TRANSLATION, CULTURAL ADAPTATION AND REVALIDATION OF THE REINTEGRATION TO NORMAL LIVING (RNL)-INDEX FOR USE IN SPAIN

Ana María Rodríguez, B.Sc. (P.T.)
School of Physical and Occupational Therapy
Faculty of Medicine
McGill University, Montreal, Quebec, Canada
December, 2007

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Science in Rehabilitation Sciences

Your file Votre référence
ISBN: 978-0-494-51331-6
Our file Notre référence
ISBN: 978-0-494-51331-6

ABSTRACT

The "International Classification of Functioning, Disability and Health" (ICF) is a model that classifies health and health-related function. Of the ICF domains, least is known about participation. Although a few measures have been developed in English that tap the construct "Participation", none exist in Spanish. The Reintegration to Normal Living (RNL)-Index is the measure of reference to evaluate participation. The general objective of the present study was to translate, culturally adapt, and assess the psychometric properties the RNL-Index in Spanish for use in Spain. The translation and the cultural adaptation of the RNL-Index consisted of a five-step process, leading to the Spanish version of the RNL-Index (SRNL-Index). Psychometric evaluation consisted of a cross-sectional study design, with a longitudinal design used for test-retest evaluation. Thirty-two subjects with stroke and 36 with total knee arthroplasty completed the SRNL-Index twice, as well as the Six Minute Walk Test (6MWT), the Short-Form 36 Health Survey (SF-36), and for stroke subjects, the Barthel Index (BI). Descriptive statistics, one-way analysis of variance and post-hoc t tests were calculated, as well as Cronbach's alpha, Pearson's, and Intraclass correlation coefficients. The SRNL-Index was found to be internally consistent. Test-retest reliability was quite poor, being moderately low for the TKA sample and moderate for the stroke sample. Discriminant validity was demonstrated by the correlations between the SRNL-Index's Perception of Self and Daily Activity subscales with the Mental and Physical Component Summary scores of the SF-36. The SRNL-Index has convergent validity as shown by the high correlations between of the SRNL-Index and the 6MWT. Known-groups validity was shown in people with stroke grouped by Barthel Index scores, and in both samples when grouped according to gait speed. The SRNL-Index demonstrated acceptable validity and internal reliability for subjects with stroke and total knee arthroplasty. However, further studies are needed to reassess external validity, reliability, and responsiveness in other populations and other Spanish-speaking countries.

ABBRÉGÉ

La « Classification internationale du fonctionnement, du handicap et de la santé » (CIF) est un modèle qui classifie les composantes de la santé et de certains éléments du bien-être connexes à la santé. De tous les domaines de la CIF, le moins connu est celui de la participation. Quelques mesures ont été développées, principalement en anglais, pour évaluer le domaine de la « participation », mais aucune n'existe en espagnol. L'indice de réintégration d'une vie normale (Indice RVN) est une mesure de référence de la participation. L'objectif général de cette étude était de traduire, d'adapter culturellement et d'évaluer les propriétés psychométriques de la version espagnole de l'Indice RVN pour un usage en Espagne. La traduction et l'adaptation culturelle de l'Indice RVN consista en un processus de cinq étapes qui résulta en la version espagnole de l'Indice RVN (Indice RVNE). Les études psychométriques étaient de type transversal, excepté pour la fiabilité test-retest, qui était de type longitudinal. Trente-deux sujets avec un accident vasculaire cérébral (AVC) et trente-six avec une arthroplastie du genou (AG) complétèrent l'Indice RVNE à deux reprises, ainsi que le test de marche de six minutes (TM6M), le « Short Form 36 Health Survey » (SF-36) et pour les sujets avec un AVC, l'Indice de Barthel (IB). Les statistiques descriptives, les analyses de variance à un seul facteur suivis de tests post-hoc furent calculés, ainsi que les coefficients alpha de Cronbach et de correlation intraclasse. L'Indice RVNE démontra avoir une consistance interne. La fiabilité test-retest démontra être plutôt pauvre dans l'échantillon des AG et modérée dans celle des AVC. Une analyse discriminante de la validité fût démontrée à l'aide de corrélations entre les échelles de la Perception de Soi et des Activités Quotidiennes de l'Index RVNE et les ponctuations sommaires des composantes Physiques et Mentales du SF-36. La validité convergente de l'Indice RVNE fût démontrée par les corrélations élevées entre l'Index RVNE et le TM6M. Une validité des groupes connus fût démontrée dans l'échantillon des AVC lorsque regroupés selon leurs scores d'IB et dans les deux échantillons lorsque groupés selon la vitesse de marche. Ainsi, l'Indice RVNE possède une validité interne et une fiabilité interne acceptables pour des sujets avec un AVC ou une arthroplastie du genou. Par contre, des études additionnelles sont nécessaires pour évaluer la validité, la fiabilité et la sensibilité au changement avec des conditions et des pays de langue hispanophone additionnels.

ACKNOWLEDGEMENTS

First and foremost, I wish to thank my supervisor, Dr. Sharon Wood-Dauphinee, for her guidance, encouragement, support, and assistance with this thesis. She inspired me to wish to learn more and pursue my studies in the realm of quality of life research. I also wish to thank my co-supervisor, Dr. Nancy Mayo, for her guidance, her enthusiasm, and her love of good research, which encouraged me to want to pursue a research career. I also owe thanks to Dr. Jordi Alonso and Dr. Montserrat Ferrer of the Institut Municipal d'Investigació Mèdica for the warm welcome to their institution, for their guidance and assistance, and for enabling me to participate in unique learning opportunities.

I would like to thank the translators involved in this project, Carlos Primo and Patricia Monroy Córdoba as well as the backtranslators, Margaret Roper, and Albert Sanchez. Special thanks must be given to Dr. Roser Belmonte, Cristina Cervantes, and Nohora Rueda Moreno, without whom the project could not have been realised at the Hospital de l'Esperança.

Additionally, I would like to thank Dr. Lois Finch, Lynne Nadeau, Susan Scott, and Gemma Vilagut for their assistance with the statistical analyses and encouragement along the journey of learning and using the SAS statistical software.

I would like to thank the subjects who gave their time and effort during the evaluation process of the translated questionnaire. Without their cooperation, this project could not have been realised.

Thanks are also due to the students and research personnel of the Division of Clinical Epidemiology of the Royal Victoria Hospital, as well as those of the Health Services Research Unit of the Institut Municipal d'Investigació Mèdica, for their warm welcome into their department and for their encouragements and advices.

Finally, I would like to thank my family and Sami Tali for their support during the course of this project. The expenses for this study were self-funded.

TABLE OF CONTENTS

INDEX OF TABLES

INDEX OF FIGURES

PREFACE

In 1987, the Reintegration to Normal Living (RNL)-Index was published[2]. It was developed as part of the study "Evaluation of Factors Affecting Reintegration of Cancer patients to Normal Working Patterns" that was funded by the National Cancer Institute of Canada in 1982. The RNL-Index was intended to complement the Quality of Life (QL)-Index previously developed and psychometrically tested by Spitzer and colleagues[3]. Both measures were designed to move assessment beyond the traditional areas of morbidity and mortality to incorporate how people felt and were able to function in their daily lives. These two measures, originally developed for use with people who had cancer, now have a much broader application.

Developed through the use of interviews with health professionals, subjects with different diagnoses and their significant others, the RNL-Index includes items reflecting mobility (indoor, community and long distance), self care, daily activities (work, school), social and recreational pursuits, coping skills, and interpersonal interactions and relationships[2]. It has been tested in various settings and is reliable, valid and responsive[2, 4-7]. Over the years, requests to use this measure were received by the principal developer primarily from researchers in the rehabilitation community. But, it was not until 2001, when the World Health Organization (WHO) published the International Classification of Functioning, Disability and Health (ICF)[8], that requests started to rise dramatically.

As will be more fully described later in the thesis, the ICF is one of the WHO families of classifications of health and classifies health and health-related function in domains that describe Body functions and structures, Activities of the person and Participation in life situations. Since an individual's functioning and disability occurs in a context, the ICF also includes a set of Environmental Factors, and acknowledges the contribution of Personal Factors to health. Collectively, these domains constitute a biopsychosocial model reflecting the body, the person and involvement in society[1]. In accordance with this model, "functioning" is an overarching term encompassing body structures and functions, activities and participation, while "disability" is the term used for impairments, activity limitations or participation restrictions[1]. The elements in the ICF model of

functioning and disability constantly influence and modify each other[9]. Following a disabling condition, the combination of impairments, activity limitations, personal and environmental factors affect an individual's ability to resume normal living patterns. A restriction in participation affects a person's ability to reintegrate his or her various roles in society and consequently, promotes complications, affecting mood and other aspects of the individual, leading to further disability, activity limitation, and participation restriction. With this model, the type of opportunities and resources necessary for well-being vary across individuals depending on their life-circumstances and personal characteristics[10].

The ICF model provides a common language for describing health conditions and their impact in terms of functioning, health, and disability. It facilitates communication and exchange between health care practitioners, researchers, and affected individuals at an international level, regardless of personal and educational backgrounds or of social and cultural factors. As international collaborations become increasingly important in the clinical and research settings, the ICF could provide common theoretical framework that facilitates exchange of thoughts. A challenge remains: the measures used to assess the theoretical constructs within the ICF need to be adapted both linguistically and culturally for true exchange to take place.

Contribution of Authors

There were many steps in the development of this thesis. A protocol was written by Ana María Rodríguez, assisted by Sharon Wood-Dauphinee and Nancy Mayo. Assistance for the study to take place at the involved institutions and for ethical approval was provided by Jordi Alonso and Montserrat Ferrer. Data collection, development of databases, and statistical analyses were done by Ana María Rodríguez. The manuscript was written by Ana María Rodríguez with editing by Sharon Wood-Dauphinee.

Organization of the thesis

This manuscript-based thesis is organized into 5 chapters. The first chapter is divided into two distinct areas of interest. The first section contains an introduction to the translation

and cultural adaptation of measures, as well as an overview of the guidelines that have been suggested for this process. The second section addresses Patient-Reported Outcomes (PROs), introduces Participation as the construct of interest, and explores several PRO measures of this construct.

Chapter 2 provides a review of the Reintegration to Normal Living (RNL)-Index describing its development and its psychometric properties. It ends with a general rationale for the Spanish cultural adaptation of the RNL-Index.

Chapter 3 focuses on the research question and states the global objectives of the study as well as the specific objectives and hypotheses of the project.

Chapter 4 presents the manuscript within this thesis. It is formatted according to the requirements of the "Journal of International Rehabilitation" and includes a background section, a statement of the rationale and main objectives of the study, a description of the methods and analyses employed. The manuscript then summarizes the results obtained and discusses them. This chapter ends by presenting the limitations of the study and the main conclusions obtained. Acknowledgements are finally given, followed by the results tables and the appendices for the manuscript.

Finally, Chapter 5 summarizes the findings and the conclusions of the thesis, describes its limitations, and suggests areas of further research.

The appendices contain information of interest.

Due to the manuscript format and in order to follow the regulations of the Graduate and Postdoctoral Studies of McGill University, there is some duplication of material throughout the thesis.

CHAPTER 1

1.1 Translation and Cultural Adaptation of Measures

1.1.1 The need for translation and cultural adaptations in health

The number of international research projects and collaborations in clinical research has grown exponentially over the past two decades. Arising from these international collaborations has been the development and use of health-related quality of life (HRQOL) measures. Cultures vary in terms of disease expression, in their view of HRQOL and in its role in evaluating health care. Hence, there has been a need to adapt health-related measures into different languages and cultures[11-13]. Even within countries, translations of widely used measures in studies of health are required[12]. The various measures simply can not be assumed to be equivalent across cultures[11].

Readying measures for international use has followed three approaches: parallel, simultaneous, or sequential[14]. In the parallel approach, used for example in the development of the EQ-5D, a standardized instrument for the measure of health outcomes, the multinational group EuroQol agreed on the construct and the items, and then translated them into several languages. In the simultaneous approach, a multinational group agrees on a construct and then identifies common elements to form the core of the cross-cultural instrument; each country representative then creates its own items[11]. This method was employed during the development of the World Health Organisation Quality of Life measure (WHOQOL)[11]. With few exceptions, the majority of the measures were initially developed in English[11, 12], and thus the sequential approach is used most often. In this approach, the measure is first developed in one language and culture, and is later translated and culturally adapted in a different language. Only once a measure has been used and recognised as being valid, reliable and responsive, and has a role in the clinical or research setting, does the need for a cultural adaptation manifest itself. The sequential approach was employed in the adaptation of the Short Form (SF)-36, the Nottingham Health Profile and the Sickness Impact Profile[15, 16]. A risk of this approach is that a questionnaire developed in one country may require such extensive adaptation when used

in another country that the resulting measure bears little resemblance to the original instrument.

Notwithstanding the above, the translation and adaptation of a measure to the cultural context of another population has several advantages: "it provides a common measure for the investigation of HRQOL within different contexts; it offers a standard measure for use in international studies; it allows comparisons between national and cultural groups relying on a standard measure designed and adapted to measure the phenomenon cross-culturally; and it allows the inclusion of immigrants avoiding the frequent bias of representing only the dominant culture of the country"[13] (p.1429). Adapting a measure developed in a different language and culture is more time and cost-effective than developing a new one. Nonetheless, the cross-cultural adaptation of HRQOL measures requires careful attention, involves numerous people and is still time consuming[13].

Carefully translated measures facilitate 'culturally sensitive research', an approach to research which considers cultural and language issues in all phases of research and assumes non-equivalence of concepts and measures between cultures[17]. Equivalence is defined as "the unbiased measurement between two instruments such that any differences detected are the result of true differences between the groups being assessed and not the result of differences inherent in the measurement tool used to gather the data."[17] (p.213). Hence, if equivalence between the original source and the translated version is reached, it will help maintain the content validity of the instrument across different cultures[12]. Equivalence is necessary to allow the comparison of patient-reported outcome results using the same instrument in different cultural groups. As "it is impossible to attain 100% equivalence, the goal is to minimize bias and come as close to this level of equivalence as possible"[17] (p. 214). A correct adaptation requires a broad design that takes into account not only the linguistic but also the technical and conceptual aspects involved in measuring the construct of interest[18].

To acquire equivalence, the cross-cultural adaptation of a health measure necessitates not only a translation, but also an entire re-evaluation of its psychometrics properties in the

new culture in which the measure is intended to be used. Considerable efforts have been made to standardize the evaluation of the psychometric properties of instruments (reliability, validity, and responsiveness) in the cross-cultural adaptation process[19]. These procedures ensure that versions adapted for use in different languages and cultural settings are as equivalent to the original questionnaires as possible.

1.1.2 Guidelines for the translation and revalidation of measures

Different methods attempted to define the best procedure to follow for the cultural adaptation process. An example for the cross-cultural adaptation of HRQOL measures was suggested by Guillemin and colleagues[13]. After a review of previous psychology and sociology literature and of published methodological frameworks, they suggested a set a standardized guidelines. They defined cross-cultural adaptation as having two components, one the translation of the measure, and the other its adaptation, or the combination of the "literal translation of words with the adaptation depending on the idiom, the cultural context, and the lifestyle"[13] (p. 1421). These guidelines include five steps: translation, back-translation, committee review, pre-testing for equivalence and re-examination of the weighting of scores if necessary.

