UNIVERSITE DE NANCY I FACULTE DE MEDECINE

DOCTORAT D'UNIVERSITE

GROUPE DE FORMATION DOCTORALE : RECHERCHE EN SANTE PUBLIQUE

MESURE ET DETERMINANTS DE L'INCAPACITE FONCTIONNELLE DANS LA POLYARTHRITE RHUMATOIDE RECENTE.

LA COHORTE EURIDISS.

Thèse présentée et soutenue publiquement

le **20 Décembre 1993** par

Francis GUILLEMIN

Né le 29 Decembre 1957 à Nancy (M&M)

COMPOSITION DU JURY:

Pr Michel MANCIAUX.....	Président
Pr Jean-François VIEL, Besançon.....	Rapporteur
Pr Wim van den HEUVEL, Groningen (Pays-Bas).....	Rapporteur
Pr Serge BRIANCON.....	Directeur de recherche
Pr Jacques POUREL.....	Juge
Pr Joseph LELLOUCH, Paris.....	Juge

SOMMAIRE

Pages

SYNTHESE

MESURE ET DETERMINANTS DE L'INCAPACITE FONCTIONNELLE DANS LA POLYARTHRITE RHUMATOIDE RECENTE - LA COHORTE EURIDISS.

L'incapacité fonctionnelle se définit comme une réduction partielle ou totale de la capacité à accomplir une activité d'une façon et dans les limites considérées comme normales pour un être humain.
Suivant le modèle de WOOD, elle s'inscrit dans le chemin de la déficience (conséquence de lésions organiques) au handicap (retentissement de l'incapacité dans la vie sociale). Pour prendre en compte le véritable poids de cette incapacité fonctionnelle dans la vie de tous les jours, et parce que la santé et ses représentations se situent plus dans le domaine des perceptions et des capacités que dans celui des constatations et des performances, elle sera envisagée comme un élément perçu par le patient, par opposition au jugement externe d'un observateur (proche, médecin). Ceci souligne toute la différence entre ce que l'un ressent avec la composante psychologique de sa propre appréciation, et ce que l'autre observe. Considérée de cette façon, l'incapacité fonctionnelle est bien une dimension majeure de la qualité de vie dans le champ des maladies rhumatismales. C'est donc en tant que dimension ou composante de la qualité de vie qu'elle doit être étudiée.

Ce travail a pour objectif d'améliorer la connaissance de l'incapacité fonctionnelle dans la polyarthrite rhumatoïde (PR) au stade de début en utilisant des instruments valides dans cette période de la maladie qui permette de caractériser cette incapacité fonctionnelle lorsqu'elle apparaît, d'identifier ses déterminants pour en déduire des stratégies précoces en vue d'une meilleure prise en compte des différents aspects de la maladie et de l'accompagnement des patients. Il développe tout d'abord une méthodologie permettant la mesure de l'incapacité fonctionnelle de façon équivalente à travers des cultures différentes et de l'appliquer à la polyarthrite rhumatoïde (PR) récente. L'application d'un instrument de mesure valide de l'incapacité fonctionnelle, le Health Assessment Questionnaire (HAQ), dans une

cohorte européenne de polyarthrite rhumatoïde évoluant depuis moins de 5 ans, la cohorte EURIDISS, illustre ces méthodes.

MESURE DE L'INCAPACITE

L'utilisation de mesures de qualité de vie développées dans une culture donnée pose le problème de leur validité dans une autre culture. Le besoin d'établir des comparaisons internationales, le développement des études multicentriques justifiaient la mise au point d'une méthodologie permettant, grâce à un processus d'adaptation dépassant le cadre d'une simple traduction, de disposer d'instruments de mesure équivalents (et non identiques) mesurant les mêmes concepts chez des sujets de culture différentes. Une revue des méthodologies développées en sociologie et en psychologie a permis d'identifier des étapes nécessaires à cette équivalence et de formuler des recommandations. Une classification taxonomique des niveaux d'équivalence a été choisie: équivalence sémantique, idiomatique, événementielle et conceptuelle. La validité et la faisabilité de ces recommandations a été testée sur des études récemment publiées et l'accord entre juges (mesure par l'indice de concordance kappa et le coefficient de corrélation intra-classe) témoigne de leur faisabilité (**Journal of Clinical Epidemiology**).

Ces recommandations ont été appliquées à l'adaptation en français du Health Assessment Questionnaire (HAQ[1]), mettant ainsi un instrument validé en français à la disposition des chercheurs (**Revue du Rhumatisme**). Un préalable à la mesure de l'incapacité fonctionnelle dans la PR récente a consisté à vérifier la validité de cet instrument à ce stade précoce de la maladie, information dont on ne disposait pas avec certitude dans la littérature. Les propriétés originelles de l'instrument - validité de contenu, de structure, reproductibilité, validité externe et discriminante - étaient conservées de façon satisfaisante (**Disability and Rehabilitation**).

La comparaison de cet instrument spécifique de la PR avec un instrument générique (Nottingham Health Profile, NHP) a montré que leurs propriétés de validité étaient similaires et que le NHP mesurait plus complètement la qualité de vie au sens large, notamment par la prise en compte d'autres dimensions de cette qualité de vie. Cependant, le HAQ était plus sensible au changement au cours du temps (**Arthritis and Rheumatism**) et conviendrait mieux aux objectifs d'une étude longitudinale.

---

[1]Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. **Arthritis Rheum** 1980;23:137-145.

POLYARTHRITE RHUMATOIDE RECENTE

La PR récente désigne la période de début de la maladie et varie selon les auteurs entre 2 et 5 ans après l'apparition des premiers symptômes. L'étude de cette période de la maladie est cruciale, car les mécanismes psychologiques d'ajustement du sujet à sa nouvelle situation de malade chronique s'y mettent en place. Pour identifier la spécificité du phénomène d'incapacité fonctionnelle à ce stade, l'étude comparative de l'incapacité à cette période avec une période plus tardive où la maladie est établie a été réalisée à travers 2 échantillons d'une durée moyenne respective de 2 et de 8 ans. Ceci a permis de montrer qu'elle évoluait selon 2 modèles différents à ces stades différents a partir d'un modèle avec interaction (**Journal of Rheumatology**): au stade précoce (moins de 5 ans) le HAQ est lié à l'indice de Ritchie, la vitesse de sédimentation et aux lésions articulaires radiographiques, tandis qu'il est lié à la durée de la maladie, la présence de lésions extra-articulaires et aux lésions articulaires radiographiques lorsque la maladie est plus ancienne.

DETERMINANTS DE L'INCAPACITE : LA COHORTE EURIDISS

L'étude EURIDISS (European Research on Incapacitating Diseases and Social Support) a constitué une cohorte européenne rassemblée par 10 universités de 9 pays dont 2 en France (Reims, Nancy I). L'objectif est d'étudier le rôle du soutien social sur la qualité de vie et le développement du handicap dans la PR au stade récent (moins de 5 ans) (**International Journal of Health Sciences**). Le Department of Health Science de l'université de Groningen (Pays-Bas) et le Département de Santé Publique de l'université de Nancy I (France) ont élaboré et mis en oeuvre cette étude internationale. Le Département de Santé Publique de Nancy réalise l'étude en Lorraine, coordonne l'étude française qui comprend également la Champagne-Ardennes et assure la centralisation et le contrôle de qualité des données internationales (**Clinical Rheumatology**).

