) A Measure of Primary Sociobiological Functions. *International Journal of Health Services* **6,** 493-507.

139. Kauhanen M.L., Korpelainen J.T., Hiltunen P., Brusin E., Mononen H., Maatta R., Nieminen P., Sotaniemi K.A., & Myllyla V.V. (1999) Poststroke Depression

Correlates with Cognitive Impairment and Neurological Deficits. *Stroke* **30,** 1875-1880.

140. Kelly-Hayes M. & Paige C. (1995) Assessment and psychologic factors in stroke rehabilitation. *Neurology* **45,** S29-S32.

141. Kertesz A. (1982) Western Aphasia Battery. Grune & Stratton, New York.

142. Kind P., Rosser R. & Williams A. (1982) Valuation of Quality of Life: Some Psychometric Evidence. In: *The Value of Life and Safety* (ed Jones-Lee M.W.), pp. 159-170. Elsevier North-Holland, Amsterdam.

143. King R.B. (1996) Quality of Life After Stroke. *Stroke* **27,** 1467-1472.

144. Kitamura T., Kijima N., Watanabe K., Takezaki Y., & Tanaka E. (1999) Precedents of perceived social support: personality and early life experiences. *Psychiatry Clin Neurosci* **53,** 649-654.

145. Kline Leidy N., Jonas D.L., Silberman C.C., Margolis M.K., & Heyes A. (2001) What Role Do Symptoms Play In Health Related Quality of Life. *Quality of Life Research* **10,** 213.

146. Knapp P. & Hewison J. (1999) Disagreement in Patient and Carer Assessment of Functional Abilities After Stroke. *Stroke* **30,** 938.

147. Knapp P., Young J., House A., & Forster A. (2000) Non-drug strategies to resolve psycho-social difficulties after stroke. *Age Ageing* **29,** 23-30.

148. Kotila M., Numminen H., Waltimo O., & Kaste M. (1998) Depression after Stroke: Results of the FINNSTROKE Study. *Stroke* **29,** 372.

149. Kwa V.I., Limburg M., & de Haan R. (1996) The Role of Cognitive Impairment in the Quality of Life After Ischaemic Stroke. *Journal of Neurology* **243,** 599-604.

150. Labi M.L.C., Phillips T.F., & Gresham G.E. (1980) Psychosocial Disability in Physically Restored Long Term Stroke Survivors. *Archives of Physical Medicine and Rehabilitation* **61,** 561-565.

151. LaPointe L.A. (1999) Quality of Life with Aphasia. *Seminars in Speech and Language* **20,** 5-17.

152. Lawrence L. & Christie D. (1979) Quality of Life After Stroke: A Three-year Follow-up. *Age Ageing* **8,** 167-172.

153. Lawton M.P. (1975) The Philadelphia Geriatric Centre Morale Scale: A Revision. *Journal of Gerontology* **1,** 89.

154. Lazarus R.S. & Folkman J. (1984) Stress Appraisal and Coping. Springer, New York.

155. Le Dorze G. & Brassard C. (1995) A Description of the Consequences of Aphasia on Aphasic Persons and their Relatives and Friends, Based on the WHO Model of Chronic Diseases. *Aphasiology* **9,** 239-255.

156. Linn M.W., Sculthorpe W.B., Evje M., & et al. (1969) Social Dysfunction Rating Scale. *Journal of Psychiatric Research* **6,** 299-306.

157. Lipsey J.R., Robinson R.G., Pearlson G.D., Rao K., & Price T.R. (1984) Nortriptyline treatment of post-stroke depression: a double-blind study. *Lancet* **1,** 297-300.

158. Lofgren B., Gustafson Y., & Nyberg L. (1999) Psychological Well-being 3 Years After Stroke. *Stroke* **30,** 567-572.

159. Lydick E. (2000) Approaches to the Interpretation of Quality of Life Scales. *Medical Care* **38,** 180-183.

160. Lyon J.G., Cariski D., Keisler L., Rosenbek J., Levine R., Kumpula J., Ryff C., Coyne S., & Blanc M. (1997) Communication Partners: Enhancing Participation in Life and Communication for Adults with Aphasia in Natural Settings. *Aphasiology* **11,** 693-708.

161. Maclean N., Pound P., Wolfe C., & Rudd A. (2000) Qualitative analysis of stroke patients' motivation for rehabilitation. *BMJ* **321,** 1051-1054.