The first step is the translation into the target language and is considered best when done by at least two independent translators who should translate the measure into their mother tongue. Ideally, one translator should know the measure and its objectives, while the other should not be aware of them[13]. For the back-translation step, or the translation from the target to the original language, the guidelines suggest that as many back-translations as the number of forward translations should be done independently of each other. Additionally, back-translators should translate the measure into their mother tongue. In this step, it is recommended that the translators should be aware of the objectives and concepts of the measure. Next, a bilingual committee with expertise in both the target diseases and the concepts of the measure should compare the original and the translated versions to produce a reconciled version. The goals at this stage are to modify or reject inappropriate items, to generate new items, to ensure that the translation is fully comprehensible at the level of a 10 to 12 year-old child, and to verify the cross-cultural

equivalence of the original and translated version. At this point, the committee could decide whether to repeat the translation and back-translation process for problematic items. A decentring technique, where both the original and the translated version of the measure can be modified during the translation procedure, is sometimes appropriate if done in close collaboration with the developers of the measure. The last step is to verify the equivalence between the translated and original versions through a pre-test. In this step, a sample population replies to the questionnaire to check for errors and problems with the translation. Guillemin and colleagues suggest using either a probe technique for this step, where subjects complete the measure and are then encouraged to explain the meaning of each item, or using a bilingual comparison, where the source and final version can be submitted to bilingual subjects to detect possible discrepancies[13]. Finally, when applicable, the response options for scoring should be considered to be adapted to the cultural context, as the response options from the original measure might not apply to the new cultural situation.

Another set of guidelines for the cultural adaptation process was developed by the investigators in the International Quality of Life Assessment (IQOLA) project (1991-1995). This group developed a general three-step process to produce cross-culturally comparable translations of patient-reported measures, in particular the Short Form-36 item (SF-36) Health Survey[20]. These steps included the forward-backward translation and the cross-cultural adaptation, the verification of the scaling assumptions, and the examination of the normative values for the new version[12, 17, 21]. The IQOLA group recommended that its first general step, the translation and cultural adaptation step, should consist of five stages: the translation, the synthesis, the back-translation, the expert committee review and the pre-testing. The recommendations by the IQOLA group are similar to those suggested by Guillemin and colleagues. A difference between the two groups includes an additional step prior to the back-translation, where the IQOLA group suggests that two translators and a recording observer synthesize the previous translations. The procedures for the psychometric evaluation and normative data are further outlined by Gandek and Ware in several publications[17, 21].

Another group with a unique approach to the cultural adaptation process is the Functional Assessment of Chronic Illness Therapy (FACIT) group. They suggested a universal translation approach, where "a single translation (is) developed that can be appropriate for different regions where it is intended to be used"[17] (p. 214). The FACIT group prefers the universal translation as the committee approach can be problematic, due to group pressure for premature agreement on translations[17, 22]. The authors claim that in the universal translation approach, translators have maximum independence and retain their objectivity by repeated input and modification. The translators originating from each major country where the language is spoken decide what terminology is most likely understood, avoiding colloquial phrasing[17]. This method is considered more rigorous than the double-back translation method[17]. The final translated version is then used across countries where the same language is spoken[17] and has the advantage that less bias is introduced when using the same translation across cultural groups than if applying country-specific versions. Besides the universal translation approach, most of the steps suggested are similar to the methodologies presented above. The suggested cultural adaptation process by the FACIT group consists of five stages: the translation by two independent professional translators; the reconciliation of the translations by a third independent translator; the back-translation into the original language; the analysis of the forward and back-translations, and reconciliations by three independent reviewers until consensus is reached; and finally, the pre-test by five subjects. One difference from the other methodologies is that only one back-translation is deemed necessary, as the FACIT group considered that the improvement from a second back-translation compared to one was not justified[17]. Their suggested pre-testing includes retrospective debriefing interviews, where the subjects are asked if items were difficult to understand, were irrelevant or offensive, and if additional topics needed to be added, or a cognitive debriefing technique, where again the subjects were asked, on an item-to-item basis, to restate the item in their own words and to interpret items that were found to be difficult in the translation process[17].

The ISPOR task force for Translation and Cultural Adaptation is yet another group that provided guidelines for the translation and cultural adaptation process for patient-reported measures[23]. Again, there are many similarities to those above. This group performed a

careful review of 12 major sets of previously published guidelines for translations and cultural adaptations of measures, including those from the American Association of Orthopaedic Surgeons, the FACIT group, the IQOLA group and the World Health Organization. Some translation steps across the guidelines were considered consistent, while others had considerable differences[23].

The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Task Force suggested ten steps in the cultural adaptation process: the preparation, where a "project manager" would obtain permission and explore the meaning of the concepts with the developer, as well as recruit the necessary people involved; the forward translation by at least two independent translators; the reconciliation step, where the key in-country persons work with the forward translators and the "project manager"; the literal, rather than conceptual back-translation; the back-translation review; the harmonization step; a cognitive debriefing step by five to eight respondents, native speakers of the target language in the in the target country; the review of the cognitive debriefing results against the original version by the project manager; the proofreading; and the final report that provides the description for all translation and cultural adaptation decisions[23]. What sets this group aside from the others is that the ISPOR task force realized that there was more than one way to achieve a particular step, as in the literature review, there was little agreement found in the recommendations provided. The ISPOR group considers that applying specific guidelines might be impossible in certain situations. An example is the harmonization step, where some guidelines suggest a harmonization meeting, when everybody involved meets and compares the translation back to the original, while others advocate including a harmonization component within each step of the translation process[23]. The ISPOR task force does not support a single approach to harmonization, as harmonization meetings might be difficult or even impossible to conduct[23]. Similarly, in the cognitive debriefing step, the ISPOR task force decided not to impose strict guidelines as it wanted them to remain practical. The flexibility offered by these guidelines is particularly aimed at private industry. Indeed, the need for cultural adaptations has arisen predominantly in the pharmaceutical industry. It is the use of patient-reported outcomes in pharmaceutical clinical trials, where the subject is required to evaluate the effects and

consequences of a pharmaceutical treatment on the aspects of his or her health related quality of life, and the development of measures for that specific purpose in clinical trials, which has rendered the need for cultural adaptation of measures a necessity.

1.2 Patient-Reported Outcomes as Measures of Participation

1.2.1 Patient-Reported Outcomes

Patient-reported outcomes (PROs), defined as the measurement of any aspect of a patient's health status coming directly from persons whose lives have been affected by illness, injury or its treatments, are becoming increasingly important in helping health professionals determine the impact of their interventions[24]. Not only are PROs answered directly by patients, but also by family members and expert health care professionals, who often contribute to the content of these measures. This means that PROs directly assess the patients' perceptions of the impact of the condition and its treatment in areas that are of uttermost importance to them.

While clearly not the only approach to assessing the impact of treatment, the use of PROs has added value under certain conditions[25], for example when there are no objective markers of symptoms or their impact, such as in sexual dysfunction or insomnia. Moreover, PROs add additional information to traditional clinical endpoints in conditions for which the primary goal of treatment is to maintain or improve function rather than prolong survival as in the case of cerebrovascular, neurological, cardiac, rheumatological and orthopaedic conditions. PROs also complement traditional outcomes by characterizing the effects of treatments for practionners. Because they reflect the patient's point of view, they may also facilitate the involvement of patients in making decisions about their own care. The completion of a PRO measure by a patient helps him or her reflect on health-related events and their impact on daily living. The responses may suggest new areas of discussion between the patient and the practionner thereby enhancing communication, and the information may help monitor the course of the disease and suggest revisions to treatment plans.

A PRO instrument can be used to measure the impact of an intervention on a single or multiple aspects of patients' health status, ranging from symptoms to more complex concepts. Examples include measures of symptoms, functional status, health status, well being, and integration or reintegration as well as adherence to and satisfaction with prescribed treatment and the intrusiveness of care. As well, multidimensional concepts with physical, psychological, and social components, such as quality of life and HRQOL, are often measured with PROs. These are certainly the most widely known PROs, but many other complex health-related constructs are assessed in this way. Of particular interest are the components of the ICF, where several PROs have been developed for its individual components, environmental factors or personal factors, body functions and structures, activities, or participation in life situations.

1.2.2 Participation

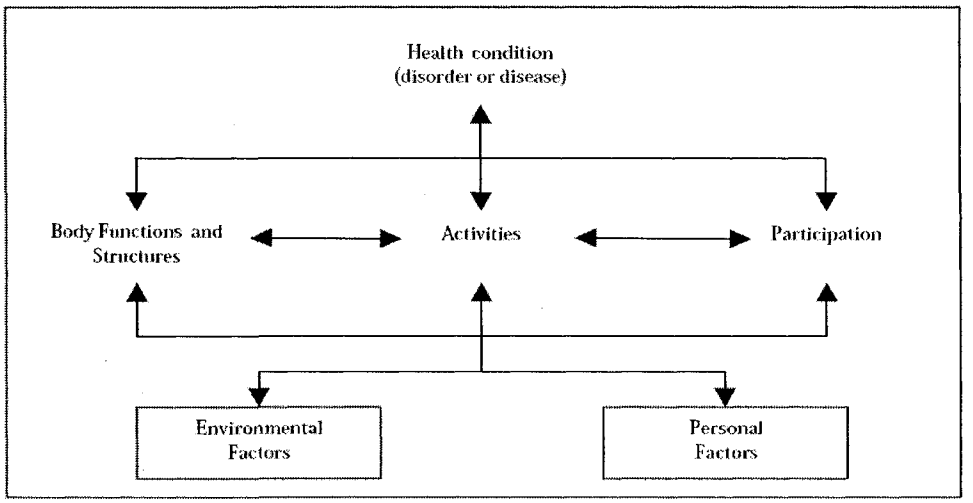
As mentioned previously, the ICF suggests that aspects of a person's Environmental and Personal factors influence how health-related function and dysfunction manifests itself in terms of Body functions and structures, Activities of the person and Participation in life situations. Collectively, these domains constitute an interacting biopsychosocial model reflecting the body, the person and involvement in society (Figure 1.1).

The definition of participation recently given in the ICF classification is "involvement in life situations". This is elaborated in a footnote: "the definition of Participation brings in the concept of involvement. Some proposed definitions of involvement incorporate taking part, being included or engaged in an area of life, being accepted or having access to needed resources"[1].

Participation can take the form of interpersonal interactions and relationships; major life areas, as well as various aspects of community, social, and civic life[1]. Interpersonal interactions and relationships have been defined as 'general' or 'particular'. General interpersonal interactions are defined as interacting with people in a contextually and socially appropriate manner, as well as "maintaining and managing interactions with other people"[1]. Examples include showing respect, warmth, appreciation and tolerance in

relationships; responding to criticism and social cues in relationships; forming and terminating relationships; regulating behaviors within interactions; interacting according to social rules; and maintaining social space[1].

Figure 1.1 The International Classification of Functioning, Disability and Health (ICF) Model[1]



Particular interpersonal interactions have been described as relations with strangers or formal relationships, including relating with persons in authority, with subordinates and with equals. Informal social relationships are another type of particular interpersonal interactions and are defined as relationships with others, such as casual relationships with people living in the same community or residence, or with co-workers, students, or people with similar backgrounds or professions, such as friends, neighbors, acquaintances, co-inhabitants and peers[1]. Family relationships are an additional form of particular interpersonal interactions. These include parent-child and child-parent relationships, sibling and extended family relationships. Finally, intimate relationships constitute the last type of particular interpersonal interactions and these are close or romantic relationships between individuals, such as husband and wife, lovers or sexual partners.

Another main aspect of the ICF's participation model is the major life areas domain. It includes education, work and employment, and economic life. Education is described as informal education or learning at home or in some other non-institutional setting, preschool education, school education, vocational training, and higher education[1]. Work

and employment are described as apprenticeship; acquiring, keeping and terminating a job; remunerative employment and non-remunerative employment. Finally, economic life includes basic economic transactions, such as using money to purchase food or bartering, exchanging goods or services; or saving money; complex economic transactions such as exchanging capital or property, and the creation of profit or economic value; and economic self-sufficiency, which includes personal economic resources and public economic entitlements[1].

The last component of the ICF Participation domain is Community, Social and Civic Life. This component relates to the actions and tasks required to engage in organized social life outside the family, in community, social and civic areas of life [1]. It includes community life, such as engaging in charitable organizations, service clubs or professional social organizations. Recreation and leisure is another of its elements, which includes play, sports, arts and culture, crafts, hobbies and socializing. Religion and spirituality, as organized religion and spirituality is yet another aspect of Community, Social and Civic Life. Finally, the last aspects of this major area of participation are human rights or the right to self-determination or autonomy; and the right to control one's destiny, as well as political life and citizenship[1].

Together, these domains constitute the description of the participation construct of the ICF framework, and each of these domains might constitute the source of a problem for individuals with a disability where participation might be affected or restricted.

According to the ICF framework, participation restrictions have been defined as "problems an individual may experience in involvement in life situations"[1]. Participation restriction has been proposed as a principal component of HRQOL[26-28]. However, of the three levels of disability proposed in the ICF, we know least about participation restriction[29, 30]. The measure of participation restriction is of particular interest as it is more important for clinicians and for patients to evaluate how the latter function in the context of their lives rather than the impairments or difficulties they might have with a particular task[30]. This takes on even greater importance in the case of chronic health

conditions, as patients adapt to their condition and are often able to maintain their level of participation, by adapting the contextual factors of a task (environmental or personal) to their particular disability situation. Therefore, the degree of participation restriction may not simply follow the level of impairments and activity limitations. "Even when impairments and activity limitations are irreversible, participation may still be a target for improvement"[27] (p.1890).When patients are prompted, by PROs, to judge for themselves whether they are restricted or not, individuals with the same health condition, impairment or limitations often arrive at different conclusions[30]. This is what makes the evaluation of participation restriction complex and often paradoxical, as is the case for several generic or specific HRQOL measures. The level of participation seems to be based on personal factors, personal experience, and individual expectations.

1.2.3 Measures of Participation

A few measures have been developed to assess participation and its restriction. The Impact on Participation and Autonomy (IPA) Questionnaire was developed in Dutch in 1999. The IPA Questionnaire is self-administered and addresses autonomy and participation through 31 items, divided into 5 domains: autonomy indoors and outdoors, family role, social relations, work and educational opportunities. Each item has two separate sets of responses with a range of 1-5 and 0-2, respectively, a higher score representing a greater participation restriction[31]. The IPA Questionnaire was originally tested in five patient populations: those with neuromuscular disease, stroke, spinal cord injury, rheumatoid arthritis, and fibromyalgia[32]. It has been shown to be internally reliable, with Cronbach's alphas ranging from 0.81 to 0.91[32]. Convergent validity and discriminant validity were demonstrated through correlations with the London Handicap Scale (LHS)[32]. The IPA Questionnaire is also test-retest reliable, with intraclass correlations coefficients (ICCs) ranging from 0.83 to 0.91[32], and with adequate responsiveness[33]. Since its development, the measure has been culturally adapted and validated in English in a United Kingdom population with a variety of conditions from multiple sclerosis, rheumatoid arthritis, to spinal cord injury, and persons with moderate or no disability[34]. A Swedish version of the IPA has also been developed and culturally adapted with good psychometric properties for use in people with spinal cord injury[35].

Another measure recently developed in England is the Keele Assessment of Participation (KAP). It is comprised of 11 items: mobility inside and outside the home, self-care, looking after the home and the belongings, dependents interpersonal interaction, managing money, work, education, and social activities[27]. The total score ranges from no participation restriction (0) to maximum participation restriction (11). The KAP was tested in people older than 50 years with a range of health problems, including rheumatological conditions, as well as in healthy persons. Convergent and discriminant validity of the KAP were demonstrated by matching items from the KAP to items of the IPA Questionnaire and the RNL-Index[27]. Repeatability was estimated by actual agreement and agreement beyond chance for all response options: mean observed agreement was 75.1% and chance-corrected agreement ranged from slight (weighted Kappa (Kw) =0.34) to moderate (Kw=0.64)[27]. Additionally, face validity of the four filter questions was demonstrated[27].

The "Participation Scale" is another recently developed instrument to measure the participation construct of the ICF. It was developed in six languages simultaneously (Hindi, Bengali, Telugu, Tamil, Brazilian Portuguese, and English). The scale has been validated for use with people affected by leprosy, spinal chord injuries, polio, and other disabilities[36]. It consists of 18 items with 5 response options, the total score varying from 0 to 12 (no significant restriction), 12 to 22 (mild restriction), 23 to 32 (moderate restriction), 33 to 52 (severe restriction) and 53 to 90 (extreme restriction)[36]. It is internally consistent (Cronbach's alpha= 0.92) and externally valid (Spearman's Rho=0.44) when compared to an undefined "expert score"[36]. Inter-rater and intra-rater reliability were demonstrated (r=0.80 and 0.83, respectively)[36].