Le HAQ, instrument de mesure de l'incapacité fonctionnelle adapté et validé, a été utilise pour en étudier les déterminants dans la cohorte EURIDISS.
Dans l'analyse transversale des données initiales de 2 pays (Pays-Bas, France), il apparaît que cette incapacité est marquée dès le début de la maladie, avec un score moyen de 1 sur l'échelle HAQ qui va de 0 à 3, et dépend essentiellement de la durée et de caractéristiques d'activité de la maladie (indice de Ritchie et vitesse de sédimentation) (**Journal of Rheumatology**). Les variables démographiques telles que l'age, le sexe ou le niveau d'éducation ne sont pas significativement liées au niveau

d'incapacité. Une deuxième approche a montré le lien qui existe avec les événements de vie stressants, ainsi qu'avec le réseau de soutien social et les membres qui le composent. Cette analyse fournit également des arguments en faveur d'un effet-tampon du réseau social sur le stress, déjà mis en évidence dans plusieurs autres pathologies (**European Journal of Public Health** (soumis)).

Dans l'analyse longitudinale possible grâce aux données françaises après un an de suivi, l'évolution de cette incapacité est modérée (+ 5 % en un an) et de même ampleur quelle que soit l'ancienneté de la maladie (1 à 5 ans). Plusieurs méthodes d'analyse de ces données longitudinales ont été utilisées pour cette approche préliminaire: régression linéaire simple et conditionnellé au niveau d'incapacité initiale; utilisation également simple et conditionnelle de mesures résumées des changements au cours du temps, analyse de variance avec mesures répétées. Cette dernière méthode, qui apparaît la plus pertinente en raison de l'importante corrélation intra-sujet des variables dépendantes et de certaines variables indépendantes, indique que les déterminants de cette incapacité sont surtout la durée de la maladie, l'indice de Ritchie et la vitesse de sédimentation entre sujet, tandis qu'il n'y a pas de changement significatif du niveau de score HAQ intra-sujets (manuscrit en préparation).

PERSPECTIVES

En résumé, le développement d'une méthodologie d'adaptation culturelle a permis de disposer d'un instrument de mesure de l'incapacité fonctionnelle valide et reproductible, équivalent à l'instrument d'origine déjà largement utilisé dans les études publiées dans la littérature.
L'étude de la cohorte EURIDISS se poursuit en Europe. Elle ouvre d'intéressantes perspectives méthodologiques et offre une meilleure connaissance de la maladie à ce stade précoce, en raison du nombre important de patients rassemblés (plus de 700 sujets inclus). L'application des méthodes d'analyse longitudinale permettra d'affiner l'identification des facteurs prédictifs de la qualité de vie et de l'incapacité fonctionnelle en particulier. Ces déterminants sont de nature biomédicale, mais pourront également être recherchés dans le domaine psychosocial, ainsi que le suggèrent les premières données transversales.
A travers les modes de prise en charge, c'est aussi une description et une comparaison des systèmes de santé européens vis-à-vis de cette maladie chronique qui ont déjà été initiées et seront poursuivies (**Revue d'Epidémiologie et de Santé Publique**).

La mesure de l'incapacité à travers la perception du sujet présente incontestablement des avantages, malgré les difficultés méthodologiques rencontrées pour mettre au point les instruments de mesure adéquats. Ils permettent au médecin qui se familiarise à leur usage une meilleure connaissance du vécu et donc des difficultés de son patient dans sa vie quotidienne, dont l'observation ne rend compte qu'imparfaitement.

Il faut signaler que ces déterminants sont pour partie les mêmes que ceux qui permettent de prédire d'autres paramètres de gravité dont l'usage est plus familier au rhumatologue (destructions articulaires radiographiques, par exemple).

Cette progression dans la connaissance des facteurs de l'incapacité est un outil précieux pour le clinicien et pour les autres acteurs de santé publique, qui peuvent ajuster la prise en charge individuelle de chaque patient en fonction de ses atteintes cliniques et de leur retentissement psychologique, étendre la prise en compte du sujet à son environnement et à son réseau social formel et informel, ou agir à l'échelle de la collectivité pour améliorer les composantes de cet environnement et de certains éléments de ce réseau.

BIBLIOGRAPHIE

F Guillemin, C Bombardier, D Beaton. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. **J Clin Epidemiol** 1993 (Sous presse)

F Guillemin, S Briançon, J Pourel. Mesure de la capacité fonctionnelle dans la polyarthrite rhumatoïde : Adaptation française du Health Assessment Questionnaire. **Rev Rhum Osteoartic** 1991; 58: 459-465.

F Guillemin, S Briançon, J Pourel. Validity and discriminant ability of the Health Assessment Questionnaire in early rheumatoid arthritis. **Disability and Rehabilitation** 1992; 14:71-77.

F Guillemin, D Bucquet, S Briancon, B Cadet, J Pourel. Generic or specific health status measurement: relationship with clinical variables in rheumatoid arthritis. **Arthritis Rheum** 1992; 35 (Suppl 5): 6N-R38

F Guillemin, S Briançon, J Pourel. Disability in rheumatoid arthritis: two models in early and established disease. **J Rheumatol** 1992; 19: 366-369.

EURIDISS. European Research on Incapacitating Diseases and Social Support. By: S Briançon, D Doeglas, F Guillemin, B Krol, R Sanderman, T Suurmeijer, W van den Heuvel. **Int J Health Sciences** 1990; 1: 217-228.

F Guillemin. Impact of social support and care on prognosis of the early rheumatoid arthritis: hypothesis and preliminary results of the EURIDISS project. **Clin Rheumatol** 1990; 9: 27-28.

F Guillemin, T Suurmeijer, B Krol, C Bombardier, S Briançon, D Doeglas, R Sanderman, W van den Heuvel. Functional disability in early rheumatoid arthritis: description and risk factors. **Journal of Rheumatology** (Sous presse)

F Guillemin, B Krol, S Briancon, R Sanderman, T Suurmejeir, D Doeglas, W van den Heuvel. Stressful life events and disability in early rheumatoid arthritis. **Eur J Public Health** (Soumis)

S Briançon, F Guillemin, C Bouchet et le groupe EURIDISS. (A Bjelle, D Doeglas, LM Smedstad, B Krol, T Kvien, R Sanderman, T Suurmeijer, W van den Heuvel). Handicap, qualité de vie et consommation de soins dans la polyarthrite rhuamtoïde: une cohorte européenne. **Rev Epidém Santé Publ** 1993;41(Suppl 1):S82.