162. Mahoney F.I., Wood O.H., & Barthel D.W. (1958) Rehabilitation of Chronically Ill Patients: The Influence of Complications on the Final Goal. *South Medical Journal* **51,** 605-609.

163. Marshall G. & Roberts S. (1996) Social Class and Underclass in Britain and the USA. *British Journal of Sociology* **47,** 22-44.

164. Mastekaasa A. & Moum T. (1984) The Perceived Quality of Life in Norway: Regional Variations and Contextual Effects. *Social Indicators Research* **14,** 385-419.

165. Mathias S.D., Bates M.M., Pasta D.J., Cisternas M.G., Feeny D., & Patrick D.L. (1997) Use of the Health Utilities Index with stroke patients and their caregivers. *Stroke* **28,** 1888-1894.

166. Maynard A. (1993) Requirements for health care purchasers. In: *Quality of Life Assessment: Key Issues in the 1990's* (ed Walker S.R.), pp. 413-426. Kluwer Academic Publishers, Dordrecht.

167. Mayou R. & Bryant B. (1993) Quality of Life in Cardiovascular Disease. *British Medical Journal* **69,** 460-466.

168. McDowell I. & Newell C. (1996) Measuring Health: A Guide to Rating Scales and Questionnaires, second edn. Oxford University Press, New York.

169. McHorney C. & Tarlov A. (1995) Individual-patient Monitoring In Clinical Practice: Are Available Health Surveys Adequate? *Quality of Life Research* **4,** 293-307.

170. 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.

171. Meenan R.F. (1982) The AIMS Approach to Health Status Measurement: Conceptual Background and Measurement Properties. *Journal of Rheumatology* **9,** 788.

172. Megone C. (1994) The Quality of Life. Starting from Aristotle. In: *The Quality of Life. Perspectives and Policies.* Routledge, London.

173. Menard S. (1995) Applied Logistic Regression Analysis. Sage, Thousand Oaks, CA.

174. Montgomery S.A. & Asberg M. (1979) A New Depression Scale Designed to Be Sensitive to Change. *British Journal of Psychiatry* **134,** 382-389.

175. Morris P.L.P., Robinson R.G., Raphael B., & Bishop D. (1991) The Relationships Between the Perception of Social Support and Post-Stroke Depression in Hospitalised Patients. *Psychiatry* **54,** 306-315.

176. Muldoon M.F., Barger S.D., Flory J.D., & Manuck S.B. (1998) What are quality of life measurements measuring? *BMJ* **316,** 542-545.

177. Muller D.J., Code C., & Mugford J. (1983) Predicting Psychosocial Adjustment to Aphasia. *British Journal of Disorders of Communication* **18,** 23-29.

178. Muthny F.A., Koch U., & Stump S. (1990) Quality of Life in Oncology Patients. *Psychotherapy and Psychosomatics* **54,** 145-160.

179. Neau J.P., Ingrand P., Mouille-Brachet C., Rosier M.P., Couderq C., Alvarez A., & Gil R. (1998) Functional Recovery and Social Outcome After Cerebral Infarction in Young Adults. *Cerebrovasc Dis* **8,** 296-302.

180. Nelson E., Wasson J., Kirk J., Keller A., Clark D., Dietrich A., Stewart A., & Zubkoff M. (1987) Assessment of function in routine clinical practice: description of the COOP Chart method and preliminary findings. *J Chronic Dis* **40 Suppl 1,** 55S-69S.

181. NHS Executive. Promoting Clinical Effectiveness: A Fremework for Action In and Through the NHS. 1996. Leeds, Department of Health. Ref Type: Report

182. NHS Executive. Clinical Governance in the new NHS. 1999. London, NHS Executive Quality Management Branch. Ref Type: Report

183. Niemi M.L., Laaksonen R., Kotila M., & Waltimo O. (1988) Quality of Life 4 Years After Stroke. *Stroke* **19,** 1101-1107.

184. Nunnally J.C. (1970) Introduction to Psychological Measurement. McGraw-Hill, New York.

185. Nunnally J.C. & Bernstein I.H. (1994) Psychometric Theory, 3rd edn. McGraw-Hill, New York.

186. O'Boyle C.A., McGee H., Hickey A., O'Malley K., & Joyce C.R. (1992) Individual quality of life in patients undergoing hip replacement. *Lancet* **339,** 1088-1091.