Similarly, the Participation Survey/Mobility (PARTS/M) measure was developed as a self-report survey addressing participation in major life activities by people with mobility impairments and limitations, as a means to resolve the fact that there were no available measures of participation[37]. The PARTS/M was developed and evaluated with in sample of 604 people with mobility limitations and diagnoses of spinal cord injury, multiple

sclerosis, cerebral palsy, stroke, or post-poliomyelitis. The PARTS/M contains 20 items assessing "major life activities" placed into 6 domains of the ICF: self-care; mobility; domestic life; interpersonal interactions and relationships; major life areas; and community, social, and civic life. Four- or 5-option Likert type response sets are used for most items, except the health-related limitation question, which has a dichotomous response option. The PARTS/M has been shown to be internally consistent, Cronbach's alpha ranging from 0.85 (civic life domain) to 0.91 (self-care domain)[37]. Additionally, it has been shown to have convergent validity, as it demonstrated a canonical correlation of 0.71 when compared to the RNL-Index. The length and the time required to complete the survey (up to 90 minutes) were recognised as being prohibitive to use in some settings by the developers[37].

Yet another measure recently developed to assess the participation construct of the revised ICF Model is the Participation Measure for Post-Acute Care (PM-PAC)[38]. The PM-PAC measures participation outcomes after rehabilitation for community-based individuals. It contains 51 items that assess participation in nine domains: mobility; role functioning; community, social, and civic life; domestic life/self-care; economic life; interpersonal relationships; communication; work; and education[38]. The current version consists of 18 main questions, many of which are further subdivided from two to 14 questions. Responses options are in a three- to six-option Likert format. Although still under refinement, the measure was analysed psychometrically using item response models, exploratory factor analysis, and by evaluating test-retest reliability and known-groups validity. Data came from 395 subjects recruited from outpatient rehabilitation units or home-care settings who had neurological, musculoskeletal or medically complex conditions (debility due to illness, cardiopulmonary conditions, surgery). Intraclass correlation coefficients ranged from 0.61 to 0.86 for test-retest reliability[38]. Diagnostic groups, ranked on the Rankin score (a measure of functional disability), were able to distinguish all scales excepted domestic life[38].

Other measures, based on the earlier ICF version from the WHO, have focused on integration, a construct similar to participation. For example, the Community Integration

Measure (CIM) is a 10-item, client-centered measure used with a broad range of patients with disabilities. The CIM has a single summary score (between 10 and 50), that is the sum of the 10 items; each with five response options[39]. It has been shown to be valid and reliable only in moderate to severe brain injured patients. Similarly, the Community Integration Questionnaire (CIQ) is a 15-item measure designed for people with acquired brain injuries. It is divided into 3 subscales (home and family, social and vocational independence). It is a widely used, valid, and reliable measure in such patients[40].

In summary, although a few measures have been developed to assess the Participation construct of the ICF model, none of them have been used universally for this purpose.

CHAPTER 2:

2.1 The Reintegration to Normal Living (RNL)-Index

2.1.1 Development of the Measure

Although impairment and activity limitations have been measured extensively in both English and other languages, the focus on measuring participation is recent[27, 29]. In 1987, the Reintegration to Normal Living (RNL)-Index was published[2]. At the time of development, "reintegration" was conceptualized as the reorganization of physical, psychological and social characteristics so that a person could resume a well adjusted life following an incapacitating illness or injury. It was designed as a PRO. The measure was originally developed in Canadian English and Canadian French using information from persons with cancer or cardiac conditions who successfully returned to normal living patterns, a few subjects with other conditions, professionals who routinely worked with such people, and lay individuals. It was subsequently tested by the developers for its psychometric properties in subjects with cancer and heart disease[2, 4]. Since that time the RNL-Index has been revalidated and/or used in diverse groups of subjects including those with stroke, traumatic brain injury spinal cord injury, those with amputations and the elderly[41]. With the recent publication of the revised ICF, the RNL-Index has gained popularity as the construct "Reintegration" has been equated with that of "Participation". Since then, the developers have received a number of requests to use the English and French versions of the measure, and a few more to translate it into other languages.

The RNL-Index measures a person's return to his or her usual physical and social activities following illness or trauma. It includes 11 items, covering the following domains: indoor, community, and distant mobility; self-care; daily activity (work and school); recreation and social activities; general coping skills; family roles; and interpersonal interactions and relationships. There are two subscales: "Daily Activity" (mobility, participation in work, family roles, social and recreational activities) and "Perception of Self" (comfort with relationships and coping skills). Each domain is accompanied by a visual analogue scale or a three- or four-point Likert categorical scale

with response descriptors. The items add up to a total score, dependent on the response set used [2].

2.1.2 Psychometric Properties

In the original psychometric testing of the RNL-Index, the sample (two primary care hospitals and one rehabilitation hospital: conditions of the subjects not defined) scored 83.9±14.7 on a score of 100 and 90.2±22.9 in a sample of "subjects of newly diagnosed, hospitalized, treated, and discharged subjects who had cancer (63%) or a myocardial infarction (37%)". Internal consistency as estimated by Cronbach's alpha was above 0.90 for each subscale, and item to total correlations ranged from 0.39 to 0.76 [4]. Content validity was established by the method of development (literature reviews, experience of investigators, application of open- and close-ended questionnaires to subjects, health care professionals, and significant others). As no gold-standard was available to assess reintegration to normal living, convergent validity was examined. Moderately strong, in the expected direction, and statistically significant correlations were obtained between the RNL-Index score and the Quality of Life (QL) Index Score (r=0.68), as well as with a subjective evaluation of psychological well being (positive well-being (r=0.32), negative well-being (r=-0.41), general well-being (r=0.41)). Additionally, responsiveness was analysed in persons with myocardial infarction and cancer newly diagnosed, hospitalized, treated and discharged from hospital, that were followed up at three months. The RNL-Index demonstrated a significant change in subjects' total index scores and in the "Daily Activity" subscale, but not in the "Perceptions of Self" scores [4].

A number of studies have further assessed the psychometric properties of the RNL-Index. Test-retest reliability was established as strong (r = 0.83) for community-dwelling elderly [42], but poor (r=0.12) for adults with traumatic brain injury[6]. Known-groups validity was demonstrated in subjects at 3 months and 1 year post-stroke grouped by levels of impairment, by the presence or absence of depression, by levels of physical disability, and by cognitive disability[7].

Other studies have used the RNL-Index as an outcome in persons with a variety of conditions: malignant tumours, degenerative heart disease, arthritis, fractures, amputations[43], musculoskeletal disorders[44], stroke[45-47], spinal cord injury[48], brain injury[49-51], ruptured and unruptured cerebral aneurysms[41], victims of motor vehicle accidents[52], and community-dwelling elderly[41, 42].

The strong psychometric properties suggest that the RNL-Index is a valid, reliable and responsive measure reflecting participation in life activities. Moreover, its widespread use across persons with diverse diseases and conditions demonstrates its important role in the clinical and research setting. Hence, the need for a cultural adaptation of the RNL-Index is manifest.

2.2 RATIONALE

There are a few measures developed in English that tap the ICF's construct "Participation". A review of the literature, although not systematic, was done to assess the existence of a Participation measure in Spanish. The review, along with input from the leader of the cultural adaptation of PRO measures in Spain, confirmed that to date, there is no measure that has been developed or translated into Spanish to assess this construct. Such a measure for use in the Spanish population would be essential to understand how a condition or its treatment affects a person's ability to resume usual living patterns.

Only recently have measures been developed to specifically measure the ICF's "Participation" construct, but their use in a clinical or research setting has yet to be fully recognized. The RNL-Index, with close to 20 years of use, is the measure of reference to evaluate participation. It is recognized as the criterion measure against which the more recently developed measures are tested. Its widespread use across persons with diverse diseases and conditions has demonstrated its important role in the clinical and research setting and supports the necessity of culturally adapting the RNL-Index as a measure of participation.

The translation and cultural adaptation of the widely used measure, the Reintegration to Normal Living Index, will allow Spanish health care practitioners and researchers to use a standardized instrument to evaluate reintegration and participation in life events following a disabling health condition. It will also facilitate a comparison of the Spanish population to other populations in terms of its reaction to disability and the ability to assume usual living patterns. Additionally, it will allow Spanish researchers to compare their findings to other published studies. Finally, it will join a growing collection of measures that have been translated into Spanish, the third most widely spoken language in the world (by over 300 million people[53]. These measures include the Short-Form-36 Health Survey (SF-36) [54, 55], the Lehman's Quality of Life Interview[56], the Nottingham Health Profile[57] and the EuroQol 5D[58].

CHAPTER 3:
RESEARCH QUESTION, OBJECTIVES AND HYPOTHESES

3.1 Research Question

In light of the rationale, we ask: In a Hispanic population living in Spain with total knee replacement or stroke, to what extent does the Spanish translation of the RNL-Index yield estimates of validity and reliability that are considered to be sufficient and comparable to those of the original RNL-Index?

3.2 Global Objective

The objective of this study is to translate, culturally adapt, and retest the psychometric properties of the RNL-Index to obtain a Spanish version of the measure that is conceptually equivalent to the original and acceptable for use with persons in Spain.

3.3 Specific Objectives and Hypotheses

The specific objectives are:

- To estimate the internal consistency of the Spanish RNL-Index (SRNL-Index) and its subscales.
- To estimate the test-retest reliability of the SRNL-Index.
- To review the content of the SRNL-Index in terms of cultural appropriateness.
- To determine the degrees of association between the SRNL-Index subscales "Daily Activity" and "Perception of Self" subscales and the Short Form 36-item Health Survey (SF-36) "Mental Component Summary (MCS) Score" and the "Physical Component Summary (PCS) Score" *(Discriminant construct validity).*
- To determine if the SRNL-Index score is correlated with the Six Minute Walk Test (6MWT) *(Convergent construct validity).*
- To determine if SRNL-Index scores can differentiate subjects with stroke, grouped by level of functional independence as measured by the Bathel Index (BI), and all subjects by walking speed *(Known-groups/discriminative construct validity).*

We hypothesize that:

- The total SRNL-Index score and its subscale scores will demonstrate acceptable levels of internal reliability (Cronbach's alpha≥0.70) *(Internal consistency).*
- The SRNL-Index will have a test-retest reliability acceptable at the group level (r≥0.70) *(Test-retest reliability).*
- The SRNL-Index "Daily Activity" subscale score will correlate more strongly with the PCS Score of the SF-36 than with the MCS Score *(Discriminant construct validity).*
- The SRNL-Index "Perception of Self" subscale score will correlate more strongly with the MCS Score of the SF-36 than with the PCS Score *(Discriminant construct validity).*
- The total score of the SRNL-Index will correlate moderately and positively with the 6MWT (r=0.40 to 0.59) *(Convergent construct validity).*
- Scores of the SRNL-Index will discriminate between persons with stroke, grouped by independence according to the Barthel Index (BI), and all subjects by gait speed *(Known-groups/discriminative construct validity).*

CHAPTER 4: MANUSCRIPT

4.1 Rationale

In the first chapters, we highlighted the importance of translating and culturally adapting health measures, the recently developed ICF Model, with particular focus on the Participation construct and the measures developed to assess it.

There has been no measure developed or translated into Spanish to assess Participation. Such a measure for use in the Spanish population is essential to allow Spanish health care practitioners and researchers to use a standardized instrument to evaluate reintegration and participation in life events following a disabling health condition

The Reintegration to Normal Living (RNL)-Index, with close to 20 years of use across persons with diverse diseases and conditions has demonstrated its important role in the clinical and research setting. This supports the necessity to culturally adapt the RNL-Index as a measure of participation.

The purpose of this paper is to summarize the process and the results obtained during the translation, cultural adaptation, and assessment of the psychometric properties of the RNL-Index into Spanish.

4.2 TRANSLATION, CULTURAL ADAPTATION AND REVALIDATION OF THE REINTEGRATION TO NORMAL LIVING (RNL)-INDEX FOR USE IN SPAIN

Ana María Rodríguez, B.Sc. [1]
Sharon Wood-Dauphinee, Ph.D. [1]
Nancy E. Mayo, Ph.D.[1]
Montserrat Ferrer, M.D. [2]
Jordi Alonso, M.D. [2]

[1] School of Physical and Occupational Therapy, Faculty of Medicine. McGill University, 3654 Prom Sir William Osler, Montreal, Quebec, H3G 1Y5, Canada.

[2] Health Services Research Unit, Institut Municipal d'Investigació Mèdica (IMIM), Barcelona Biomedical Research Park, C/ Doctor Aiguader, 88, 08003 Barcelona, Spain.

Manuscript prepared for submission to the journal entitled "International Journal of Rehabilitation"

Running title: Translation and cultural adaptation of the RNL-Index into Spanish

Communication addressed to: Sharon Wood-Dauphinee
School of Physical and Occupational Therapy
Faculty of Medicine, McGill University
3654 Prom Sir William Osler
Montreal, Quebec, H3G 1Y5
CANADA
Telephone: (514) 398-5326
Fax: (514) 398-6360
e-mail: sharon.wood.dauphinee@mcgill.ca

ABSTRACT

Background. Participation has been proposed as a principal component of health-related quality of life and is an important construct to measure, but no measure of participation exists in Spanish. A widely used, valid, reliable, and responsive measure of participation is the Reintegration to Normal Living (RNL)-Index. The general objective of the present study was to translate, culturally adapt, and assess the psychometric properties the RNL-Index in Spanish. *Methods.* The translation and the cultural adaptation of the RNL-Index consisted of a five-step process, leading to the Spanish version of the RNL-Index (SRNL-Index). Psychometric evaluation employed a cross-sectional study design, with a longitudinal design used for test-retest evaluation. Thirty-two subjects with stroke and 36 with total knee arthroplasty completed the SRNL-Index twice, as well as the Six Minute Walk Test (6MWT), the Short-Form 36-item Health Survey (SF-36), and for stroke subjects, the Barthel Index (BI). Descriptive statistics, one-way analysis of variance and post-hoc t tests were calculated, as well as Cronbach's alpha, Pearson's, and Intraclass correlation coefficients. *Results.* The SRNL-Index was found to be internally consistent (global Cronbach's alpha: 0.94 for stroke and 0.88 for TKA samples). Test-retest reliability was quite poor within our samples, being moderately low for the TKA sample (ICC=0.22) and moderate for the stroke sample (ICC=0.55). Discriminant validity was demonstrated by the correlations between the SRNL-Index's Perception of Self and Daily Activity subscales with the Mental and Physical Component Summary scores of the SF-36. The SRNL-Index has convergent validity as shown by the high correlations between of the SRNL-Index and the 6MWT (r=0.70 for stroke and 0.61 for TKA samples). Known-groups validity was seen in people with stroke grouped by Barthel Index scores, except for the most independent group, and in both samples when grouped according to gait speed. *Conclusions.* The SRNL-Index demonstrated acceptable validity and internal reliability for subjects with stroke and total knee arthroplasty. However, further studies are needed to reassess validity, external reliability, and responsiveness in other populations and other Spanish-speaking countries.

BACKGROUND

The "International Classification of Functioning, Disability and Health" (ICF) is a model that classifies health and health-related function in domains that constantly influence and modify each other[9]. These domains are the Body functions and structures, Activities of the person and Participation in life situations, and are influenced by Environmental and Personal Factors[1]. In this model, the type of opportunities and resources necessary for well-being vary between persons dependent on their life-circumstances and personal characteristics[10].

The ICF's definition of participation is "involvement in life situations"[1]. Participation includes interpersonal interactions and relationships; major life areas (education, employment, economic life), and community, social, and civic life[10]. Participation has been described as a principal component of health-related quality of life[26-28], but little is known about participation[29, 30]. Measures of participation are of particular interest, as clinicians and patients need to evaluate function in the context of patients' lives not just the difficulties they might have with a particular task[30]. The level of participation is based on personal factors, personal experience, and personal expectations. It is not simply the result of the level of impairments and/or activity limitations[30].