TRAVAUX ET PUBLICATIONS

J Clin Epidemiol Vol. 46, No. 12, pp. 1417–1432, 1993

0895-4356/93 $6.00 + 0.00

CROSS-CULTURAL ADAPTATION OF HEALTH-RELATED QUALITY OF LIFE MEASURES: LITERATURE REVIEW AND PROPOSED GUIDELINES

FRANCIS GUILLEMIN,[1]* CLAIRE BOMBARDIER[2,3] and DORCAS BEATON[3]

[1]School of Public Health, University of Nancy, France, [2]Clinical Epidemiology Division, Wellesley Hospital, University of Toronto and [3]Ontario Worker Compensation Institute, Toronto, Canada

(Received in revised form 11 June 1993)

Abstract—Clinicians and researchers without a suitable health-related quality of life (HRQOL) measure in their own language have two choices: (1) to develop a new measure, or (2) to modify a measure previously validated in another language, known as a cross-cultural adaptation process. We propose a set of standardized guidelines for this process based on previous research in psychology and sociology and on published methodological frameworks. These guidelines include recommendations for obtaining semantic, idiomatic, experiential and conceptual equivalence in translation by using back-translation techniques and committee review, pre-testing techniques and re-examining the weights of scores. We applied these guidelines to 17 cross-cultural adaptation of HRQOL measures identified through a comprehensive literature review. The reporting standards varied across studies but agreement between raters in their ratings of the studies was substantial to almost perfect (weighted $\kappa = 0.66$–0.93) suggesting that the guidelines are easy to apply. Further research is necessary in order to delineate essential versus optional steps in the adaptation process.

Quality of life Health status index Cross-cultural comparison Culture Validity Guidelines

RATIONALE

A large body of research has recently been devoted to the development of health-related quality of life (HRQOL) measures. In 1991 alone, over 160 different measures were used in the published literature [1]. Such techniques are increasingly used in clinical trials [2,3] to determine the impact of medical intervention on quality of life (QOL), and by public health researchers [4] to assess the outcome of health care services. With a few exceptions [5,6] all the measures so far developed are in the English language and are intended for use in English-speaking countries. There is nonetheless a need for measures specifically designed to be used in non English-speaking countries and also among immigrant populations, since cultural groups vary in disease expression and in their use of various health care systems. This need has become more acute with the growing number of large multicenter multicountry trials.

In order to meet that need, two options are available: 1) to develop a new measure, and 2) to use a measure previously developed in another language. The first option, the generation of a new HRQOL measure is a time-consuming process in which the bulk of the effort is devoted to the conceptualization of the measure and the selection and reduction of its items. In the second option, if the transposition of a measure from its original cultural context is done by simple translation it is unlikely to be successful because of language and cultural differences [7]. Furthermore, the perception of QOL and the ways in which health problems are expressed vary from culture to culture [8]. To be successfull this option requires a systematic approach to the translation and to the cross-cultural adaptation process of HRQOL measures. A recent effort was made by Hunt and the European Group for Health Measurement and Quality of Life Assessment through a cross-cultural adaptation of the Nottingham Health Profile (NHP) to several European countries using a systematic method [9]. This effort of standardisation needs to be expanded.
In this paper, we propose a set of standardized guidelines for the cross-cultural adaptation of HRQOL measures based on previous research in psychology and sociology [10-18] and on published methodological frameworks for HRQOL validity [19,20]. We review the published literature on the cross-cultural adaptation of HRQOL instruments and evaluate the practicality of our proposed guidelines to this literature.

METHODS

Literature search strategy

Relevant papers reflecting the methods used for cross-cultural adaptation were identified from three databases: *Medline* (1966-1992), *Health Planning and Administration* (1975-1992) and *Embase* (Excerpta Medica) (1990-1992). The search strategy used was to identify articles with 'quality of life', 'health status', 'health status indicator', 'functional status', 'questionnaires' and 'interviews' as main subject headings (exploded) or text words in titles and abstracts. This was matched with 'cross-cultural', 'cross-cultural comparison', 'translation' and 'languages' as main subject headings (exploded) or text words. Papers published between January 1966 and October 1992 were considered, without language restriction. All references obtained were entered into Reference Manager computer software [21] to check for duplication.

Development of the guidelines

The literature review identified several publications in the field of psychology and sociology addressing the methodology of cross-cultural adaptation. For example, the General Health Questionnaire has been translated into at least 36 languages [10] and the State Trait Anxiety Inventory into at least 21 languages [11]. Although these papers were excluded from our formal evaluation because they addressed mainly mental health, their work suggests methodological approaches, developed to overcome the inadequacy of simple translation, which may be useful in the cross-cultural adaptation of HRQOL measures. We have developed guidelines and a scoring method which can be applied in a standardized manner to evaluate the quality of cross-cultural adaptations of HRQOL measures. This system was based on both empirical and theoretical findings extracted from the literature. The empirical basis was derived from a systematic review of the published work on cross-cultural adaptation. Theoretical foundations were gained from guidelines on the methodology of assessing the validity of HRQOL measures [19,20].

Selection of articles for review

Our literature search identified 712 references (Table 1). Their titles and abstracts were reviewed by one author (FG) for relevance to the study. Papers were included if they contained a description of the methodological process used to adapt a

HRQOL measure from a source to a target culture. Papers were excluded if they presented only results of cross-cultural comparisons, or simply mentionned the use of HRQOL measures in different countries or in international trials without describing the translation and adaptation process. Papers concerning instruments to measure only pain, symptoms or mental status/disorders as well as utilities were also excluded.

During the selection process, any reference to methodology quoted in these papers were used to conduct a supplementary search in the *Science Citation Index* (1980-1992) for additional material on cross-cultural adaptation of HRQOL instruments.

From 1966 to 1992, only 32 papers met the inclusion and exclusion criteria for review. Six of these publications were in abstracts form [22-27] and were not included in the present review since not enough information was available to asses the quality of the adaptation process. Seven other papers [28-34] were rejected during the review process as they were judged by subsequent raters (CB,DB) not to have met the original selection criteria, i.e. either not deqling with a quality of life measure or not with the cross-cultural adaptation of such measure or not containing a description of the method used for cross-cultural adaptation. Of the remaining 19 articles, only 2 were excluded [35,36] by restricting our review to English language papers. Thus, the review of seventeen studies was completed [37-53].

Application of the guidelines to assess quality of studies

The proposed guidelines include 5 sections: 1) translations and 2) back-translations by qualified people, 3) committee review of those translations and back-translations, 4) pre-testing for equivalence using adequate techniques (with bilingual or monolingual individuals), and 5) reexamination of the weighting of scores, if relevant. If an instrument was adapted from one culture to another using a similar language (from American English, USA, to British English, UK), the steps of translation and back-translation obviously were not required and therefore were not assessed.