187. O'Mahony P.G., Thomson R.G., Dobson R., Rodgers H., & James O.F. (1999) The prevalence of stroke and associated disability. *J Public Health Med* **21,** 166-171.

188. Office of National Statistics (2000) Income and Lifestyles. *Regional Trends* **35,** 128.

189. Osberg J.S., DeJong G., Haley S.M., Seward M.L., McGinnis G.E., & Germaine J. (1988) Predicting Long-term Outcome among Post-rehabilitation Stroke Patients. *American Journal of Physical Medicine and Rehabilitation* **67,** 94-103.

190. Parr S., Byng S., & Gilpin S. (1997) Talking about Aphasia. Open University Press, Buchingham.

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

192. Patrick D.L. & Erickson P. (1993) Assessing Health-Related Quality of Life for Clinical Decision Making. In: *Quality of Life Assessment: Key Issues in the 1990's* (ed Walker S.R.), pp. 11-63. Kluwer Academic Publishers, Dordrecht.

193. Pearlman R.A. & Uhlmann R.F. (1988) Quality of Life In Chronic Diseases: Perceptions of Elderly Patients. *Journal of Gerontology* **43,** M25-M30.

194. Pearlman R.A. & Uhlmann R.F. (1991) Quality of Life in Elderly, Chronically Ill Outpatients. *Journal of Gerontology* **46,** M31-M38.

195. Pound C., Parr S., Lindsay J., & Woolf C. (2000) Beyond Aphasia: Therapies for Living With Communication Disability. Speechmark, Bicester, UK.

196. Pound P., Gompertz P., & Ebrahim S. (1994) Patients' Satisfaction with Stroke Services. *ClinicalRehabilitation* **8,** 7-17.

197. Pound P., Tilling K., Rudd A.G., & Wolfe C.D.A. (1999) Does Patient Satisfaction Reflect Differences in Care Received after Stroke. *Stroke* **30,** 49-55.

198. Radloff L.S. & Locke B. (1986) The Community Mental Health Assessment Survey and the CES-D Scale. In: *Community Surveys* (ed Weissman M.), pp. 177-189. Prodist Publisher, New York.

199. Raven J., Raven J.C., & Court J.H. (2000) Standard Progressive Matrices. Oxford Psychologists Press, Oxford.

200. Raven J.C. (1962) Coloured Progressive Matrices Sets A, Ab, B. Lewis, London.

201. Raven J.C., Court J.H., & Raven J. (1995) Coloured Progressive Matrices. Oxford Psychologists Press, Oxford.

202. Reding M.J., Orto L.A., Winter S.W., Fortuna I.M., Di Ponte P., & McDowell F.H. (1986) Antidepressant therapy after stroke. A double-blind trial. *Arch.Neurol* **43,** 763-765.

203. Robinson R.G., Murata Y., & Shimoda K. (1999) Dimensions of social impairment and their effect on depression and recovery following stroke. *Int Psychogeriatr.* **11,** 375-384.

204. Rogerson R.J. (1995) Environmental and Health-Related Quality of Life: Conceptual and Methodological Similarities. *Social Science and Medicine* **41,** 1373-1382.

205. Rose D. & O'Reilly K. (1997) Constructing Classes. Towards a New Social Classification for the UK. ESRC/ONS, Swindon.

206. Rose D. (1997) Recommendations to ONS on 2001 Census. In: *Constructing Classes. Towards a New Social Classification for the UK* (eds Rose D. & O'Reilly K.), pp. 178-182. ESRC/ONS, Swindon.

207. Rosenberg R.S. (1995) Health-Related Quality of Life Between Naturalism and Hermeneutics. *Social Science and Medicine* **41,** 1411-1415.

208. Rosenthal R. & Rosnow R.L. (1991) Essentials of Behavioral Research: Methods and Data Analysis, second edn. McGraw-Hill Publishing Company, New York.

209. Rosser R. & Kind P. (1978) A Scale of Valuations of States of Illness: Is There A Social Consensus? *International Journal of Epidemiology* **7,** 347-358.

210. Roth M., Tyme E., Mountjoy C.Q., Hupert F.A., Hendrie A., Verma S., & Goddard R. (1986) CAMDEX: A Standardised Instrument for The Diagnosis of Mental Disorders in The Elderly with Special Reference ot the Early Detection of Dementia. *British Journal of Psychiatry* **149,** 698-709.