A few measures have been developed, mostly in English, to assess participation[27, 36-38]. As participation is a relatively new concept, the measures that assess it have yet to be recognized in a clinical or research setting. None are universally used to measure Participation.

To date, there is no measure that has been developed or translated into Spanish to assess this construct. Such a measure for use in the Spanish population would be essential to understand how a condition or its treatment affects a person's ability to resume usual living patterns.

The Reintegration to Normal Living (RNL)-Index, with close to 20 years of use, is the measure of reference to evaluate participation[2, 4, 27, 37]. Its widespread use across persons with diverse diseases and conditions [2, 4, 41, 41-52], as well as its established validity,

reliability, and responsiveness [2, 4, 6, 7, 42], has demonstrated its important role in the clinical and research setting and supports the necessity of culturally adapting the RNL-Index as a measure of participation.

OBJECTIVES

The general objectives of the present study were to translate, culturally adapt, and assess the psychometric properties the RNL-Index in Spanish, in order to obtain a measure of participation that is acceptable for use with people in Spain.

METHODS

The study has two distinct parts: 1) the translation and cultural adaptation of the RNL-Index into Spanish, 2) the psychometric evaluation of the Spanish version of the RNL-Index.

Translation Process

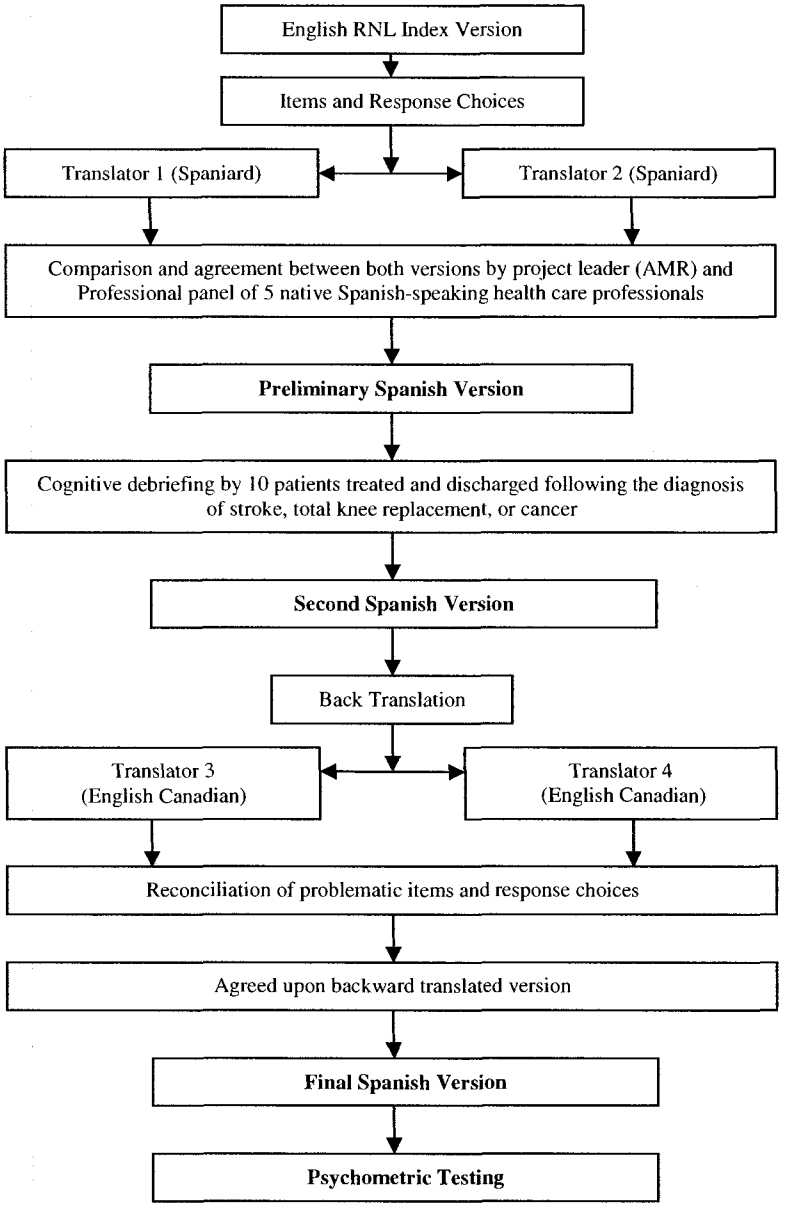
The methods that were used for the translation and the cultural adaptation of the RNL-Index for use in Spain loosely followed the process set out in the IQOLA Project protocol[59]. The translation procedure algorithm is presented in Figure 1 and is described below.

Forward translation

Two individuals translated individually the items, response choices and instructions of the English version of the RNL-Index. The translators were selected on the basis of professional qualifications and/or experience in translation, knowledge of the English language and for having Spanish (from Spain) as a mother tongue. The latter criterion was applied because the choice of specific words and phrases needed to reflect the Spanish language as spoken in Spain. When completed, the researcher (AMR) and 5 native Spanish-speaking health care professionals (2 physicians, 2 physical therapists, 1 investigator in public health) compared the two translations and came to a consensus that

the items reflected Spanish language, life and thinking. This constituted the Preliminary Spanish Version of the RNL-Index.

Figure 1. Translation Procedure Algorithm



Cognitive debriefing

A group of 10 subjects who had been treated and discharged after a stroke, a total knee replacement or cancer were 'cognitively debriefed' by the researcher (AMR) concerning the content, the wording and the meaning of the items and the instructions in the Preliminary Spanish Version of the RNL-Index. Subjects were recruited by physician referral from the Hospital de l'Esperança. They were males or females, native Spaniards of different ages and levels of education. Each person individually completed an item of the Index and was asked: 1) Did you have problems understanding the item? 2) What does this item mean to you? 3) Please explain the item using your own words 4) Do you have any suggestions to improve the item? 5) Are the response options clear and appropriate for the item? The process was repeated for each item and resulted in the Second Spanish Version of the RNL-Index.

Backtranslation

The reconciled Spanish version was then given to two translators, selected on the basis of professional qualifications, knowledge of Spanish, having Canadian English as a mother tongue, and who were unaware of the RNL-Index. The backtranslators independently translated this Second Spanish version of the RNL-Index into Canadian English. The backward translations were then compared to the original RNL-Index by the researcher (AMR) and the original developer of the RNL-Index (SWD). The final Spanish version of the RNL-Index was entitled "Índice de reintegración a la vida normal" (Índice RVN). To facilitate reading, throughout this text, it is referred to as the Spanish RNL-Index (SRNL-Index).

Settings

The forward translations of the RNL-Index into Spanish were performed in Montreal, Canada. The consolidation of the two forward translations into a first Spanish version of the RNL-Index took place at the Institut Municipal d'Investigació Mèdica (IMIM), in Barcelona, Spain. The IMIM is a public centre focused on scientific research into Biomedicine and Health Science fields. The IMIM is organized into multidisciplinary

research units and groups and includes research personnel from the IMIM itself and clinical researchers from the Institut Municipal d'Assistència Sanitària (IMAS) centers[60].

The cognitive debriefing was performed at the Hospital de l'Esperança, in Barcelona, Spain, an IMAS center. This 900-bed hospital specializes in surgery, particularly for osteo-articular, chronic and degenerative conditions. Its day hospital with 28,000 patient-visits provides general as well as specialized post-surgical rehabilitation[61].

The back-translations into English and consolidation into a final Spanish version were done via electronic collaboration between centers in Montreal and in Barcelona.

Psychometric Testing of the SRNL-Index

Study design

A cross-sectional study design was used to meet the study objectives, with the exception of the test-retest portion, which had a longitudinal design.

Participant Selection

Subjects who had been diagnosed, treated and discharged as inpatients from the Hospital de l'Esperança for either a total knee replacement or a stroke, but who continued to attend the hospital as outpatients at the rehabilitation department, were asked to participate. Even though the RNL-Index is valid, reliable and responsive for a variety of disabling conditions, the choice of selecting people with stroke was because known-groups validity had been studied with this population[7]. The choice of selecting participants with total knee replacement was based on having these readily available, and because it had previously been used to assess the psychometric properties of the original RNL-Index. The subjects were selected according to availability, and aimed to represent a range of severity, ages and gender, and to reflect typical subjects with these diagnoses.

For inclusion in this study, participants had to have Spanish as a mother tongue, to live in Spain, had to have had a total knee replacement or a stroke; and had to have been discharged as an inpatient within ten days to three months from the Hospital de

l'Esperança. Persons with serious co-morbid life-threatening conditions that could affect return to normal living patterns, a poor cognitive status or inability to read limiting their completion of the measures, or who were less than 18 years of age were excluded. Given the average reading ability of the study population, we agreed that we would read the questionnaires to the subjects if it wasn't understood. Lack of comprehension was especially a concern for the elderly subjects with stroke.

Sample Size

Sample size was determined as recommended by Hobart[86]. This author undertook an empirical examination of the impact of sample size on reliability and validity estimates. Psychometric analyses in small samples (reliability: n>20, validity: n>40) was determined to generate valuable information. However, in samples of n>35, results reflected those generated in the original sample[86]. We therefore aimed at recruiting 35 subjects with each condition.

Procedures

An eligible list of subjects was given to the researcher (AMR) by the rehabilitation unit director at the Hospital de l'Esperança. The researcher then directly approached the eligible subjects at the rehabilitation unit of the hospital after being introduced by the treating physiotherapist or occupational therapist. The subjects were informed individually about the study and written information was obtained. If the subjects verbally agreed to participate in the study, an appointment was given for the assessment to take place. On the day of the assessment, a consent form was given for all participants to sign, and any questions that they had about the study were answered. They were then evaluated by the researcher and main evaluator (AMR), a physiotherapist familiar with the questionnaires. Subsequent to consenting, subjects were administered the SRNL-Index (score 1), the Short Form 36-item Health Survey (SF-36), and the Six Minute Walk Test (6MWT) and the evaluator recorded sociodemographic information such as gender, age, marital status, living arrangements, education, and current employment status of the subjects. The Barthel Index score was obtained from the stroke subjects' medical files. Lastly, subjects were given an additional copy of the SRNL-Index, with the instruction to

complete it and return it by mail five days later, along the answer to a written question on whether their health status had changed or not since the last assessment and the direction of the change, if any. If subjects did not return it within ten days, they were contacted by telephone and reminded to do so. If they failed to return the completed form after the reminder, their data were not included in the test-retest reliability analysis. Even though a different mode of administration between the first and the second administration of the SRNL-Index was less than ideal, many subjects had almost completed their rehabilitation session as an outpatient, and it would have been impossible to asses them in the hospital premises five days later.

Settings

The psychometric testing of the final Spanish version of the RNL-Index took place at the Hospital de l'Esperança. Subsequent analyses of the data were done in Montreal, Canada.

Ethics

Ethical approval was obtained from the Institut Municipal d'Assistència Sanitària in Barcelona. Each participant provided informed written consent.

Instrumentation

The Reintegration to Normal Living (RNL)-Index

The Reintegration to Normal Living (RNL)-Index is a standardised self-administered measure to assess participation. The measure was originally developed in Canadian English and Canadian French with persons with cancer or heart disease[2, 4]. The RNL-Index has been shown to be valid[2, 7], reliable[42] and responsive[2] and has been revalidated or used with diverse conditions including malignant tumours, degenerative heart disease, arthritis, fractures, amputations, musculoskeletal disorders, stroke, traumatic brain injury, cerebral aneurysms, spinal cord injury, general physical disability, motor vehicle accidents, and with community-dwelling elderly[41-52].

The RNL-Index has 11 items that assess mobility, self-care, daily activity, recreation and social activities, family roles and relationships, as well as general coping skills and

interpersonal relationships[2]. Each item is scored with a 100 point visual analogue scale or a three- or four-point Likert categorical scale. The four-point categorical scale was used in this study, scored 1-4 for a total score range of 11 to 44. Scores can also be obtained for two subscales, "Daily Activity" and the "Perception of Self".

Short Form 36-item Health Survey (SF-36)

The SF-36 is a generic Health-Related Quality of Life measure that assesses 8 health concepts: limitations in physical functioning or social activities due to health problems, limitations in usual roles activities due to physical health problems or emotional problems, pain, vitality, mental health, and perception of health in general. Each item is scored on a dichotomous, three or five-point categorical scale; subscale scores range from 0 to 100. Physical and mental summary scores can also be constructed[62]. It was culturally adapted into Spanish by Alonso and colleagues as part of the International Quality of Life Assessment (IQOLA) project (1991-1995)[54, 55].

Six Minute Walk Test (6MWT)

The Six Minute Walk Test (6MWT) is a submaximal functional test of walking endurance[63]. The 6MWT has been studied in several different populations and is a valid and reliable measure[64], that is commonly used in Spain. In this study, the 6MWT was administered with a marked 15-meter path in a hospital corridor. The instructions to the subjects were: "Walk as quickly as you can for 6 minutes to cover as much ground as possible in this 15-meter path. You may rest if you need to, but continue walking as soon as possible". The distance walked was recorded. All subjects were ambulators, with or without walking aids, and were hence able to complete the 6MWT.

Walking Speed

Gait speed is a physical characteristic derived from directly measuring the parameters of distance and time. Gait speed is considered a valid measure of walking ability as it correlates with functional mobility, degree of independence in walking, and many different gait parameters[65-67]. It has been associated with strength of the affected lower extremity, cadence and stride length, balance, degree of lower extremity motor recovery,

and functional mobility[65, 68-70]. Standardized instructions are to walk at a "comfortable" or "maximum" speed along a walkway typically ranging from 2 to 20 meters[71]. In this study, the distance walked during the 6MWT was divided by the time to arrive at a walking speed. Subjects were instructed to walk at a comfortable pace and walked over a distance of 15 meters.

Barthel Index (BI)

The Barthel Index (BI), a commonly used measure of activities of daily living in people with stroke, has well established reliability and validity[72]. There are 10 items: feeding, grooming, bathing, dressing, controlling bowel and bladder, personal toilet, ambulation, transfers and stair climbing. It has a maximal score of 100[72] and has been fully implemented clinically and in research in Spain[73, 74].

ANALYSIS

Descriptive data

To describe the sociodemographic characteristics of the participants, means, standard deviations, counts and percentages were calculated. Floor and ceiling effects of the samples were obtained.

Reliability

Two types of reliability, internal consistency and test-retest, were assessed. Internal consistency was estimated using Cronbach's alpha coefficients. Test-retest reliability was calculated by comparing the consistency of scoring of the SRNL-Index at the two times via a one-way analysis of variance design of the Intraclass Correlation Coefficient (ICC 1,2[75]). An ICC of 0.70 is acceptable to make decisions at a group level[76], and is, therefore, the lowest ICC for the SRNL-Index to be considered reliable. An ICC of 0.90 is considered appropriate at the individual level[76], and would be even more preferable.

Validity

In an effort to confirm construct validity, discriminant, convergent, and known-groups validity were assessed. Discriminant construct validity was estimated by Pearson correlation coefficients between the SRNL-Index subscales and the SF-36 component summary scores. It was hypothesized that: 1) The SRNL-Index "Daily Activity" subscale would correlate more strongly with the PCS of the SF-36 than with the MCS; 2) The SRNL-Index "Perception of Self" subscale score would correlate more strongly with the MCS of the SF-36 than with the PCS.

Convergent construct validity was estimated by Pearson correlation coefficients between the SRNL-Index and the 6MWT. It was hypothesized that the correlation between the SRNL-Index and the 6MWT would be positive and moderate (between 0.40 and 0.59)[76].

Participants with stroke were grouped in three categories based on their Barthel Index (BI) scores to estimate known-groups validity. The cut-off scores defining each level of independence were determined using the BI classification of functional independence for stroke; that is scores ≥ 91 were classified as independent, scores of 61 to 90 were defined as slightly to moderately dependent in self-care and mobility, and score of ≤ 60 were defined as moderately to markedly or totally dependent in self-care and mobility [72]. A general linear model with post-hoc t tests estimated whether the SRNL-Index could discriminate between subjects across three levels of independence measured by the BI with an alpha level set at 0.05. It was hypothesized that the mean scores of subjects in the 'independent' category would demonstrated the highest scores in the SRNL-Index, those in the 'moderately to totally dependent' category would score the lowest, and those in the 'slightly to moderately dependent' category would score in the middle.