The quality of each study was assessed by two of the investigators (DB,CB) blinded to authors'names, journal titles, and city or area of the study. Each investigator was provided with an operational definition of the evaluation criteria for each section and was asked to rate each section as "good", "moderate", "poor" quality or "not done" using a standard data extraction form (Appendix). Agreement between the two judges in these ratings was assessed using the weighted kappa statistic for categorical judgement [54]. For each section, a mean score across studies was calculated as the mean of quality ratings assigned for that section with the following

values: good=3; moderate=2; poor=1; if the section was rated as "not done" the study was not included in this calculation.
An overall score was also calculated for each study by adding the scores across all sections for that study and dividing by the number of sections. For this calculation a value of 0 was assigned to sections rated as "not done". If translation and rescoring were not considered relevant for a particular study, these sections were omitted from the calculation of the overall score for that study. Agreement between judges on these overall scores was assessed using the intra-class correlation coefficient for continuous data [55]. Recognizing that the concepts and techniques involved in adapting measures for people using similar language in another culture, using another language in another country and using another language in the same country (immigrants) differ, agreement between the judges was also considered separately for each subgroup of studies.

RESULTS

Settings for cross-cultural adaptation.

Our literature review identified several settings for the cross-cultural adaptation process. A range of situations may be encountered depending on similarities and differences between the cultures and languages of the populations concerned. An instrument originally developed in the English language in the USA can readily be used by a majority of the American population (Table 2: example 1).
Immigrants using the same language may encounter particular problems in expressing themselves with regard to health and HRQOL. Therefore, particular attention should be paid to the adaptation of cross-cultural HRQOL measures to such populations. Immigrants to the USA, for instance Hispanics, will judge their health and related QOL according to their cultural origin and their degree of assimilation into the host culture. They may have been settled for long enough to have mastered the English language sufficiently well to answer the original instrument, and still refer to their Spanish culture in assessing their situation (Table 2: example 2).
An instrument used in a country other than that in which it was developed may require adaptation if the populations concerned have another culture with similar language. For instance British English should be used in Great Britain rather than American English. There are sufficiently meaningful differences between the British

and American cultures to necessitate modification of some items and validation of the measure in its new setting [37-39](Table 2: example 3).
Recently settled immigrants with a low degree of acculturation may require an instrument that is cross-culturally adapted to their Spanish (native) language and culture, but appropriate to the American situation (Table 2: example 4).
Under most circumstances, instruments require adaptation for use in a different country with both a different culture and a different language. For instance, the American measure would need to be modified for use in the French language in France or in Canada (Table 2: example 5). The degree of adaptation required depends on similarities in language structure (there are fewer differences between most of the European languages than there are between European and Arabic or Asian languages) and in culture [12].

Cross-cultural adaptation has two components: the translation of the HRQOL measure and its adaptation, i.e. a combination of the literal translation of individual words and sentences from one language to another and an adaptation with regard to idiom, and to cultural context and life-style. Translation and adaptation are required for examples 4 and 5 while only cross-cultural adaptation is necessary in examples 2 and 3. The quality of the adapted measure is then assessed with regard to its sensibility. The elements of sensibility, as defined by Feinstein, which need to be considered include the purpose of the measure, its comprehensibility, its content and face validity, its replicability and the suitability of the scales [19].

Guidelines for cross-cultural adaptation

The following guidelines concentrate on the points that must be addressed in order to preserve the sensibility of the tool in the target culture. Table 3 summarizes the steps that are essential in order to ensure the quality of the procedure.

1.Translation

Produce several translations

Translations are of higher quality when undertaken by at least two independent translators. This allows for the detection of errors and divergent interpretations of ambiguous items in the original. The quality will be even higher if each translation is undertaken by teams rather than single individuals, who are more likely to introduce personal idiosyncrasies.

Use qualified translators

The qualifications and characteristics of the translators are also important. Highly educated individuals may not be culturally representative of the target population [13]. Translators should preferably translate into their mother tongue [40]. Some of them should be aware of the objectives underlying the material to be translated and the concepts involved so as to offer a more reliable restitution of the intended measurement [56,41]. Other translators who are unaware of these objectives and concepts may usefully elicit unexpected meanings from the original tool.

2. *Back-translation*

Produce as many back-translations as translations

Back-translation, translating back from the final language into the source language, has been shown to help improve the quality of the final version [7,14]. Each first translations should be back-translated independently from each other. Misunderstandings in the first translation may be amplified in the back-translation, and thereby revealed. Failure to adapt to the cultural target context and ambiguity in the source version can also be uncovered.

Use appropriate back-translators

Back-translation is of better quality if those who do it are fluent in the idioms and colloquial forms of the source language, i.e. the result of their back-translation. Thus, they should also translate into their mother tongue. Unlike some of the first translators, back-translators should preferably not be aware of the intent and concepts underlying the material. Back-translators without a priori knowledge of the intent of the original instrument are free of biases and expectations and their back-translation may reveal unexpected meanings or interpretations in the final version.

3. *Committee review*

Constitute a committee to compare source and final versions

A committee should be constituted in order to produce a final version of the modified measure based on the various translations and back-translations obtained as described above. Part of that committee's role should also review the introduction and instruction to the questionnaire as well as review the scaling of responses to each question (i.e. the translation should maintain equivalence of steps in Likert-type scales).

Membership in the committee should be multidisciplinary

To use the analogy of the development of a new health status measure, the committee should consist of individuals expert in the disease(s) explored, and in the

intent of the measure and the concepts to be explored. Bilingual members are of particular value for such committee [15].

In case of a cross-cultural adaptation for an immigrant population, individuals representative of the target group are likely to be available. Their input is likely to result in a measure better adapted in terms of idioms and colloquialisms than that which would be produced by highly educated people [9,42]. A scale referring to the ability to speak, write, read and understand both languages has been developed [16] and can be useful in selecting these bilingual committee members.

Use structured techniques to resolve discrepancies

The committee may resolve problems by considering the material it has now collected. It may further decide to repeat the translation-back-translation process. A decentering technique [17] has also been proposed as a way of improving cross-cultural adaptation. This technique considers the source and final versions equally important. Both are open to modification during the translation procedure. In other words, the measure is not considered to centre on one of the languages. Decentering is best conducted in close collaboration with the authors. If problem items are found, the authors may provide a working version of the instrument or items, maintaining the concept of the questions, but avoiding colloquialisms. Searching for a common way to express a concept in both languages is the best way to ensure that the final version maintains content validity. It is unusual for authors to be available, and this process may need to be conducted by committee members.

Modify instructions or format, modify or reject inappropriate items, generate new items

The committee must ensure that the introduction to the research tool and the instructions for filling in the questionnaire are carefully translated in order to preserve the replicability of the measure [19]. The redundancy principle, i.e. repeating the same instruction in a different manner, may help to reduce comprehension errors [12].

The review committee is also likely to modify or eliminate irrelevant, inadequate and ambiguous items and may generate substitutes better fitting the cultural target situation while maintaining the general concept of the deleted items.

Ensure that the translation is fully comprehensible

Guidelines about how to produce translations comprehensible to a majority of people have suggested using language which can be understood by 10 to 12-year old children [12]. Recommendations include: short sentences with key words in each item as simple as possible [11]; the active rather than the passive voice; repeated nouns instead of pronouns; and specific rather than general terms. Authors should avoid using: metaphors and colloquialisms; the subjective mode; adverbs and

prepositions telling 'where' and 'when'; possessive forms; words indicating vagueness; and sentences containing two different verbs that suggest different actions.