211. Royal College of Physicians. National Clinical Guidelines for Stroke. Prepared by The Intercollegiate Working Party for Stroke. 2000. London, Royal College of Physicians. Ref Type: Report

212. Ruta D.A., Garratt A.M., Leng M., Russell I., & MacDonald L. (1994) A New Approach to the Measurement of Quality of Life:the Patient Generated Index (PGI). *Medical Care* **32,** 1109-1126.

213. Sackett D. (1979) Bias in Analytic Research. *Journal of Chronic Diseases* **32.**

214. Salonen T. (1995) Report of a Questionnaire Survey of Poststroke Patients With Aphasia and Their Families. *Topics in Stroke Rehabilitation* **2,** 72-75.

215. Sarason I.G. & Sarason B.R. (1982) Concomitants of social support: attitudes, personality characteristics, and life experiences. *J Pers.* **50,** 331-344.

216. Sarno J.E., Sarno M.T., & Levita E. (1973) The Functional Life Scale. *Archives of Physical Medicine and Rehabilitation* **54,** 214-220.

217. Sarno M.T. (1969) The Functional Communication Profile: Manual of Directions. Institute of Rehabilitation Medicine, New York.

218. Sarno M.T. (1997) Quality of Life in Aphasia in the First Post Stroke Year. *Aphasiology* **11,** 665-679.

219. Scarpa M., Colombo A., Sorgato P., & De Renzi E. (1987) The incidence of aphasia and global aphasia in left brain-damaged patients. *Cortex* **23,** 331-336.

220. Schwartz G.E. (1983) Development and Validation of the Geriatric Evaluation by Relatives Rating Instrument (GERRI). *Psychological Reports* **53,** 479-488.

221. Sherbourne C.D. & Stewart A.L. (1991) The MOS Social Support Survey. *Social Science and Medicine* **32,** 705-714.

222. Shinar D., Gross C.R., Heyman A., & et al. (1985) Inerobserver Variability in the Assessment of Neurological History and Examination in the Stroke Data Bank. *Archives of Neurology* **42,** 557-565.

223. Siegrist J. & Junge A. (1990) Measuring the Social Dimensions of Subjective Health in Chronic Illness. *Psychotherapy and Psychosomatics* **54,** 90-98.

224. Sinclair A. & Dickinson E. (1998) Effective Practice in Rehabilitation: The Evidence of Systematic Review. King's Fund, London.

225. Singleton R.A. & Straits B.C. (1999) Approaches to Social Research, third edn. Oxford University Press, Oxford.

226. Sloan J.A. (2000) Methods for Assessing Clinical Significance in Quality of Life Measurement. *Quality of Life Research* **9,** 246.

227. Sloan J.A., Cella D., Frost M., Guyatt G.H., Sprangers M., & Symonds T. (2002) Assessing clinical significance in measuring oncology patient quality of life: introduction to the symposium, content overview, and definition of terms. *Mayo Clin Proc.* **77,** 367-370.

228. Smits C.H.M., Smit J.H., van den Heuvel N., & Jonker C. (1997) Norms for an Abbreviated Raven's Coloured Progressive Matrices in an Older Sample. *Journal of Clinical Psychology* **53,** 687-697.

229. Sneeuw K.C.A., Aaronson N.K., de Haan R.J., & Limburg M. (1997) Assessing Quality of Life After Stroke. The Value and Limitations of Proxy Ratings. *Stroke* **28,** 1541-1249.

230. Spitzer W.O., Dobson A.J., Hall J., Chesterman E., Levi J., Shepherd R., Battista R.N., & Catchlove B.R. (1981) Measuring the quality of life of cancer patients: a concise QL-index for use by physicians. *J Chronic Dis* **34,** 585-597.

231. Sprangers M.A.G. & Aaronson N.K. (1992) The Role of Health Care Providers and Significant Others in Evaluating the Quality of Life of Patients with Chronic Disease: A Review. *Journal of Clinical Epidemiology* **45,** 743-760.

232. SPSS Inc. (1999) SPSS Base 10.0 Applications Guide. SPSS Inc., Chicago.

233. Staniszewska S. (1999) Patient expectations and Health Related Quality of Life. *Health Expectations* **2,** 93-104.

234. Staniszewska S. & Ahmed L. (1999) The concepts of expectation and satisfaction: do they capture the way patients evaluate their care? *J Adv Nurs* **29,** 364-372.