Additional analyses grouped subjects with stroke and with total knee arthroplasty, who were all ambulators, according to their mean walking speed during the 6MWT. Groups were determined by walking speed categories established by Perry and colleagues in 1995[77]: walking speeds < 24 m/min were defined as 'household ambulators', those between 24 and 35 m/min were defined as 'limited community ambulators', and those $\geq$

35 m/min were classified as 'community ambulators'. A general linear model with post-hoc t tests estimated whether the SRNL-Index could discriminate between subjects across three walking speed groups, with an alpha level set at 0.05. It was hypothesized that the scores of subjects in the 'household ambulator' category would show the lowest scores in the SRNL-Index, that those in the 'community ambulator' category would be the highest, and that those in the 'limited community ambulator' category would score in the middle.

Data analysis

All data were analyzed using SAS, version 9.1 (SAS Institute Inc., Cary, NC, USA).

RESULTS

Translation process

After much discussion, consensus was obtained for the translations of all items and response options. The initial translated version of the SRNL-Index included three items as worded by translator 1, five items as translator 2, and a combination of the provided translations or a completely new wording was chosen for the remaining three items.

Cognitive debriefing

The cognitive debriefing lead to the following results: the 10 subjects debriefed felt that:

- The questionnaire was comprehensible and the questions made sense.
- There were no difficulties in completing the SRNL-Index.
- The response options were clear and appropriate for each item.
- Following the subjects' comments, a minor alteration was made to the preliminary Spanish Version of the RNL-Index for items 9 and 10. The wording "me siento a gusto con..." was suggested by eight of the subjects to be changed to "me siento bien con...". The change was adapted and constituted the Second Spanish Version of the RNL-Index.

Backtranslation and reconciliation of the Final Spanish Version of the RNL-Index

The Second Spanish Version of the RNL-Index, translated into English by two independent backtranslators, was essentially the same as the original RNL-Index. No

changes were made to the Second Spanish Version of the RNL-Index, as the items and response options appeared to be conceptually equivalent to the original version.

Psychometric Testing

Participants

A total of 68 subjects agreed to participate. Of those, 32 had had a stroke and 36 had undergone a total knee arthroplasty (TKA). The sociodemographic information is shown in table 1, at the end of this manuscript. Sixty-three percent of those with stroke and 39% of those TKA were males. Twenty eight percent of the stroke and 44% of the TKA samples had not completed primary education (8^{th} grade). Most of the participants had Spanish as a first language (72% of the stroke and 83% of the TKA samples). However in all instances where Spanish was not the first language, the subject was as fluent (or more) in Spanish as in the first language (Catalan). A large proportion of the participants were from Catalonia (53% of the stroke and 39% of the TKA samples), but as the Barcelona region has always been a region of high immigration from Spain and elsewhere, participants came from a variety of other Spanish provinces. The self-reported level of activity is included as part of the standard sociodemographic questionnaire used by the research facility, however, data for this question were missing for six of the stroke patients.

Scores on the SRNL-Index

The distributions of the SRNL-Index scores are presented in table 2. The total scores for both samples were fairly similar, 32.5 and 34.5, as were the 'Daily Activity' subscale scores. The 'Perception of Self' subscale scores varied with the subjects with a TKA, scoring higher on average than those with stroke. The distribution of scores from subjects with stroke varied across the theoretical range for both the total score and the subscales. The TKA sample scored in the higher portion of the range for both the total and the subscales score. Floor effects ranged from 0 to 3.1%. Ceiling effects ranged from 3.1 to 63.9% and were highest for the 'Perception of Self' subscales, 34.4% for stroke and 63.9% for TKA samples.

Reliability

Internal Consistency

Inter-item correlations for subjects with stroke ranged from 0.13 to 0.73 and for subjects with TKA, from -0.17 to 0.73. The alpha correlations between the items and total scores varied from 0.43 to 0.78 for stroke and only two items correlated below 0.40 for TKA participants. Standardized Cronbach's alphas for each item and for the subscales were high in both samples, ranging from 0.86 to 0.94 (table 3). Global Cronbach's alphas were 0.94 for the stroke and 0.88 for the TKA samples (table 3).

Test-Retest Reliability

Test-retest reliability results are shown in table 4. The response rate for the second SRNL-Index was 87.5% for the participants with stroke and 77.8% for those with TKA. ICCs were moderately low for the TKA sample and moderate for the stroke sample.

Validity

Discriminant validity

Results for discriminant validity are presented in table 5. Pearson's correlation coefficients between the SF-36's Physical Component Summary Score (PCS) and the Daily activity subscale were 0.55 for the stroke and 0.64 for the TKA sample, and were 0.22 for stroke and 0.35 for TKA participants for the Perception of Self subscale. The correlation for the SF-36's Mental Component Summary Score (MCS) with the Daily activity subscale was 0.57 in the stroke sample, and 0.43 in the TKA sample, and was 0.57 for the stroke and 0.64 for the TKA participants for the Perception of Self subscale.

Convergent validity

The convergent validity results can be seen in table 6. The correlations between the SRNL-Index and the 6MWT were moderately high in both samples, 0.70 for stroke and of 0.61 for TKA participants. The 6MWT correlated well in both samples with the Daily Activity subscale, but only moderately with the Perception of Self subscale.

Known-groups validity

In participants with stroke, the distribution of SRNL-Index scores across groupings by levels of independence, as measured by the Barthel Index, indicated the discriminative ability of the SRNL-Index (table 7). Significant differences were seen between all groups except between the 'slightly to moderately dependent' and the 'independent' group.

TKA and stroke samples, when grouped into 'household ambulator', 'limited community ambulator', and 'community ambulator' walking speed categories, showed significant differences in SRNL-Index scores. The differences were significant across all walking speed categories of the stroke sample, and across the TKA sample except for the difference between the two fastest groups.

DISCUSSION

The RNL-Index has been validated or used with diverse groups of subjects. Our samples consisted of people with stroke and total knee arthroplasty from the Hospital de l'Esperança. The main goal of this study was to create a standardized, valid, and reliable Spanish version of the RNL-Index that could be used with these populations in Spain. In this section, the translation process and psychometric properties of the SRNL-Index are discussed.

Translation process

Although some difficulties were initially encountered in finding translators and back-translators, the procedures followed our translation logarithm. The results of the cognitive debriefing session show that the SRNL-Index was understood. Although a suggestion was made by the debriefed subjects to improve one word, they had little problem in understanding or rephrasing an item into their own words. A comparison of the backtranslated SRNL-Index into English confirmed that the items and response options were conceptually equivalent to the original version.

The samples

Both stroke and TKA samples were poorly educated (8 or less years of education). This is fairly representative of the Spanish population of this age group. Because of the Spanish civil war (1936-1939) and the difficult post-war period characterized by dictator Franco's regime (that meant famine and extreme poverty for a large percentage of Spaniards), many children stopped their schooling. Some participants (four men) mentioned going to a 'clandestine' school (night school, usually at a teacher's home) for a couple of months to a year, but this could not be maintained permanently. For girls, this usually meant the end of their studies and the beginning of domestic tasks at home. Low educational levels, seen more predominantly in the TKA sample than in the stroke sample, could perhaps be explained by the fact that this group consists of a higher percentage of women. While 46% of subjects reported having Catalan as a first language, this should not be seen as a threat to the generalization of the findings to the Spanish population, as all Catalonian citizens speak Spanish. This is especially true for this generation of Catalonians, as Catalonia was repressed during Franco's regime: it was forbidden to speak Catalan in public. It is not uncommon to see Catalonians who write and speak perfect Spanish, but who cannot do the same in their mother tongue, Catalan. Since Barcelona has always been a region of high immigration (from other regions of Spain and worldwide), participants came from a variety of other Spanish provinces. It, therefore, seems appropriate that results could be generalized to most Spanish provinces.

Ceiling and Floor Effects

As presented in the results, the total knee arthroplasty (TKA) sample demonstrated a moderate ceiling effect in both the subscales and the total score of the SRNL-Index. This was especially true for the Perception of Self subscale, where a strong ceiling effect (64%) was observed. To a lesser extent, this ceiling effect was also seen in this subscale with the stroke sample (34%). It is suspected that this was caused by a selection bias, as the ceiling effect was mostly seen in one of the study samples. To be recruited in the study, subjects were approached in person. Although not documented during the recruiting, it seemed that subjects who refused to participate were not in an as good physical condition as those who accepted. It should be noted that because of the

difficulties in recruiting patients, it was not possible to record the participation rate, refusals, and reason for refusals. Studies with bigger samples and documentation of follow-up of refusals should be conducted to determine if the ceiling effect seen in the Spanish version of the Perception of Self subscale is persistent or could be explained by a selection bias. It should be noted, that the ceiling effect and lack of variability seen in the TKA sample affects some of the psychometric results.

Reliability

Internal Consistency

The SRNL-Index demonstrated high levels of internal consistency for both the total index and its subscales within our samples. These values are similar to those reported by the developer, with internal consistency, estimated by Cronbach's alpha, above 0.9 for each scale[4]. Nonetheless, the scales and subscales with alphas higher than 0.90 could suggest redundancy of items[76]. Item-to-total correlations were generally moderate to high, which resembles the findings of the original study where item-to-total correlations ranged from 0.39 to 0.76[4].

The inter-item correlations of the stroke sample ranged from moderately low to moderately high, but remained below 0.90, suggesting that redundancy within the items of the SRNL-Index in this population is not a major issue. The inter-item correlations for the TKA sample are somewhat problematic, as they often approached 0 and were unexpectedly negative in some cases (although always close to 0). This can be explained by the lack of variability and the ceiling effect in the TKA sample's scores. This low variability nullifies the results for the TKA samples, as correlations, and all parametric tests for that matter, loose their relevance. Additional studies are required in a larger sample of TKA participants with greater variability in function. The SRNL-Index should be considered as internally reliable in the stroke sample.

Test-Retest Reliability

The ICCs for test-retest reliability for the TKA and the stroke sample were below the acceptable level of reliability for decision-making at a group level (at least 0.70)[76]. The

modes of administration and the settings between the first and second completions of the SRNL-Index were different. The first time it was administered by the researcher in the hospital setting, and the second it was self-administered by the participants in their homes. This fact, combined with the low level of education of the majority of participants, could explain why the scores at the two times were not highly correlated. As stated previously, this was a less than ideal situation, but it was accepted to recruit subjects who were almost finished their rehabilitation period and would not be in the hospital five days later. Ideally, the same mode of administration should have been employed to ensure consistency in the results. In retrospect, a telephone-based administration for the retest of the SRNL-Index by the project leader would have been preferable. A previous study, that compared a structured telephone-based with a face-to-face administration of the RNL-Index in individuals with stroke and orthopedic conditions, demonstrated substantial agreement between these modes of administration[78]. This study did not show test-retest reliability of the SRNL-Index: this should be reevaluated in subsequent studies.

Validity

Discriminant validity

As anticipated, the correlation between the Perception of Self subscale was stronger with the Mental Component Summary (MCS) score of the SF-36 than with the Physical Component Summary (PCS) score in both the stroke and the TKA samples. This was expected, as an individual with a low psychological quality of life (as measured by the MCS) should score lower in terms of the perception of him or herself than in terms of perceived physical or daily activities. Alternatively, the Daily Activity subscales correlated more strongly with the PCS score of the SF-36 than with its MCS score for TKA subjects. Again this is to be anticipated, as a subject with better physical quality of life (as measured by the PCS) should reintegrate to normal daily activity levels more easily. Correlations between the Daily Activity subscale and the SF-36's MCS and PCS scores were identical in the stroke sample. This may be due to the physical and the mental component of HRQOL (as measured by the SF-36) being equally affected in stroke subjects. A recent study reported that Spanish stroke subjects, when measured by the SF-

36, demonstrated lower perception of health in all its areas compared to the general Spanish population[79], meaning that both mental and physical aspects of HRQOL were greatly affected in the stroke sample. Moreover, depression is common in stroke survivors: "depression is the most frequently occurring psychiatric disorder after stroke and has been shown to inhibit physical and cognitive recovery"[80] (p. 1006).

Another reason that could explain why correlations between the SRNL-Index and the component summary scores of the SF-36 did not always behave as hypothesized, or were not as strong as one might expect, may be related to the composition of the SRNL-Index subscales themselves[81]. Rarely is the RNL-Index divided into subscales. A recent study exploring the structure of the RNL-Index indicated the presence of two subscales called "social" and "physical" that did not correspond to the initially suggested subscales. This finding may affect the psychometric results involving the subscales. But as the analyses involving the total SRNL-Index were of main interest in the present study, the conclusions drawn about the psychometric properties of the SRNL-Index remain unaffected by the composition of the subscales.

Convergent validity

The original RNL-Index was compared to Spitzer's Quality of Life Index [3] to assess convergent validity; this questionnaire had never been translated into Spanish so could not be used. The 6MWT is used internationally, including in Spain. As hypothesized, the correlations between the mean total score of the SRNL-Index and the mean distance walked during the 6MWT were moderately high and in the expected direction for both stroke and TKA samples. Further, when reviewing correlations between the subscales and the 6MWT, it was observed that the correlation between the Daily Activity subscale and the distance walked during the 6MWT was greater than the correlation between the Perception of Self subscale and the 6MWT in both samples. This supports what one would expect; subjects who are able to walk further in six minutes should also be more functional and able to participate in their normal daily activities, than those only able to walk shorter distances.

Known-groups validity

Significant differences were seen across all BI independence groups in the stroke sample, except for the highest functioning groups. However, it must be noted that all groups were relatively high functioning, as all subjects were able to ambulate. Additionally, the BI has a recognized ceiling effect in high functioning groups. "It is less useful for assessing minor deficits at a high functional level, and more useful for differentiating between subjects with more severe disabilities"[82] (p.2056). In this study, the BI scores were taken directly from the medical charts by the director of the rehabilitation department of the Hospital de l'Esperança. The BI score had been assigned previously by either a nurse or a rehabilitation physician; however the time when the score had been entered was not as clear. It was supposedly judged when patients were in the outpatient rehabilitation unit, but whether this was taken at the first assessment as an outpatient, in the middle of the rehabilitation period or when approaching discharge was not known. Considering that the stroke sample was evaluated, on average, 22 days after discharge as an inpatient, the BI might not have been the best measure to use to assess known-groups validity. Moreover, since the stroke sample was small, further analyses of known-groups validity were undertaken using walking speed as an outcome. A point must first be raised about walking speed. Although the literature recognizes walking speed as a valid and reliable measure of physical function, and although subjects have been classified according to their walking speed[77], the speed is usually calculated over a short distance. The decision to use walking speed as an indicator for known-groups validity was reached after the data collection phase, when walking speed for subjects was unavailable. We, therefore, estimated walking speed from the 6MWT data. This could be a source of bias, as the long distance walked could consist of different walking speeds. In the literature, some authors indicate that gait speed might be underestimated when calculated from 6MWT[83], but others state that the gait speed can be accurately estimated from this test[84, 85]. Regardless of whether there is a bias or not, psychometric analyses are still accurate as the possible bias was applied equally across all subjects, and all calculations are equally affected.

Significant differences in SRNL-Index scores were seen across groups for all the walking speed categories of the stroke sample, and for all of the TKA sample categories except

between the fastest groups. The later might be related to the transformation of 6MWT distance into walking speed, but could also be explained by the low variability and important ceiling effect in this sample.

LIMITATIONS

As in any validation study, several limitations must be considered when evaluating the results. A first limitation is related to the samples. When the Spanish psychometric study was conceived, it was planned to obtain Spanish samples that were similar to those used by the developers in order to compare data. But even with the collaboration of the involved institutions, the timing and the duration of the study made obtaining subjects with the intended diagnoses impossible. As a consequence, the sample populations cannot be assumed to be the similar to those of the original study, and direct comparisons to the original study of the RNL-Index cannot be made. Test-retest reliability results were possibly more affected those related to the examination of validity.