Verify cross-cultural equivalence of source and final versions

Several taxonomies of cross-cultural equivalence have been proposed in the psychiatry literature [12,13,60,61]. The ultimate parity is equivalence of HRQOL concepts within the cultures concerned. Translators aiming for conceptual equivalence should consider the following:

- *semantic equivalence* is equivalence in the meaning of words, and achieving it may present problems with vocabulary and grammar. For example, vocabulary problems may be encountered in the question 'are you able to bend ?' which can refer to several parts of the body, such as the arm, back or knees, and might have been intended - and translated - only to explore the ability to flex (arm), bend over (back) or squat (knees). Furthermore, some words, such as 'happy', have several subtly different meanings depending on the context.

Grammatical alterations are sometimes necessary in the construction of sentences. For example, languages without the gerund form may be more difficult to adapt [13]: activities couched in terms such as dancing, singing or eating (gerund form of to dance, sing and eat) may not be readily translatable.

- *idiomatic equivalence.* Since idioms and colloquialisms are rarely translatable, equivalent expressions have to be found or items have to be substituted. This is more likely to be necessary in the emotional and social dimensions. For example, 'Do you feel downhearted and blue ?' or 'Do you feel at home ?' are untranslatable idioms for which equivalents must be found. The item 'I am feeling on edge' in the NHP was translated into 'I have my nerves outside my skin' in Italian [9], 'I feel nervous, tense' in French [43] and 'I am afraid' in Arabic [60].
- *experiential equivalence.* The situations evoked or depicted in the source version should fit the target cultural context. This may result in the modification of an item. For example, in the Brazilian version of the HAQ, 'using public transportations' was substituted for 'using a private car', since most of the people in Brazil have no car [44]. 'I have forgotten what it is like to enjoy myself' [60] and 'How many hours a week do you have leisure activities?' do not refer to usual experiences in a number of cultures and equivalent feelings (enjoy) or activities (leisure) must be found or the items discarded.
- *conceptual equivalence* refers to the validity of the concept explored and the events experienced by people in the target culture, since items might be equivalent in semantic meaning but not conceptually equivalent.

For example, 'cousin' and 'brother' may mean more than simply second or first-degree relative of the same generation. In many cultures in developing countries they have a wider meaning within the social network. 'I have pain in my head' may translate perfectly into another language semantically, but have a totally different conceptual meaning for the target culture [60].

4. Pre-testing

Check for equivalence in source and final versions using a pre-test technique

In pre-testing, a sample population replies to the questionnaire in order to check for errors and deviations in the translation. Two techniques are available: a probe technique and appraisal by bilingual individuals. Both allow for the checking of face validity, i.e. the confirmation that questions are acceptable without arousing reluctance or hesitation. If the final version does not achieve a satisfactory level of equivalence, further revision can be performed by the committee.

Either use a probe technique

The answer to an item might appear adequate, yet be consistently misunderstood. In order to determine whether a questionnaire is being understood correctly, it can be administered to a group of patients as follows. After each answer (or a random sample of answers), the patient is asked the probe question: "What do you mean ?" and is encouraged to elucidate his or her understanding of the item in an open-ended manner [18]. This ensures that the final item is understood as having a meaning equivalent to that of the source item.

Or submit the source and final versions to bilingual lay people

The source and the final versions of a measure can be administered to a group of bilingual individuals in order to detect possible discrepancies. This method can also help pinpoint any inadequacy of the final version with the cultural context. They are asked to rate the equivalence of each item between the source and final versions. Those items with low level of equivalence or rated discrepantly by different people can still be revised at this stage [11,12]. Administration of a questionnaire to bilingual lay people is not practical in every setting but may be possible with immigrants.

The case of adaptation for immigrants requires two additional considerations:

Choose the language of measurement administration for immigrants

During questionnaire administration, some immigrant respondents express a preference for their native language. Several methods have been described for the choice of language of administration; it can be the decision of the respondent her/himself, or of the interviewer or research assistant, or it can be based on a

measure of acculturation. Several such measures have been proposed (57-59) and include language proficiency and preference, country of birth and origin, location of education, ethnic identification, contact with homeland, and ethnicity of children's friends, combined to form an acculturation score. Even if not used to determine the interview language, the score can be a useful covariable when investigating several cultural groups.

Use a dual-format measure for immigrants

On the other hand, immigrants may switch language during an interview and responses should be recorded in the language used (on a two-language form). With regard to self-administered questionnaires, the best option is to present the material in a dual-language format, either on two separate pages [62] or item by item [40]. The measure of HRQOL by proxies is not recommended, even when respondents are illiterate, since individual subjective appreciation is not reliably assessed by other raters [63]. In this situation, interview must be preferred to self-administered questionnaire. Interview may also be appropriate when questionnaires are not appropriate to the respondent's culture [9].

5.Weighting scores

Consider adapting the weights of scores to the cultural context

A scoring method using weights is provided with the source versions of some instruments (Sickness Impact Profile [64], NHP [65]) in order to combine the information in an index or in several indices (profile). However, the weighting may not apply to the new cultural situation. It can be reexamined either by judgement or using a mathematical approach. Using judgement, the cross-cultural validity of the weighting of items is reexamined by experts, who may be health care professionals, patients or lay people. Several techniques are available to elicit culture-adapted weights from expert opinion. With a mathematical approach, data obtained from a sample of patients are analyzed by various statistical techniques for scalability (Gutmann analysis) or dimensionality (factor analysis) in order to work out the best way of aggregating the information in one index or several indices.

Application of the guidelines

Study characteristics.

A description of the content of the 17 studies analyzed is given in Table 4. Some instruments have been adapted to another language by different research teams addressing different aspects of the adaptation. For instance, a Chicano version of the

Sickness Impact Profile (SIP, originated in English [64]) produced by one team [50] has been further examined for score weights by another team [42].

The methods for adaptation reported in the 17 studies were heterogenous. In the adaptation of the instruments from one culture in another with the same language (three studies), i.e. from American to British English, rewording of items was never carried out by a specified committee. Pretesting of the measure was conducted in two studies using either a probe technique [37] or a comparison of the original and final versions [38]. The weight of scores were reassessed in one study [39].

In the adaptation of instruments from one culture to another which uses a different language (8 studies), the translation techniques used varied from one literal translation to three translations performed independently by one to two translators with varying degrees of qualification. Back-translation was described in six studies. Only one study with multiple translations specified using as many back-translations as translations [41]. When a committee for review was constituted (5 studies), its composition varied from two authors to 12 people including physicians, other health professionals, patients, non-patients and bilingual people. Pretesting of the measure was conducted in 4 studies. The weighting of scores, examined in two of three studies where it was applicable [43,47] involved the use of large samples of patients and non-patients as experts. Weights were derived by expert elicitation using the Thurstone method, as in the original instrument.