235. Stevens J.P. (1992) Applied Multivariate Statistics for the Social Sciences, 2nd edn. Erlbaum, Hillsdale, NJ.

236. Stewart J.A., Dundas R., Howard R.S., Rudd A.G., & Wolfe C.D. (1999) Ethnic differences in incidence of stroke: prospective study with stroke register. *BMJ* **318,** 967-971.

237. Streiner D.L. & Norman G.R. (1995) Health Measurement Scales. A Practical Guide to Their Development and Use. Oxford University Press, New York.

238. Tabachnick B.G. & Fidell L.S. (2001) Using Multivariate Statistics, 4th edn. Allyn and Bacon, Boston.

239. Tartaglione A., Inglese M.L., Bandini F., Spadavecchia L., Hamsher K., & Favale E. (1991) Hemisphere Asymmetry in Decision Making Abilities. *Brain* **114,** 1441-1456.

240. Thelander M.J., Hoen B., & Worsley J. Report on the Evaluation of Effectiveness of a Community Program for Aphasic Adults. 1994. York-Durham Aphasia Centre. Ref Type: Report

241. Thompson J. & Enderby P.M. (1983) Is All your Schuell Really Necessary? *British Journal of Disorders of Communication* **14,** 195-201.

242. Torrance G.W., Thomas W.H., & Sackett D.L. (1972) A utility maximization model for evaluation of health care programs. *Health Serv.Res* **7,** 118-133.

243. Torrance G.W., Boyle M.H., & Horwood S.P. (1982) Application of multi-attribute utility theory to measure social preferences for health states. *Oper.Res* **30,** 1043-1069.

244. Trigg R., Wood V.A., & Langton Hewer R. (1999) Social Reintegration after Stroke: The First stages in the Development of the Subjective Index of Physical and Social Outcome (SIPSO). *Clinical Rehabilitation* **13,** 341-353.

245. Tuomilehto J., Nuottimaki T., Salmi K., Aho K., Kotila M., Sarti C., & Rastenyte D. (1995) Psychosocial and Health Status in Stroke Survivors After 14 Years. *Stroke* **26,** 971-975.

246. Turner R.J. & Noh S. (1983) Class and Psychological Vulnerability Among Women: The Significance of Social Support and Personal Control. *Journal of Health and Social Behavior* **24,** 2-15.

247. Turner R.R. (1990) Rehabilitation. In: *Quality of Life Assessment in Clinical Trials* (ed Spilker B.), pp. 247-267. Raven, New York.

248. van Straten A., de Haan R.J., Limburg M., Schuling J., Bossuyt P.M., & van den Bos G.A.M. (1997) A Stroke-Adapted 30-Item Version of the Sickness Impact Profile to Assess Quality of Life (SA-SIP30). *Stroke* **28,** 2155-2161.

249. Vieweg B.W. & Hedlund J.L. (1983) The General Health Questionnaire (GHQ): A Comprehensive Review. *Journal of Operational Psychiatry* **14,** 74-81.

250. Viitanen M., Fugl-Meyer K.S., Bernspaang B., & Fugl-Meyer A.R. (1988) Life Satisfaction in Long-term Survivors After Stroke. *Scandinavian Journal of Rehabilitation Medicine* **20,** 17-24.

251. Villardita C. (1985) Raven's Colored Progressive Matrices and Intellectual Impairment in Patients with Focal Brain Damage. *Cortex* **21,** 627-634.

252. Wade D.T., Legh-Smith J., & Langton Hewer R. (1985) Social Activities After Stroke: Measurement and Natural History Using the Frenchay Activities Index. *Int Rehabil Med* **7,** 176-181.

253. Wade D.T., Hewer R.L., David R.M., & Enderby P.M. (1986) Aphasia after stroke: natural history and associated deficits. *J Neurol Neurosurg.Psychiatry* **49,** 11-16.

254. Ware J.E., Harris W.J., Gandek B., Rogers B.W., & Reese P.R. (1997) MAP-R for Windows: Multitrait/Multi-item Analysis Program- Revised User's Guide Version 1. Health Assessment Lab., Boston,MA.

255. Ware J.E. & Guyatt G.H. (2001) Generic versus Disease-Specific Measures. Presentation at the ISOQOL Annual Conference. Amsterdam, The Netherlands.

256. Ware J.E., Jr., Snow K.K., & Kosinski M. (1993) SF-36 Health Survey: Manual and Interpretation Guide. The Health Institute, New England Medical Center, Boston, MA.