A second limitation is related to the recruitment. The recruitment and administration of the measures were done during the summer, when the reduced activity of some departments and the lack of personnel made it difficult to find participants. Moreover, subjects were not accustomed to being recruited by a non-physician, and especially by someone not involved in their direct care. The low supervision and lack of resources available to conduct the study made the recruitment especially difficult. Conditions for the recruitment of subjects were less than optimal, and may have affected the results. This did not seem to be as big an issue for stroke subjects, who were mainly referred by their occupational therapist. The different referral mode may explain the different variability observed in the two samples.

Another limitation is related to sample size. As a result of the recruitment difficulties, the sample was adequate for reliability analyses[86], but not for validity[86]. A reevaluation of the psychometric properties should, therefore, be considered with larger samples. As the reliability studies are those with poorer results, another study with a larger sample size could also help reevaluate the existence of these reliability problems.

Additionally, results cannot be generalized to other diagnoses or for additional Spanish-speaking countries. Further studies should be conducted with other groups of subjects in order to do so.

Finally, responsiveness was not evaluated in this study. This should be undertaken in subsequent studies so that all the psychometric properties of the SRNL-Index have been examined.

In summary, validity has been demonstrated for the SRNL-Index and it has shown acceptable internal consistency. Test-retest reliability was not demonstrated. Only after this property has been retested and a formal assessment of responsiveness has been undertaken, will the 'Índice de Reintegración a la Vida Normal' be advocated for use in the clinical or research setting as a measure of participation restriction for use in Spain.

CONCLUSIONS

We set out to translate and culturally adapt the RNL-Index into Spanish for use in Spain and to test its reliability and validity. The SRNL-Index has demonstrated partially acceptable measurement properties for subjects with stroke and total knee arthroplasty. However, further studies are needed to assess all psychometric properties in other populations and other Spanish-speaking countries.

ACKNOWLEDGEMENTS

I wish to thank Lynne Nadeau, Susan Scott, Gemma Vilagut, and Dr. Lois Finch for their assistance with the statistical analyses, and the translators and backtranslators, Carlos Primo, Patricia Monroy Córdoba, Margaret Roper, and Albert Sanchez. Special thanks must be given to Dr. Roser Belmonte, Cristina Cervantes, and Nohora Rueda Moreno, without whom the project could not have been realised at the Hospital de la Esperança.

TABLES

Table 1. Sociodemographic characteristics in samples with stroke and total knee arthroplasty

Characteristic	Stroke (n=32)	TKA (n=36)
Age		
Mean age (SD), yr.	63 (11)	73 (7)
Gender, n (%)		
Men	20 (62.5)	14 (38.9)
Women	12 (37.5)	22 (61.1)
Education, n (%)		
Can barely read or write	----	1 (2.8)
Incomplete primary education	9 (28.13)	16 (44.4)
Primary education completed	8 (25.0)	12 (33.3)
Part of secondary education completed	----	3 (8.3)
Pre-university education completed	5 (15.6)	----
Professional training	2 (6.25)	----
Undergraduate university studies (diploma or Bachelor's)	8 (25)	2 (5.5)
Graduate university studies	----	1 (2.8)
Marital Status, n (%)		
Never married	3 (9.4)	4 (11.1)
Married	21 (65.6)	19 (52.8)
Separated or divorced or widowed	8 (25)	11 (30.6)
Dwelling, n (%)		
At home	30 (93.8)	30 (83.3)
At a family member's house	2 (6.3)	3 (8.3)
At a residence or convalescence hospital	----	1 (2.8)
Living arrangements, n (%)		
Alone	5 (15.6)	9 (25.0)
With spouse	21 (65.6)	17 (47.2)
With family members	6 (18.8)	8 (22.2)
Maternal language, n (%)		
Spanish	23 (71.9)	24 (66.7)
Catalan	8 (25)	10 (27.8)
Other	1 (3.13)	----
Province of origin, n (%)		
Catalonia	17 (53.1)	14 (38.9)
Andalusia	3 (9.4)	9 (25.0)
Madrid	3 (9.4)	3 (8.3)
Aragon	1 (3.1)	3 (8.3)
Valencia	----	2 (5.6)
Extremadura	2 (6.25)	----
Murcia	1 (3.1)	1 (2.8)
Castile – La Mancha	1 (3.1)	----
Basque Country	1 (3.1)	----
Galicia	1 (3.1)	----

*TKA: total knee arthroplasty; n: number of subjects; SD: standard deviation; yr.: year.

Table 1 cont. Sociodemographic characteristics in samples with stroke and total knee arthroplasty

Characteristic	Stroke (n=32)	TKA (n=36)
Current employment status, n (%)		
Working	2 (6.3)	----
Not working for ≥ 3 months	12 (37.5)	1 (2.8)
Housewife	1 (3.1)	12 (33.3)
Incapacitated with permanent disability	2 (6.25)	2 (5.6)
Retired due to age	15 (46.9)	18 (50.0)
Retired early	----	2 (5.6)
Self-reported level of activity, n (%)		
Sits most of the day	1 (21.9)	8 (22.2)
Stands up most of the day, without moving	1 (3.1)	5 (13.9)
Activities without an important physical effort, but walks frequently	24 (75)	22 (61.1)
Activities with important physical effort	----	1 (2.8)
Mean (SD)		
6MWT (m)	197.3 (104.2)	144.9 (48.0)
Gait speed (m/min)	32.9 (17.4)	24.1 (8.0)
Barthel Index	76.1 (17.3)	----
Time since discharge as inpatient (days)	31.7 (19.4)	12.3 (6.8)
Time since stroke or TKA operation (days)	61.2 (22.5)	21.8 (9.4)

*TKA: total knee arthroplasty; n: number of subjects; SD: standard deviation; yr.: year; m/min: meters per minute; 6MWT: Six Minute Walk Test; BI not available for TKA sample.

Table 2. Descriptive statistics for the SRNL-Index in samples with stroke and total knee arthroplasty

SRNL-Index (0-44)	Daily Activity subscale		Perception of Self subscale		Total Score	
Study sample	Stroke (n=32)	TKA (n=36)	Stroke (n=32)	TKA (n=36)	Stroke (n=32)	TKA (n=36)
Mean (SD)	22.3 (5.9)	22.9 (5.9)	10.2 (2.2)	11.6 (0.7)	32.5 (7.8)	34.5 (6.3)
Observed range	8-32	14-32	3-12	10-12	11-44	24-44
Theoretical range	8-32	8-32	3-12	3-12	11-44	11-44
Floor Effect (%)	3.1	0	3.1	0	3.1	0
Ceiling Effect (%)	3.1	8.3	34.4	63.9	3.1	8.3

*SRNL-Index: Spanish Reintegration to Normal Living Index; TKA: total knee arthroplasty; n: number of subjects; SD: standard deviation.

Table 3. Internal consistency: Standardized item alphas and Cronbach's alphas for subscales and the SRNL-Index in samples with stroke and total knee arthroplasty

Study Sample		**Stroke (n=32)**	**TKA (n=36)**
		Standardized alpha	Standardized alpha
Item			
1	Moving in living quarters	0.94	0.88
2	Moving in community	0.93	0.87
3	Taking trips	0.93	0.87
4	Self-care needs	0.93	0.89
5	Activity necessary and important	0.94	0.88
6	Recreation	0.93	0.88
7	Social activities	0.93	0.87
8	Family role	0.93	0.88
9	Comfortable with relationships	0.93	0.89
10	Comfortable with social self	0.93	0.89
11	Dealing with life events	0.93	0.87
Subscales	Daily Activity	0.92	0.86
	Perception of Self	0.93	0.87
Standardized Global Alpha		0.94	0.88

*TKA: total knee arthroplasty; n: number of subjects.

Table 4. Test-retest reliability: Intraclass correlation coefficients for total SRNL-scores in samples with stroke and total knee arthroplasty

	Stroke (n=28)*			**TKA (n=28)***		
		95% Confidence Intervals			**95% Confidence Intervals**	
Total score	**ICC (1,2)**	**Lower limit**	**Upper limit**	**ICC (1,2)**	**Lower limit**	**Upper limit**
	0.55	0.05	0.79	0.22	-0.67	0.64

*SRNL-Index: Spanish Reintegration to Normal Living Index; TKA: total knee arthroplasty; n: number of subjects returning the 2nd SRNL-Index; ICC: Intraclass correlation coefficient.

Table 5. Discriminant validity: Correlation of SRNL-Index's Daily Activity and Perception of Self subscales with SF-36's Physical and Mental Summary Component scores in samples with stroke and total knee arthroplasty

Study Sample	**Stroke (n=32)**		**TKA (n=36)**	
SRNL-Index's	Daily Activity subscale	Perception of Self subscale	Daily Activity subscale	Perception of Self subscale
SF-36's				
Physical Component Summary (PCS) score	0.55	0.22	0.64	0.35
Mental Component Summary (MCS) score	0.57	0.57	0.43	0.64

*SRNL-Index: Spanish Reintegration to Normal Living Index; SF-36: Short Form 36-item Health Survey; TKA: total knee arthroplasty; n: number of subjects.

Table 6. Convergent validity: Correlation of SRNL-Index with 6MWT in samples with stroke and total knee arthroplasty

	6MWT (SD)	**Pearson Correlation Coefficients (with SRNL-Index)**		
		Total score	**Daily Activity**	**Perception of**
Study Sample			**Subscale**	**Self subscale**
Stroke (n=32)	197.3 (104.2)	0.70	0.76	0.45
TKA (n=36)	144.9 (47.9)	0.61	0.60	0.41

*6MWT: Six Minute Walk Test; SRNL-Index: Spanish Reintegration to Normal Living Index; TKA: total knee arthroplasty; n: number of subjects; SD: standard deviation.

Table 7. Known-groups validity: Discriminative ability of the SRNL-Index among BI groups in samples with stroke and among walking speed groups in samples with stroke and total knee arthroplasty

	Stroke (n=32) SRNL-Index (11-44)				TKA (n=36) SRNL-Index (11-44)			
	n	**BI or speed (m/min) (SD)**	**Mean**	**95% CI**	**n**	**BI or speed (m/min) (SD)**	**Mean**	**95% CI**
Barthel Index categories								
Totally dependent (BI≤60)	5	51.4 (6.6)	20.2	13.0-27.4				
Slightly dependent (61≤BI≤90)	15	77.9 (9.8)	33.9	31.1-36.7				
Independent (BI≥91)	12	97.7 (2.8)	35.8	31.5-40.0				
Walking speed categories								
Household (<24 m/min)	12	16.9 (4.2)	25.6	21.1-30.1	19	17.8 (4.1)	30.8	28.7-32.9
Limited community (24 to <35 m/min)	7	29.6 (2.3)	32.4	30.2-34.6	11	28.7 (3.1)	37.7	34.5-41.0
Community (≥35 m/min)	13	49.4(14.2)	38.8	36.3-41.4	6	35.9 (1.2)	40.0	32.9-47.3

*TKA: total knee arthroplasty; n: number of subjects; BI: Barthel Index; SD: standard deviation; CI: confidence interval; m/min: meters/minute; shaded area: BI unavailable.

APPENDIX 1 (for Manuscript):
The RNL-Index (4 point-Likert scale version)

This questionnaire asks about how you manage activities, roles and relationships on a daily basis. The information will help keep track of how you feel and how well you are able to do your usual activities since your illness or injury.

Your reply to this questionnaire is confidential. Your identity will be known only to members of the research team, and the information you provide will not be able to be traced.

Your task is to respond to each of the following statements by placing an (X) in the box that most closely describes your situation over the past month. It is important that you respond to every statement as best you can. There are no correct answers.

The RNL-Index	Does not describe my situation	Somewhat describes my situation	Mostly describes my situation	Describes my situation
1. I move around my living quarters as I feel is necessary. (Wheelchairs, other equipment or resources may be used).				
2. I move around my community as I feel is necessary. (Wheelchairs, other equipment or resources may be used).				
3. I am able to take trips out of town as I feel are necessary. (Wheelchairs, other equipment or resources may be used).				
4. I am comfortable with how my self care needs (dressing, feeding, toileting, etc.) are met. (Adaptive equipment, supervision and/or assistance may be used).				
5. I spend most of my days occupied in a work activity that is necessary or important to me. (Work activity could be paid employment, housework, volunteer work, school, etc.). (Adaptive equipment, supervision and/or assistance may be used).				
6. I am able to participate in recreational activities (hobbies, sports, reading, computers, etc.) as I want to. (Adaptive equipment, supervision and/or assistance may be used).				
7. I participate in social activities with friends, family and /or business acquaintances as is necessary or desirable to me. (Adaptive equipment, supervision and/or assistance may be used).				
8. I assume a role in my family which meets my needs and those of other family members. (Family means people with whom you live and/or relatives with whom you don't live but see on a regular basis). (Adaptive equipment, supervision and/or assistance may be used).				
9. In general, I am comfortable with my personal relationships.				
10. In general, I am comfortable with myself when I am in the company of others.				
11. I feel that I can deal with life events as they happen.				

APPENDIX 2 (for Manuscript):
The SRNL-Index (4 point-Likert scale version)

Este cuestionario pregunta sobre cómo se desenvuelve en las actividades, funciones y relaciones sociales durante su día a día. La información proporcionada ayudará a seguir una pista de cómo se siente y de su capacidad a la hora de realizar actividades cotidianas desde su enfermedad o desde que sufrió un daño físico.

Sus respuestas a este cuestionario son confidenciales. Su identidad sólo la sabrá los miembros del equipo de investigación y la información que usted provea no será utilizada para ponerse de nuevo en contacto con usted.

Debe responder a cada una de las siguientes frases escribiendo una (X) en el recuadro que mejor describa su situación durante el mes pasado. Es muy importante que su respuesta se corresponda lo mayor posible a su opinión. No hay una respuesta correcta, todas son validas.

The SRNL-Index or Índice de reintegración la vida normal (Índice RVN)	No describe mi situación	Describe un poco mi situación	Describe en gran parte mi situación	Describe bien mi situación
1. Me muevo dentro de mi casa cuando lo considero necesario. (Con sillas de ruedas u otros aparatos o recursos si es necesario).				
2. Me muevo por mi vecindario cuando lo considero necesario. (Con sillas de ruedas u otros aparatos o recursos si es necesario).				
3. Puedo viajar fuera de mi ciudad cuando lo considero necesario. (Con sillas de ruedas u otros aparatos o recursos si es necesario).				
4. Me siento a gusto con cómo mis cuidados básicos se llevan a cabo (comida, ropa, aseo, etc.). (Con equipos adaptados, supervisión o ayuda si es necesario).				
5. Paso la mayoría de mis días trabajando en actividades necesarias o importantes para mí. (Trabajo remunerado, labores de casa, voluntariado, cursos, etc.).(Con equipos adaptados, supervisión o ayuda si es necesario).				
6. Puedo participar en actividades de ocio cuando lo deseo (hobbies, deportes, lectura, ordenador, etc.). (Con equipos adaptados, supervisión o ayuda si es necesario).				
7. Participo en actividades sociales con amigos, familia o compañeros de trabajo cuando lo deseo o lo creo necesario. (Con equipos adaptados, supervisión o ayuda si es necesario).				
8. Asumo el papel dentro de mi familia que responde a mis necesidades y a las de otros miembros de mi familia. (Se entiende como familia a las personas con las que vive o los familiares con los que no vive, pero ve regularmente). (Con equipos adaptados, supervisión o ayuda si es necesario.).				
9. En general, me siento bien con mis relaciones personales.				
10. En general, me siento bien conmigo mismo cuando estoy en compañía de otros.				
11. Creo que puedo afrontar los acontecimientos de la vida a medida que suceden.				