Six papers addressed important cultural groups of immigrants in the host country: Asians in the UK and Hispanics in the USA. The translation techniques used varied considerably, from one to five translations, each involving one to three translators in one to four back-translations. A review committee of two to 15 carefully selected bilingual people was constituted in five studies. A pre-test of the new versions was conducted in three studies by 12 to 31 bilingual individuals. Weighting of score was reassessed in one of three relevant studies (SIP in Spanish [42]) using expert evaluation by a group of health care consumers, as in the original instrument.

Quality of the studies.

For each of the individual guidelines, the mean quality scores on a scale of 1=poor, 2=moderate and 3=good, ranged between 1.9 and 2.4 (Table 5). Thus, when a guideline was addressed in a study, raters judged the methods to be of moderate quality. According to Landis [66], there was substantial to almost perfect agreement between the two judges in these ratings (weighted Kappa 0.66 to 0.93).

Since not all of the relevant guidelines were considered in each study, the overall scores were much lower. In calculating the overall score when a guideline rated as "not done" is assigned a value of 0, the mean quality score overall studies was 1.3. It

was higher in immigrant populations (1.6) than in adaptation to another country using the same or another language (0.8 and 1.3 respectively). The agreement between judges for the overall quality score for the 17 studies was high (ICC=0.92). It was higher in studies where the target was another country using the same language (ICC=1) or adaptation to another language in another country (ICC=0.96) than in studies for adaptation to immigrant populations (ICC=0.87).

COMMENTS

Cross-cultural adaptation must be clearly distinguished from cross-cultural comparison since the two processes rely on different research hypotheses [67]. Adaptation is oriented towards measuring a similar phenomenon in different cultures; it is essentially the production of an equivalent instrument adapted to another culture. Cross-cultural comparison refers to the comparative study of a phenomenon across cultures in order to identify differences attributable to culture. It is possible only after the measurement tool has been adapted and is equivalent in both cultures. Thus, the cross-cultural adaptation of a measure is a prerequisite for the investigation of cross-cultural differences.

The articles reviewed were mainly describing the cross-cultural adaptation of English-language instruments into European languages in European countries. Although our search strategy included all languages, 17 of the 19 eligible articles were papers published in the English language. It may be that the use of self-administered instruments to measure QOL is an English cultural phenomena. Alternatively, non English papers may have been missed since papers published only in national journals may not be included in the three medical databases we searched. These papers would have been overlooked unless they were cited in another article included in a database [35]. Finally, much of this research may not be published because it is not the main topic of the research but only a preliminary step toward an application of a quality of life measure. However, the increasing number of publications appearing in recent years reflects the growing importance and interest attached to the methodology of cross-cultural adaptation.

Our review of the literature indicates a lack of standardized approach to the cross-cultural adaptation of HRQOL instruments.. The methodologies vary and often the authors do not give the readers essential information to understand the strength of the translation. Interestingly, citation searches (*Science Citation Index*) using

methodology papers found in the psychology and sociology literature for adaptation of tools in these fields failed to produce additional references relating to HRQOL tools. This suggests that many researchers in QOL may not be aware or do not quote this methodological work developed in the psychology and sociology literature. Based on our review of the methods of cross-cultural adaptation in the field of psychology and sociology, and our review of the HRQOL measures, we propose a set of guidelines which includes 5 essential steps for the translation and cross-cultural adaptation of HRQOL measures. The agreement between the judges using the proposed guidelines appeared substantial to almost perfect. The results were consistent within and across the sections of these guidelines and in different cross-cultural adaptation settings. This indicates that they are appropriate, easy to interpret and suggests that they can be used further with satisfactory reliability.

The quality of the methodology employed for the adaptation of HRQOL measures was rated between 1.9 and 2.4 (on a 1-3 scale) in each section. However, in each study assessed, some aspect of the adaptation process is likely to have been underreported (even if it was carried out) perhaps because it was not at the time regarded as important. Hence, a score of 0 ('not done') might have underrated a work that authors merely did not made explicit. For this reason we are not reporting on individual paper scores. For instance, the replicability of the adapted instrument, i.e. the clarity of the presentation and the thoroughness of the directions provided for its use [19] was addressed in none of the papers, although it might well have been considered by authors.

Whether reliability, validity and sensitivity to change should also be considered in the cross-cultural adaptation process is a matter of controversy. Overall the 17 studies, the reliability of the final version was assessed in 6 studies, the construct validity in 9 studies and the responsiveness in 1 study. On one hand, one may think that the full achievement of cross-cultural equivalence convey the equivalence of the original measurement properties. But on the other hand, one may argue that because of the adaptation process, the modified instrument has unknown reliability, validity or sensitivity to change in the new culture. Since this question is not clear, we did not address this methodology.

Overall, the guidelines used here cannot yet be taken as firm recommendations. Very little research has been done in this field to delineate what is essential from what is supplementary in the process of cross-cultural adaptation. It is important to stress that our guidelines address only the quality of the "process" of adaptation. The

quality of the final product (the adapted version) can only be judged by a qualified committee. A single investigator or clinician can only judge the process used in the translation and cross-cultural adaptation since a single individual cannot fulfil the linguistic and other committee qualifications. Our literature review suggests that each step adds quality to the final version in terms of equivalence of concepts explored between source and final instruments, but this benefit has to be weighed against the feasibility of the process.

Important factors in determining the feasibility of a particular technique are the constraints of time and resources. Bilingual people are often in short supply. Variations in how bilingalism is defined may mean they are even more rare[1] . This problem can be more difficult to deal with when an original instrument is developed in a language other than English. In some such situations, for example in international studies involving several countries in Europe, investigators may opt to translate the instrument into English first, as a common communication language, and then adapt it into the other language as required. If a shortage of bilingual people impedes the pre-testing with bilingual respondents, a probe technique with monolingual respondents can be used.

The preservation of the sensibility of an instrument is time-consuming but unavoidable. The only step that can possibly be shortened is the examination of the weighting of scores, if any, by simply accepting the weights of the original instrument. It should be borne in mind, however, that the validity of the final score may be diminished [43].

Further research is required to establish a method to quantify the equivalence of source and final instruments across cultures, and to identify essential versus optional steps in the adaptation process. Research is also needed to determine adaptation needs for HRQOL measures where items are selected by the patient (Patient Elicitation Technique [68], SEIQoL [69]) rather that the questionnaires addressed in this paper.

In conclusion, the adaptation of a preexisting measure to the cultural context of a target population, as described above, has several advantages:

- it provides a common measure for the investigation of HRQOL within different cultural contexts;

1. Some authors have considered as bilingual people (with one language) who have lived at least one year in another country (with another language) [42]. Other define people as bilingual only when they have been reared in two cultural and language contexts, keeping in touch with both [40]. Although the former definition probably does not allow for an understanding and mastery of idioms and colloquialisms, the latter situation may well be too rare to be useful.

- it offers a standard measure for use in international studies, many of which are now being conducted;
- it allows comparisons between national/cultural groups relying on a standard measure designed and adapted to measure the phenomenon cross-culturally;
- it allows the inclusion of immigrants avoiding the frequent bias of representing only the dominant culture of the country;
- it is less costly and time-consuming than generating a new measure. Nevertheless, it should be borne in mind that the cross-cultural adaptation of HRQOL also requires careful attention, involves numerous people and is time-consuming.