257. Weinstein M.C. & Stason W.B. (1976) Hypertension: A Policy Perspective. Harvard University Press, Cambridge, MA.

258. Wenger N.K., Mattson M.E., Furberg C.D., & Elison J. (1984) Assessment of Quality of Life in Clinical Trials of Cardiovascular Therapies. *American Journal of Cardiology* **54,** 908-913.

259. WHOQOL Group (1993) Measuring Quality of Life: The Development of the World Health Organisation Quality of Life Instrument (WHOQOL). WHO, Geneva.

260. WHOQOL Group (1995) The World Health Organisation Quality of Life Assessment (WHOQOL): Position Paper from the World Health Organisation. *Social Science and Medicine* **41,** 1403-1409.

261. WHOQOL Group (1998) The World Health Organisation Quality of Life Assessment (WHOQOL): Development and General Psychometric Properties. *Social Science and Medicine* **46,** 1569-1585.

262. Wilkinson P.R., Wolfe C.D.A., Warburton F.G., Rudd.A.G., Howard R.S., RossRussell R.W., & Beech R. (1997) Longer Term Quality of Life and Outcome in Stroke Patients: Is the Barthel Index Alone an Adequate Measure of Outcome? *Quality in Health Care* **6,** 125-130.

263. Williams A. (1993) The importance of quality of life in policy decisions. In: *Quality of Life Assessment: Key Issues in the 1990's* (ed Walker S.R.), pp. 427-439. Kluwer Academic Publishers, Dordrecht.

264. Williams L.S., Weinberger M., Harris L.E., Clark D.O., & Biller H. (1999) Development of a Stroke-Specific Quality of Life Scale. *Stroke* **30,** 1362-1369.

265. Williams L.S., Weinberger M., Harris L.E., & Biller H. (1999) Measuring Quality of Life In a Way That Is Meaningful to Stroke Patients. *Neurology* **53,** 1839-1843.

266. Willms D.G. & Johnson N.A. Essentials in Qualitative Research: A Notebook for the Field. 1993.
Ref Type: Unpublished Work

267. Wolfe C.D., Taub N.A., Woodrow J., Richardson E., Warburton F.G., & Burney P.G. (1993) Patterns of acute stroke care in three districts of southern England. *J Epidemiol.Community Health* **47,** 144-148.

268. Wolfe C.D., Taub N.A., Woodrow J., Richardson E., Warburton F.G., & Burney P.G. (1993) Does the incidence, severity, or case fatality of stroke vary in southern England? *J Epidemiol.Community Health* **47,** 139-143.

269. Wolfe C.D., Taub N.A., Bryan S., Beech R., Warburton F., & Burney G.J. (1995) Variations in the incidence, management and outcome of stroke in residents under the age of 75 in two health districts of southern England. *J Public Health Med* **17,** 411-418.

270. Wood-Dauphinee S. & Williams J.I. (1987) Reintegration to Normal Living As A Proxy To Quality of Life. *Journal of Chronic Disease* **40,** 491-499.

271. Wood-Dauphinee S. (1999) Assessing quality of life in clinical research: from where have we come and where are we going? *J Clin Epidemiol.* **52,** 355-363.

272. Woodward C.A. & Chambers L.W. (1991) Guide to Questionnaire Construction and Question Writing. Canadian Public Health Association, Ottawa.

273. World Health Organisation (1980) International Classification of Impairments, Disabilities and Handicaps: A Manual of Classification Relating to the Consequences of Disease. World Health Organisation, Geneva, Switzerland.

274. Wright J.G. & Young N.L. (1997) A Comparison of Different Indices of Responsiveness. *Journal of Clinical Epidemiology* **50,** 239-246.

275. Wyller T.B., Holmen J., Laake P., & Laake K. (1998) Correlates of Subjective Well-being in Stroke Patients. *Stroke* **29,** 363-367ss.

276. Zarit S.H., Reever K.E., & Bach-Peterson J. (1980) Relatives of the Impaired Elderly: Correlates of Feelings of Burden. *The Gerontologist* **20,** 649-655.

277. Zellars K.L. & Perrewe P.L. (2001) Affective personality and the content of emotional social support: coping in organizations. *J Appl.Psychol.* **86,** 459-467.

278. Zigmond A.S. & Snaith R.P. (1983) The hospital anxiety and depression scale. *Acta Psychiatr.Scand* **67,** 361-370.

279. Zung W.W.K. (1965) A Self-Rating Depression Scale. *Archives of General Psychiatry* **12,** 67-70.