CHAPTER 5:
CONCLUSIONS

5.1 Summary and Discussion

The definition of participation recently stated in the International Classification of Functioning, Disability, and Health (ICF) is "involvement in life situations"[1]. While there are a few measures developed in English that are able to assess restriction in "Participation", no measure has yet been developed or translated into Spanish to assess this construct. The translation and cultural adaptation of the widely used measure, the Reintegration to Normal Living (RNL-) Index[2, 4], will provide a standardized Spanish instrument to evaluate reintegration and participation in life events following a disabling health condition.

The study had two distinct parts: 1) to translate and culturally adapt the RNL-Index into Spanish, 2) to evaluate the psychometric properties of the Spanish version of the RNL-Index (SRNL-Index).

The methods used for the translation and the cultural adaptation of the RNL-Index included two independent forward translations into Spanish; a review of the translations by a professional panel and consolidation into a preliminary version; a cognitive debriefing of this version by 10 subjects with stroke, a total knee replacement or cancer; two backtranslations of the debriefed version into Spanish; and a careful comparison with the original version. In this process, consensus was obtained by the professional panel for all items and response choices by either selecting one of the translations provided or a combination of both. Cognitive debriefing with ten subjects led to the change of a word in items 9 and 10. The backtranslations yielded no differences in content with the original RNL-Index, and as a consequence no changes were made to the Spanish version after the cognitive debriefing.

The psychometric evaluation of the SRNL-Index was a cross-sectional study (with the exception of the test-retest portion, which was longitudinal) of a sample of 32 subjects

with stroke and 36 subjects with total knee arthroplasty (TKA), assessed ten days to three months following discharge from Hospital de l'Esperança. Sociodemographic information was recorded, and subjects were evaluated with the SRNL-Index, the Short Form 36-item Health Survey (SF-36), and the Six Minute Walk Test (6MWT). Subjects with stroke also had their Barthel Index score recorded from their medical files. Walking speed was calculated from the 6MWT results for all participants. An additional copy of the SRNL-Index was given to the participants with the instruction to complete it and return it five days later.

Participants came from a variety of other Spanish provinces, so results can be generalized to Spain. The average SRNL-Index scores for persons with stroke and with total knee arthroplasty (TKA) were fairly similar, 32.5 and 34.5 out of a maximum score of 44, respectively. People with stroke scored throughout the theoretical range; however, a ceiling effect was seen with the TKA sample, particularly in the Perception of Self subscale. This was probably caused by a selection bias, as the ceiling effect was mostly seen in only one of the samples. Studies with bigger samples should be conducted to determine if the ceiling effect seen in the Spanish version of the Perception of Self subscale is persistent.

Two types of reliability, internal consistency and test-retest, were assessed. Internal consistency reliability was estimated using Cronbach's alpha coefficients. Correlations were also calculated. For the TKA sample, inter-item correlations were variable, sometimes null and even negative. This can be explained by the low variability and the ceiling effect seen in the TKA sample. The overall Cronbach's alphas for both samples were high, 0.94 and 0.88 respectively. Considering item-to-total, subscales-to-total and Cronbach's alpha values, the SRNL-Index is internally reliable.

Test-retest reliability was calculated using the scores of the first administration of the SRNL-Index and the scores that were self-administered by the subjects, on average within five days. Eighty-eight percent of the participants with stroke and 78% of those with TKA returned the second SRNL-Index questionnaire. Intraclass correlation coefficients (ICC)

were moderately low for the TKA sample and moderate for the Stroke sample. This can be explained by the different modes of administration and the different settings of the first and the second administrations of the SRNL-Index. In retrospect, a telephone-based administration for the retest of the SRNL-Index would have been preferable. Test-retest reliability was not demonstrated, and should be reevaluated in subsequent studies under more controlled conditions.

The types of construct validity assessed were discriminant, convergent, and known-groups validity. Discriminant validity was estimated by administering the SRNL-Index and the SF-36, using Pearson correlation coefficients. As anticipated, the correlation between the Perception of Self subscale score was stronger with the Mental Component Summary (MCS) Score of the SF-36 than with its Physical Component Summary (PCS) Score for both the stroke and the TKA samples. The Daily Activity subscales correlated more strongly with the PCS Score of the SF-36 than with its MCS Score for TKA subjects. However, correlations between the Daily Activity subscale and the SF-36 scores were nearly identical for the stroke sample. This may be explained by the nature of the condition, as depression and other psychogenic disorders are common following a stroke and could affect the emotional quality of life of the stroke sample. It could also be explained by the composition of the RNL-Index subscales themselves, as recent studies put into question the originally suggested subscales[81].

Convergent validity was estimated by correlating scores of the the SRNL-Index and the 6MWT, using Pearson correlation coefficients. As hypothesized, the correlations between the mean total score of the SRNL-Index and the mean distance walked during the 6MWT were moderately high and in the expected direction for both stroke and TKA samples, 0.70 and 0.61 respectively. Correlations of the 6MWT with the Daily Activity subscale were high for both samples, but only a moderate correlation was observed in both samples with the Perception of Self subscale. This was expected, as subjects who are able to walk faster and cover more distance in six minutes should also be more functional and able to participate more in their normal daily activities.

Known-groups validity was assessed twice. In the first instance, subjects with stroke were classified in three categories according to their Barthel Index (BI) score, using known cut-off scores[72]. A general linear model with post-hoc t tests estimated whether the SRNL-Index could discriminate between subjects across three levels of independence measured by the BI with an alpha level set at 0.05. Significant differences were seen between all groups except between the middle and highly independent group, which could be explained by the ceiling effect in high functioning stroke subjects when measured by the BI[82].

Since the stroke sample was small, further analyses of known-groups validity were undertaken using walking speed as an outcome. Significant differences in SRNL-Index scores were seen across groups for all the walking speed categories of the stroke sample, and for all of the TKA sample categories except between the fastest groups. The later might be related to the transformation of 6MWT distance into walking speed, but could also be explained by the low variability and important ceiling effect in this sample.

5.2 Limitations and Future Research

As in any study, limitations must be considered when examining the results. The main difficulty is related to doing research in a foreign country where physicians are essentially the only researchers. Although some difficulties occurred because of the timing and the duration of the study, and even with the collaboration of the involved institutions, subjects were not accustomed to being recruited for a study by a non-physician, especially if not involved in their direct care. Conditions for the recruitment of subjects were less than optimal, and may have affected the results obtained.

Another limitation is related to sample size. As noted earlier, reevaluation of the psychometric properties should be considered with larger samples. Moreover, results cannot be generalized to other diagnoses and for other Spanish-speaking countries. Finally, responsiveness was not evaluated in this study. This should be undertaken in subsequent studies so that all the psychometric properties of the SRNL-Index have been examined.

Nonetheless, even with the limitations, sufficient validity and internal consistency has been demonstrated for an eventual implementation of the SRNL-Index into clinical and research practice in Spain. Test-retest reliability and responsiveness studies should, however, be undertaken.

REFERENCES

(1) World Health Organization. *International classification of impairments, disabilities, and handicaps: a manual of classification relating to the consequences of disease*. Geneva: World Health Organization; 1980.

(2) Wood-Dauphinee S, Williams JI. Reintegration to Normal Living as a proxy to quality of life. *J Chronic Dis* 1987; 40(6):491-502.

(3) Spitzer WO, Dobson AJ, Hall J, Chesterman E, Levi J, Shepherd R, Battista RN, Catchlove BR. Measuring the quality of life of cancer patients: a concise QL-index for use by physicians. *J Chronic Dis* 1981; 34(12):585-97.

(4) Wood-Dauphinee SL, Opzoomer MA, Williams JI, Marchand B, Spitzer WO. Assessment of global function: The Reintegration to Normal Living Index. *Arch Phys Med Rehabil* 1988; 69(8):583-90.

(5) Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to their Development and Use*. Oxford: Oxford University Press; 1995.

(6) Trombly CA, Radomski MV, Davis ES. Achievement of self-identified goals by adults with traumatic brain injury: Phase I. *Am J Occup Ther* 1998; 52(10):810-8.

(7) Clarke PJ, Black SE, Badley EM, Lawrence JM, Williams JI. Handicap in stroke survivors. *Disabil Rehabil* 1999; 21(3):116-23.

(8) Doward LC, McKenna SP. Defining patient-reported outcomes. *Value Health* 2004; 7 Suppl 1:S4-S8.

(9) Bilbao A, Kennedy C, Chatterji S, Ustun B, Barquero JLV, Barth JT. The ICF: Applications of the WHO model of functioning, disability and health to brain injury rehabilitation. *Neurorehab* 2003; 18(3):239-50.

(10) Granlund M, Eriksson L, Ylven R. Utility of International Classification of Functioning, Disability and Health's participation dimension in assigning ICF codes to items from extant rating instruments. *J Rehabil Med* 2004; 36(3):130-7.

(11) Bullinger M, Alonso J, Apolone G, Leplege A, Sullivan M, Wood-Dauphinee S, Gandek B, Wagner A, Aaronson N, Bech P, Fukuhara S, Kaasa S, Ware JE, Jr. Translating health status questionnaires and evaluating their quality: the IQOLA Project approach. International Quality of Life Assessment. *J Clin Epidemiol* 1998; 51(11):913-23.

(12) Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000 15; 25(24):3186-91.

(13) Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol* 1993; 46(12):1417-32.

(14) Bullinger M. The challenge of cross-cultural quality of life assessment. *Psychology & Health* 1997; 12(6):815-25.

(15) Ware JE, Gandek B. Methods for testing data quality, scaling assumptions, and reliability: The IQOLA project approach. *J Clin Epidemiol* 1998; 51(11):945-52.

(16) Kucukdeveci AA, McKenna SP, Kutlay S, Gursel Y, Whalley D, Arasil T. The development and psychometric assessment of the Turkish version of the Nottingham Health Profile. *Int J Rehabil Res* 2000; 23(1):31-8.

(17) Eremenco SL, Cella D, Arnold BJ. A comprehensive method for the translation and cross-cultural validation of health status questionnaires. *Eval Health Prof* 2005; 28(2):212-32.

(18) Hunt SM, Alonso J, Bucquet D, Niero M, Wiklund I, McKenna S. Cross-cultural adaptation of health measures. European Group for Health Management and Quality of Life Assessment. *Health Pol* 1991; 19(1):33-44.

(19) McKenna SP, Doward LC. The translation and cultural adaptation of patient-reported outcome measures. *Value Health* 2005; 8(2):89-91.

(20) Jounal of Clinical Epidemiology, *J Clin Epidemiol* 1997; 50(11): 1189-1310.

(21) Gandek B, Ware JE. Methods for validating and norming translations of health status questionnaires: The IQOLA project approach. *J Clin Epidemiol* 1998; 51(11):953-9.

(22) Hilton A, Skrutkowski M. Translating instruments into other languages: Development and testing processes. *Cancer Nurs* 2002;25(1):1-7.

(23) Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, Erikson P. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health* 2005;8(2):94-104.

(24) Patrick DL, Gagnon DD, Zagari MJ, Mathijs R, Sweetenham J. Assessing the clinical significance of health-related quality of life (HrQOL) improvements in anaemic cancer patients receiving epoetin alfa. *Eur J Cancer* 2003;39(3):335-45.

(25) Hareendran A, Abraham L. Using a treatment satisfaction measure in an early trial to inform the evaluation of a new treatment for benign prostatic hyperplasia. *Value Health* 2005; 8:S35-S40.

(26) Wilkie R, Peat G, Thomas E, Croft PR. Measuring the consequences of osteoarthritis and joint pain in population-based studies: can existing health measurement instruments capture levels of participation? *Arthritis Rheum* 2004; 51(5):755-62.

(27) Wilkie R, Peat G, Thomas E, Hooper H, Croft PR. The Keele Assessment of Participation: a new instrument to measure participation restriction in population studies. Combined qualitative and quantitative examination of its psychometric properties. *Qual Life Res* 200; 14(8):1889-99.

(28) Ware JE, Jr. Conceptualization and measurement of health-related quality of life: comments on an evolving field. *Arch Phys Med Rehabil* 2003; 84(4 Suppl 2):S43-S51.

(29) Harwood RH, Prince M, Mann A, Ebrahim S. Associations between diagnoses, impairments, disability and handicap in a population of elderly people. *Int J Epidemiol* 1998; 27(2):261-8.

(30) Wilkie R, Peat G, Thomas E, Croft P. The prevalence of person-perceived participation restriction in community-dwelling older adults. *Qual Life Res* 2006; 15(9):1471-9.

(31) Cardol M, de Jong BA, van den Bos GA, Beelen A, de Groot IJM, de Haan RJ. Beyond disability: perceived participation in people with a chronic disabling condition. *Clin Rehabil* 2002; 16(1):27-35.

(32) Cardol M, de Haan RJ, de Jong BA, van den Bos GA, de G, I. Psychometric properties of the Impact on Participation and Autonomy Questionnaire. *Arch Phys Med Rehabil* 2001; 82(2):210-6.

(33) Cardol M, Beelen A, van den Bos GA, de Jong BA, de G, I, de Haan RJ. Responsiveness of the Impact on Participation and Autonomy questionnaire. *Arch Phys Med Rehabil* 2002; 83(11):1524-9.

(34) Sibley A, Kersten P, Ward CD, White B, Mehta R, George S. Measuring autonomy in disabled people: validation of a new scale in a UK population. *Clin Rehabil* 2006; 20(9):793-803.

(35) Lund ML, Fisher AG, Lexell J, Bernspang B. Impact on Participation and Autonomy Questionnaire: internal scale validity of the Swedish version for use in people with spinal cord injury. *J Rehabil Med* 2007; 39(2):156-62.

(36) van Brakel WH, Anderson AM, Mutatkar RK, Bakirtzief Z, Nicholls PG, Raju MS, Das-Pattanayak RK. The Participation Scale: Measuring a key concept in public health. *Disabil Rehabil* 2006; 28(4):193-203.

(37) Gray DB, Hollingsworth HH, Stark SL, Morgan KA. Participation survey/mobility: Psychometric properties of a measure of participation for impairments and limitations. *Arch Phys Med Rehabil* 2006; 87(2):189-97.

(38) Gandek B, Sinclair SJ, Jette AM, Ware JE. Development and initial psychometric evaluation of the participation measure for post-acute care (PM-PAC). *Am J Phys Med Rehabil* 2007; 86(1):57-71.

(39) McColl MA, Davies D, Carlson P, Johnston J, Minnes P. The community integration measure: development and preliminary validation. *Arch Phys Med Rehabil* 2001; 82(4):429-34.

(40) Willer B, Ottenbacher KJ, Coad ML. The community integration questionnaire. A comparative examination. *Am J Phys Med Rehabil* 1994; 73(2):103-11.

(41) Finch E, Canadian Physiotherapy Association. *Physical rehabilitation outcome measures : a.* 2 ed. 2002.

(42) Steiner A, Raube K, Stuck AE, Aronow HU, Draper D, Rubenstein LZ, Beck JC. Measuring psychosocial aspects of well-being in older community residents: performance of four short scales. *Gerontol* 1996; 36(1):54-62.

(43) Nissen SJ, Newman WP. Factors influencing reintegration to normal living after amputation. *Arch Phys Med Rehabil* 1992; 73(6):548-51.

(44) Stock SR, Cole DC, Tugwell P, Streiner D. Review of applicability of existing functional status measures to the study of workers with musculoskeletal disorders of the neck and upper limb. *Am J Ind Med* 1996; 29(6):679-88.

(45) Trigg R, Wood VA, Hewer RL. Social reintegration after stroke: the first stages in the development of the Subjective Index of Physical and Social Outcome (SIPSO). *Clin Rehabil* 1999; 13(4):341-53.

(46) Bethoux F, Calmels P, Gautheron V. Changes in the quality of life of hemiplegic stroke patients with time: a preliminary report. *Am J Phys Med Rehabil* 1999; 78(1):19-23.

(47) Carter BS, Buckley D, Ferraro R, Rordorf G, Ogilvy CS. Factors associated with reintegration to normal living after subarachnoid hemorrhage. *Neurosurg* 2000; 46(6):1326-33.

(48) Daverat P, Petit H, Kemoun G, Dartigues JF, Barat M. The long term outcome in 149 patients with spinal cord injury. *Parapleg* 1995; 33(11):665-8.