SUPPORT. F G was supported by a Bourse Lavoisier from the French Ministry of Foreign Affairs and by the Fondation pour la Recherche Therapeutique, France.

ACKNOWLEDGEMENTS. The authors are grateful to Dr Stephen Stansfeld and Dr James Wright for fruitful comments on earlier drafts of this manuscript.

REFERENCES

1. Spilker B, Simpson RL, Tilson HH. Quality of life bibliography and indexes: 1991 update. J Clin Res Pharmacoepidemiol 1992; 6: 205-66.

2. Aaronson NK, Meyerowitz BE, Bard M et al. Quality of life research in Oncology. Past achievements and future priorities. Cancer 1991; 67: 839-843.

3. Fitzpatrick R, Fletcher A, Gore S, Jones D, Spiegelhalter D, Cox D. Quality of life measures in health care. I: Applications and issues in assessment. BMJ 1992; 305: 1074-1077.

4. Advances in Health Status Assessment: Proceedings of a conference. Med Care 1992; 30: MS1-MS293.

5. Gerin P, Dazord A, Cialdella P, Leizorovicz A, Boissel JP. Le questionnaire "Profil de la qualité de la vie subjective". Premiers éléments de validation. Therapie 1991; 46: 131-138.

6. Kempen GI, Suurmeijer TP. The development of a hierarchical polychotomous ADL-IADL scale for noninstitutionalized elders. Gerontologist 1990; 30: 497-502.

7. Berkanovic E. The effect of inadequate language translation on Hispanics' responses to health surveys. Am J Pub Health 1980; 70: 1273-1276.

8. Kleinman A, Eisenberg L, Good B. Culture, illness and care: clinical lessons from anthropologic and cross-cultural research. Ann Intern Med 1978; 88: 251-258.

9. Hunt SM, Alonso J, Bucquet D, Niero M, Wiklund I, McKenna S. Cross-cultural adaptation of health measures. Health Policy 1991; 19: 33-44.

10. Goldberg D, Williams P. A user's guide to the General Health Questionnaire. Windsor: NFER-NELSON; 1990.

11. Spielberger CD, Sharma S. Cross-cultural measurement of anxiety. In: Spielberger CD, Diaz-Guerrero R, Eds. Cross-cultural anxiety. Washington: Hemisphere Pub Corp; 1976: 13-25.

12. Brislin RW, Lonner WJ, Thorndike RM. Questionnaire wording and translation. In: Cross-cultural research methods. New-York: John Wiley & Sons; 1973: 32-58.

13. Secherst L, Fay TL, Zaidi SMH. Problems of translation in cross-cultural research. J Cross-cultural Psychology 1972; 3: 41-56.

14. Brislin RW. Back-translation for cross-cultural research. J Cross-cultural Psychol 1970; 1: 185-216.

15. Deutscher I. Asking questions cross-culturally: some problems of linguistic comparability. In: Comparative research methods. Warwick DP, Oskerson S, Eds. New Jersey: Prentice Hall; 1973: 163-186.

16. Gonzalez-Reigosa F. The anxiety-arousing effect of taboo words in bilinguals. In: Spielberger CD and Diaz-Guerrero R, Eds. Cross-cultural anxiety. Washington: Hemisphere Pub Corp; 1976: 89-105.

17. Werner O, Campbell DT. Translating, working through interpreters, and the problem of decentering. In: Narroll R, Cohen R, Eds. A handbook of method in cultural anthropology. New York: American museum of natural history Press; 1970: 398-420.

18. Schuman H. The random probe: A technique for evaluating the quality of closed questions. American Sociological Review 1966; 31: 218-222.

19. Feinstein AR. The theory and evaluation of sensibility. In: Feinstein AR, Ed. Clinimetrics. New Haven and London: Yale University Press; 1987: 141-166.

20. Bombardier C, Tugwell PX. Methodological consideration in functional assessment. J Rheumatol 1987; 14 (suppl 15): 6-10.

21. Reference Manager, version 4. Carlsbad, California: Research Information System, Inc; 1989.

22. Perez ER, Mackenzie CR, Ryan C. Development of a Spanish version of the Modified Health Assessment Questionnaire. Arthritis Rheum 1990; 33 (suppl 9): S100 (Abstract).

23. Cardiel MH, Abello-Banfi M, Ruiz-Mercado R, Alarcon-Segovia D. Quality of life in rheumatoid arthritis: validation of a Spanish version of the disability index of the Health Assessment Questionnaire. Arthritis Rheum 1991; 34 (suppl 9): S183 (Abstract).

24. Raynauld JP, Singh G, Shiroky JB, Haraoui B, Choquette D, Blier C, Fortin P, Pelletier JP, Gascon J, Goulet JR, Neville C, Fries JF. A French-Canadian version of the Health Assessment Questionnaire. Arthritis Rheum 1992; 35 (suppl 9): S125 (Abstract).

25. Archenholtz B, Bjelle A. Reliability and validity of a Swedish version of the Arthritis Impact Measurement Scales. Scand J Rheumatol 1990; 20: 219 (Abstract).

26. Abello-Banfi M, Cardiel MH, Ruiz-Mercado R, Alarcon-Segovia D. Quality of life in rheumatoid arthritis: validation of a Spanish version of the Arthritis Impact Measurement Scales. Arthritis Rheum 1991; 34 (suppl 9): S183 (Abstract).

27. Esteva M, Gonzalez N, Ruiz M. Reliability and validity of a Spanish version of the Sickness Impact Profile. Arthritis Rheum 1992; 35 (suppl 9): S219 (Abstract).

28. Chang RH, Dodder RA. The modified purpose in life scale: a cross national validity study. Int J Aging and Human Development 1983-84; 18: 207-217.

29. Kaplan RM, Bush JW, Berry CC. The reliability, stability and generalizability of a health status index. Proc Social Statistics Section. American Statistical Association 1978: 704-709.

30. Brazier JE, Harper R, Jones NMB, O'Cathain A, Thomas KJ, Usherwood T, Westlake L. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. BMJ 1992; 305: 160-164.

31. Siegert CEH, Vleming LJ, Vandenbroucke JP, Cats A. Measurement of disability in Dutch rheumatoid arthritis patients. Clin Rheumatol 1984; 3: 305-309.

32. Van der Heidje DMFM, Van Riel PLCM, Van de Putte LBA. Sensitivity of a Dutch Health Assessment Questionnaire in a trial comparing hydroxychloroquine vs sulphasalazine. Scand J Rheumatol 1990; 19: 407-412.

33. Cavalieri F, Salaffi F, Ferracioli GF. Relationship between physical impairment, psychosocial variables and pain in rheumatoid disability. An analysis of their relative impact. Clin Exp Rheumatol 1991; 9: 47-50.

34. Wiklund I, Romanus B, Hunt SM. Self-assessed disability in patients with arthrosis of the hip joint. Reliability of the Swedish version of the Nottingham Health Profile. Int Disabil Studies 1988;10:159-163.