(49) Ruffolo CF, Friedland JF, Dawson DR, Colantonio A, Lindsay PH. Mild traumatic brain injury from motor vehicle accidents: factors associated with return to work. *Arch Phys Med Rehabil* 1999; 80(4):392-8.

(50) Boake C. Supervision rating scale: a measure of functional outcome from brain injury. *Arch Phys Med Rehabil* 1996; 77(8):765-72.

(51) Dawson DR, Levine B, Schwartz M, Stuss DT. Quality of life following traumatic brain injury: A prospective study. *Brain Cogn* 2000; 44(1):35-9.

(52) Friedland JF, Dawson DR. Function after motor vehicle accidents: a prospective study of mild head injury and posttraumatic stress. *J Nerv Ment Dis* 2001; 189(7):426-34.

(53) World Languages, April 20 2006, from http://www2.ignatius.edu/faculty/turner/worldlang.htm.

(54) Alonso J, Prieto L, Anto JM. [The Spanish version of the SF-36 Health Survey (the SF-36 health questionnaire): an instrument for measuring clinical results]. *Med Clin (Barc)* 1995; 104(20):771-6.

(55) Vilagut G, Ferrer M, Rajmil L, Rebollo P, Permanyer-Miralda G, Quintana JM, Santed R, Valderas JM, Ribera A, Domingo-Salvany A, Alonso J. [The Spanish version of the Short Form 36 Health Survey: a decade of experience and new developments.]. *Gac Sanit* 2005; 19(2):135-50.

(56) Chavez LM, Canino G, Negron G, Shrout PE, Matias-Carrelo LE, guilar-Gaxiola S, Hoppe S. Psychometric properties of the Spanish version of two mental health outcome measures: World Health Organization Disability Assessment Schedule II and Lehman's Quality Of Life Interview. *Ment Health Serv Res* 2005;7(3):145-59.

(57) Alonso J, Anto JM, Moreno C. Spanish Version of the Nottingham Health Profile - Translation and Preliminary Validity. *Am J Public Health* 1990; 80(6):704-8.

(58) Badia X, Roset M, Montserrat S, Herdman M, Segura A. [The Spanish version of EuroQol: a description and its applications. European Quality of Life scale]. *Med Clin (Barc)* 1999;112 Suppl 1:79-85.

(59) Ware JE, Jr., Keller SD, Gandek B, Brazier JE, Sullivan M. Evaluating translations of health status questionnaires. Methods from the IQOLA project. International Quality of Life Assessment. *Int J Technol Assess Health Care* 1995;11(3):525-51.

(60) IMIM, May 11, 2006, from www.imim.es/imim/eng.htm.

(61) IMAS, March 13, 2007, from www.imasbcn.com/hospitals/hospital-esperan/index.html.

(62) Wood-Dauphinee S, Mayo N, Jung H, Shen N. Does the SF-36 reflect changes in the health of people with stroke? Baltimore, Maryland, November 15-17, 1998. 5th Annual Conference of the International Society for Quality of Life Research. 1998.

(63) Cooper KH. A means of assessing maximal oxygen intake. Correlation between field and treadmill testing. *JAMA* 1968;203(3):201-4.

(64) Solway S, Brooks D, Lacasse Y, Thomas S. A qualitative systematic overview of the measurement properties of functional walk tests used in the cardiorespiratory domain. *Chest* 2001; 119(1):256-70.

(65) Podsiadlo D, Richardson S. The Timed Up and Go - A Test of Basic Functional Mobility for Frail Elderly Persons. *J Am Geriatr Soc* 1991; 39(2):142-8.

(66) Holden MK, Kozlowski BA. Relationship of Reciprocal Torque Production to Gait in Hemiplegia. *Phys Ther* 1984; 64(5):714.

(67) Holden MK, Gill KM, Magliozzi MR. Gait Assessment for Neurologically Impaired Patients - Standards for Outcome Assessment. *Phys Ther* 1986; 66(10):1530-9.

(68) Nakamura R, Handa T, Watanabe S, Morohashi I. Walking Cycle After Stroke. *Tohoku J Exp Med* 1988; 154(3):241-4.

(69) Bohannon RW, Andrews AW. Correlation of Knee Extensor Muscle Torque and Spasticity with Gait Speed in Patients with Stroke. *Arch Phys Med Rehabil* 1990; 71(5):330-3.

(70) Bohannon RW, Walsh S. Nature, Reliability, and Predictive Value of Muscle Performance-Measures in Patients with Hemiparesis Following Stroke. *Arch Phys Med Rehabil* 1992; 73(8):721-5.

(71) Fransen M, Crosbie J, Edmonds J. Reliability of gait measurements in people with osteoarthritis of the knee. *Phys Ther* 1997; 77(9):944-53.

(72) Granger CV, Dewis LS, Peters NC, Sherwood CC, Barrett JE. Stroke Rehabilitation - Analysis of Repeated Barthel Index Measures. *Arch Phys Med Rehabil* 1979; 60(1):14-7.

(73) Baro E, Ferrer M, Vazquez O, Miralles R, Pont A, Esperanza A, Cervera AM, Alonso J. Using the Nottingham Health Profile (NHP) among older adult inpatients with varying cognitive function. *Qual Life Res* 2006; 15(4):575-85.

(74) Gallego CF, Roger MR, Bonet IU, Vinets LG, Ribas AP, Pisa RL, Oriol RP. Validation of a questionnaire to evaluate the quality of life of nonprofessional caregivers of dependent persons. *J Adv Nurs* 2001; 33(4):548-54.

(75) Portney LG, Watkins MC. Statistical Measures of Reliability. *Foundations of Clinical Research: Applications to Clinical Practice*. 2 ed. New Jersey: Prentice Hall Health; 2000. p. 557-86.

(76) Nunally JC. *Psychometric theory*. 3 ed. New York: McGraw-Hill; 1994.

(77) Perry J, Garrett M, Gronley JK, Mulroy SJ. Classification of Walking Handicap in the Stroke Population. *Stroke* 1995; 26(6):982-9.

(78) Korner-Bitensky N, Wood-Dauphinee S, Siemiatycki J, Shapiro S, Becker R. Health-related information postdischarge: telephone versus face-to-face interviewing. *Arch Phys Med Rehabil* 1994; 75(12):1287-96.

(79) Marco E, Duarte E, Santos JF, Boza R, Tejero M, Belmonte R, Muniesa JM, Sebastia E, Samitier CB, Pou M, Guillen A, Escalada F. Short form 36 health questionnaire in hemiplegic patients 2 years after stroke. *Neurologia* 2006; 21(7):348-56.

(80) Chemerinski E. Neuropsychiatric disorders following vascular brain injury. Levine SR, editor. *Mount Sinai J Med* 2006; 73(7): 1006-1014.

(81) Stark SL, Edwards DF, Hollingsworth H, Gray DB. Validation of the reintegration to normal living index in a population of community-dwelling people with mobility limitations. *Arch Phys Med Rehabil* 2005; 86(2):344-5.

(82) Weimar C, Stausberg J, Kraywinkel K, Wagner M, Busse O, Haberl RL, Diener HC. Diagnosis related groups in stroke treatment. An analysis from the Stroke Data Bank of the German Stroke Foundation. *Dtsch Med Wochenschr* 2002; 127(31-32):1627-32.

(83) Dean CM, Richards CL, Malouin F. Walking speed over 10 metres overestimates locomotor capacity after stroke. *Clin Rehabil* 2001; 15(4):415-21.

(84) Barbeau H. Comparison of spees used for the 15.2-meter and 6-minute walks over the year after an incomplete spinal cord injury: the SCILT trial. Elashoff R, Deforge D, Ditunno J, Sauline M, Dobkin BH, editors. *Neurorehabil Neur Rep* 2007; 21(2): 302-306.

(85) Van Hedel, H.J.A, Dieyz V, Curt A. Assessment of walking speed and distance in subjects with an incomplete spinal cord injury *Neurorehabil Neur Rep* 2007; 21 (4): 295-301.

(86) Hobart J. What sample sizes for reliability and validity studies? *Qual of Lif Resear* 2002; 11: 636.

APPENDICES

A 1.0 MEASURES USED FOR PSYCHOMETRIC ANALYSES

The Reintegration to Normal Living (RNL)-Index

The Reintegration to Normal Living (RNL)-Index is a standardised self-administered measure to assess participation. The measure was originally developed in Canadian English and Canadian French with persons with cancer or heart disease[2, 4]. The RNL-Index has been shown to be valid[2, 7], reliable[42] and responsive[2] and has been revalidated or used with diverse conditions including malignant tumours, degenerative heart disease, arthritis, fractures, amputations, musculoskeletal disorders, stroke, traumatic brain injury, cerebral aneurysms, spinal cord injury, general physical disability, motor vehicle accidents, and with community-dwelling elderly[41-52].

The RNL-Index has 11 items that assess mobility, self-care, daily activity, recreation and social activities, family roles and relationships, as well as general coping skills and interpersonal relationships[2]. Each item is scored with a 100 point visual analogue scale or a three- or four-point Likert categorical scale. The four-point categorical scale was used in this study, scored 1-4 for a total score range of 11 to 44. Scores can also be obtained for two subscales, "Daily Activity" and the "Perception of Self".

Short Form 36-item Health Survey (SF-36)

The SF-36 is a generic Health-Related Quality of Life measure that assesses 8 health concepts: limitations in physical functioning or social activities due to health problems, limitations in usual roles activities due to physical health problems or emotional problems, pain, vitality, mental health, and perception of health in general. Each item is scored on a dichotomous, three or five-point categorical scale; scores range from 0 to 100. Physical and mental summary scores can also be constructed[62]. It was culturally adapted into Spanish by Alonso and colleagues as part of the International Quality of Life Assessment (IQOLA) project (1991-1995)[54, 55].

Six Minute Walk Test (6MWT)

The Six Minute Walk Test (6MWT) is a submaximal functional test of walking endurance[63]. The 6MWT has been studied in several different populations and is a valid and reliable measure[64], that is commonly used in Spain. In this study, the 6MWT was administered with a marked 15-meter path in a hospital corridor. The instructions to the subjects were: "Walk as quickly as you can for 6 minutes to cover as much ground as possible in this 15-meter path. You may rest if you need to, but continue walking as soon as possible". The distance walked was recorded. All subjects were ambulators, with or without walking aids, and were hence able to complete the 6MWT.

Walking Speed

Gait speed is a physical characteristic derived from directly measuring the parameters of distance and time. Gait speed is considered a valid measure of walking ability as it correlates with functional mobility, degree of independence in walking, and many different gait parameters[65-67]. It has been associated with strength of the affected lower extremity, cadence and stride length, balance, degree of lower extremity motor recovery, and functional mobility[65, 68-70]. Standardized instructions are to walk at a "comfortable" or "maximum" speed along a walkway typically ranging from 2 to 20 meters[71]. In this study, the distance walked during the 6MWT was divided by the time to arrive at a walking speed. Subjects were instructed to walk at a comfortable pace and walked over a distance of 15 meters.

Barthel Index (BI)

The Barthel Index (BI), a commonly used measure of activities of daily living in people with stroke, has well established reliability and validity[72]. There are 10 items: feeding, grooming, bathing, dressing, controlling bowel and bladder, personal toilet, ambulation, transfers and stair climbing. It has a maximal score of 100[72] and has been fully implemented clinically and in research in Spain[73, 74].

A 2.0 INCLUSION AND EXCLUSION CRITERIA

A2.1 Inclusion criteria

- Spanish as a mother tongue and living in Spain;
- Diagnosis of total knee replacement or stroke;
- Discharged from a period of ten days to three months from the Hospital de l'Esperança.

A2.2 Exclusion criteria

- Presence of life-threatening co-morbid conditions that could affect the return to normal living patterns;
- Poor cognitive status or inability to read that could affect completion of the measures;
- Age less than 18 years.

A 3.0 SAMPLE SIZE

Sample size was determined as recommended by Hobart (Hobart, 2002). This author undertook an empirical examination of the impact of sample size on reliability and validity estimates. Data obtained from a sample of 713 people with Multiple Sclerosis who completed the MS Impact Scale were analyzed. This sample was reduced continually by decrements of 50%, both randomly and non-randomly. All samples were analyzed for indicators of reliability and validity, and the original and reduced samples were compared in how they differed in their magnitude and psychometric interpretation. Psychometric analyses in small samples (reliability: n>20, validity: n>40) was determined to generate valuable information. However, in samples of n>36, results reflected those generated in the original sample (Hobart, 2002). Since missing data will not be included in the study and will be discarded, the 68 subjects enrolled in this study to make up for possible loss of subjects is sufficient.

A4.2 Consent form

HOJA DE INFORMACIÓN AL PACIENTE

Título: "TRADUCCIÓN, ADAPTACIÓN CULTURAL y REVALIDACIÓN DEL "REINTEGRATION TO NORMAL LIVING INDEX" PARA SU USO EN ESPAÑA"

La unidad de Investigación en Servicios Sanitarios del Institut Municipal d'Investigació Mèdica está realizando un estudio al cual le invitan a participar.

Se trata de un estudio para evaluar la traducción al español de un cuestionario que evalúa el retorno a la vida diaria después de una enfermedad o de un trauma, en el que participan los siguientes centros: Consorci Sanitari de Barcelona (Hospital Clínic Provincial de Barcelona) e IMAS (Hospital de la Esperanza). En este estudio, se evalúa la versión traducida al castellano de un cuestionario ya utilizado en varios países en inglés y en francés.

Para dicho estudio, se le realizará dos sesiones: una de aproximadamente 60 minutos, y otra que efectuará en su domicilio de unos 5 minutos. Su participación en el estudio consistirá en acudir al hospital y contestar, por escrito, a tres cuestionarios con preguntas sobre varios aspectos de su salud y a preguntas sobre sus características sociodemográficas. También se le hará una evaluación física, donde tendrá que caminar a un ritmo cómodo para usted durante seis minutos, y esto, tomando todas las pausas que usted necesite. A finalizar la sesión, le entregaremos un sobre prepagado con un segundo cuestionario que rellenará cinco días después del día de la sesión. Una vez completado el cuestionario, nos lo enviará por correo.

Este estudio no supone ningún gasto para usted, y trata de intentar conocer su grado de calidad de vida y aumentar la de futuros pacientes. Sus datos personales serán tratados con confidencialidad absoluta ya que no se le identificará de forma personalizada en ningún informe final ni en publicación alguna. Estos datos serán sólo tratados por el grupo de investigadores. Se podrá retirar del estudio cuando lo desee, sin tener que dar explicación alguna por ello. En todo tiempo, usted tendrá el derecho al acceso, a la rectificación y a la cancelación de sus datos personales.

Su participación en el estudio es totalmente voluntaria y, si decide no participar, no se verá afectada en ningún modo la asistencia que reciba, se decidirá el tratamiento en conformidad con usted, ya que como le hemos explicado cualquiera de los tratamientos son correctos para su caso.

Si está de acuerdo en formar parte del estudio, por favor firme el consentimiento informado.

CONSENTIMIENTO INFORMADO POR ESCRITO
(según modelo del Anexo 6, punto 2 del Real Decreto 561/1993 de 16-04-1993)

Título del estudio: "TRADUCCIÓN, ADAPTACIÓN CULTURAL y REVALIDACIÓN DEL "REINTEGRATION TO NORMAL LIVING INDEX" PARA SU USO EN ESPAÑA"

Yo, (nombre y apellidos)

_____

He leído la hoja de información que se me ha entregado.

He podido hacer preguntas sobre el estudio.

He recibido suficiente información sobre el estudio.

He hablado con (nombre del investigador):

_____

Comprendo que mi participación es voluntaria.

Comprendo que puedo retirarme del estudio:

1° Cuando quiera.

2° Sin tener que dar explicaciones

3° Sin que esto repercuta en mis cuidados médicos.

Presto libremente mi conformidad para participar en el estudio.

_____, de de 200__
(ciudad)

Firma del paciente _____

Firma del médico _____
Nombre y apellidos
N° de colegiado
Fecha