35. Ferracioli GF, Cavalieri F, Salaffi F, Mercadanti M. Compromissione funzionale, invalidita ed handicap nell'artrite reumatoide. Studio prospettico mediante questionari orientari. Reumatismo 1988;40:45-56.

36. Jacobs HM, Luttik A, Touw-Otten FW, de Melker RA. The Sickness Impact Profile (SIP); results of an evaluation study of the Dutch version. Nederlands Tijdschift Voor Geneeskunde 1990; 134: 1950-1954.

37. Hill J, Bird HA, Lawton CW, Wright V. The Arthritis Impact Measurement Scales: An anglicized version to assess the outcome of British patients with rheumatoid arthritis. Br J Rheumatol 1990; 29: 193-196.

38. Kirwan JR, Reeback JS. Stanford Health Assessment Questionnaire modified to assess disability in British patients with RA. Br J Rheumatol 1986; 25: 206-209.

39. Patrick DL, Sittampalam Y, Sommerville SM, Carter WB, Bergner M. A cross-cultural comparison of health status values. Am J Public Health 1985; 75: 1402-1408.

40. Hendricson WD, Russel IJ, Jacobson JM, Rogan H, Bishop GD, Castillo R. Development and initial validation of a dual language English-Spanish format for the Arthritis Impact Measurement Scales. Arthritis Rheum 1989; 32: 1153-1159.

41. Guillemin F, Briançon S, Pourel J. Validity and discriminant ability of a French version of the Health Assessment Questionnaire in early RA. Disability and Rehabilitation 1992; 14: 71-77.

42. Deyo RA. Pitfalls in measuring the health status of Mexican Americans: comparative validity of the English and Spanish Sickness Impact Profile. Am J Public Health 1984; 74: 569-573.

43. Bucquet D, Condon S, Ritchie K. The French version of the Nottingham Health Profile. A comparison of items weights with those of the source version. Soc Sci Med 1990; 30: 829-835.

44. Bosi-Ferraz M, Magalhaes Oliveira L, Araujo PMP, Atra E, Tugwell P. Cross-cultural reliability of the physical ability dimension of the Health Assessment Questionnaire. J Rheumatol 1990; 17: 813-817.

45. Sampalis JS, Pouchot J, Beaudet F, Carette S, Gutkowski A, Harth M, Myhal D, Senecal JL, Yeadon C, Williams JI, Esdaile JM. Arthritis Impact Measurement Scales: reliability of a French version and validity in adult Still's disease. J Rheumatol 1990; 17: 1657-1661.

46. Ekdahl C, Eberhardt K, Andersson ST, Svensson B. Assessing disability in patients with rheumatoid arthritis. Scand J Rheumatol 1988; 17: 263-271.

47. Hunt SM, Wiklund I. Cross-cultural variation in the weighting of health statements: a comparison of English and Swedish valuations. Health Policy 1987; 8: 227-235.

48. Burckhardt CS, Archenholtz B, Bjelle A. Measuring the quality of life of women with rheumatoid arthritis or systemic lupus erythematosus: a Swedish version of the Quality of Life Scale (QOLS). Scand J Rheumatol 1992; 21: 190-195.

49. Alonso J, Anto JM, Moreno C. Spanish version of the Nottingham Health Profile: translation and preliminary validity. Am J Public Health 1990; 80: 704-708.

50. Gilson BS, Erickson D, Chavez CT, Bobbitt RA, Bergner M, Carter W. A Chicano version of the Sickness Impact Profile. Cult Med Psychiatry 1980; 4: 137-150.

51. Coulton CJ, Hyduk CM, Chow JC. An assessment of the Arthritis Impact Measurement Scales in 3 ethnic groups. J Rheumatol 1989; 16: 1110-1115.

52. Hendricson WD, Russel IJ, Prihoda TJ, Jacobson JM, Rogan A, Bishop GD. An approach to developing a valid Spanish language translation of a health status questionnaire. Med Care 1989;27:959-966.

53. Ahmad WIU, Kernohan EEM, Baker MR. Cross-cultural use of socio-medical indicators with British Asians. Health Policy 1989; 13: 95-102.

54. Fleiss JL. The design and analysis of clinical experiments. Wiley, New York 1986.

55. Dunn G. Design and analysis of reliability studies. Oxford University Press, New York 1989.

56. Del Greco L, Walop W, Eastridge L. Questionnaire development. 3.Translation. Can Med Assoc J 1987; 136: 817-818.

57. Chesney AP, Chavira JA, Hall RP, Gary HE. Barriers to medical care of Mexican-Americans: the role of social class, acculturation and social isolation. Med Care 1982; 20: 883-891.

58. Schribner R, Dwyer JH. Acculturation and low birthweight among Latinos in the Hispanic HANES. Am J Public Health 1989; 79: 1263-1267.

59. Marks G, Solis J, Richardson JL, Collins LM, Birba L, Hisserich JC. Health behavior of elderly Hispanic women: does cultural assimilation make a difference ? Am J Public Health 1987; 77: 1315-1319.

60. Hunt SM. Cross-cultural issues in the use of socio-medical indicators. Health Policy 1986; 6: 149-158.

61. Flaherty JA, Gaviria FM, Pathak D, Mitchell T, Wintrob R, Richman JA, Birz S. Developing instruments for cross-cultural psychiatric research. J Nervous and Mental disease 1988; 176: 257-263.

62. Karno M, Burnam A, Escobar JI, Hough RL, Eaton WW. Development of the Spanish-language version of the national institute of mental health diagnostic interview schedule. Arch Gen Psychiatry 1983; 40: 1183-1188.

63. Sprangers MA, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: a review. J Clin Epidemiol 1992; 45: 743-760.

64. Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: Development and final revision of a health status measure. Med Care 1981; 19: 787-805.

65. Hunt SM, McKenna SP, McEwen J. A quantitative approach to perceived health status: a validation study. J Epidemiol and Community Health 1980; 34: 281-286.

66. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-174.

67. Campos SS, Johnson TM. Cultural considerations. In: Spilker B, Ed. Quality of life in clinical trials. Raven Press, New-York; 1990: 163-170.

68. Tugwell PX, Bombardier C, Gent M, Bennett KJ, Bensen WG, Carette S, Chalmers A, Esdaile JM, Klinkhoff AV, Kraag GR, Ludwin D, Roberts R. Low-dose cyclosporine versus placebo in patients with rheumatoid arthritis. Lancet 1990;335:1051-1055.

69. O'Boyle CA, McGee H, Hickey A, O'Malley K, Joyce CRB. Individual quality of life in patients undergoing hip replacement. Lancet 1992;339:1088-1091.

70. Meenan RF, Gertman PM, Mason JH, Dunaif R. Measuring health status in arthritis: The Arthritis Impact Measurement Scales. Arthritis Rheum 1980; 23: 146-152.

71. Fries JF, Spitz PW, Young DY. Measurement of health outcome in arthritis. Arthritis Rheum 1980; 23: 137-145.

72. Flanagan JD. A research approach to improving our quality of life. Am Psychol 1978; 33: 138-147.

Table 1. Selection of articles in the databases

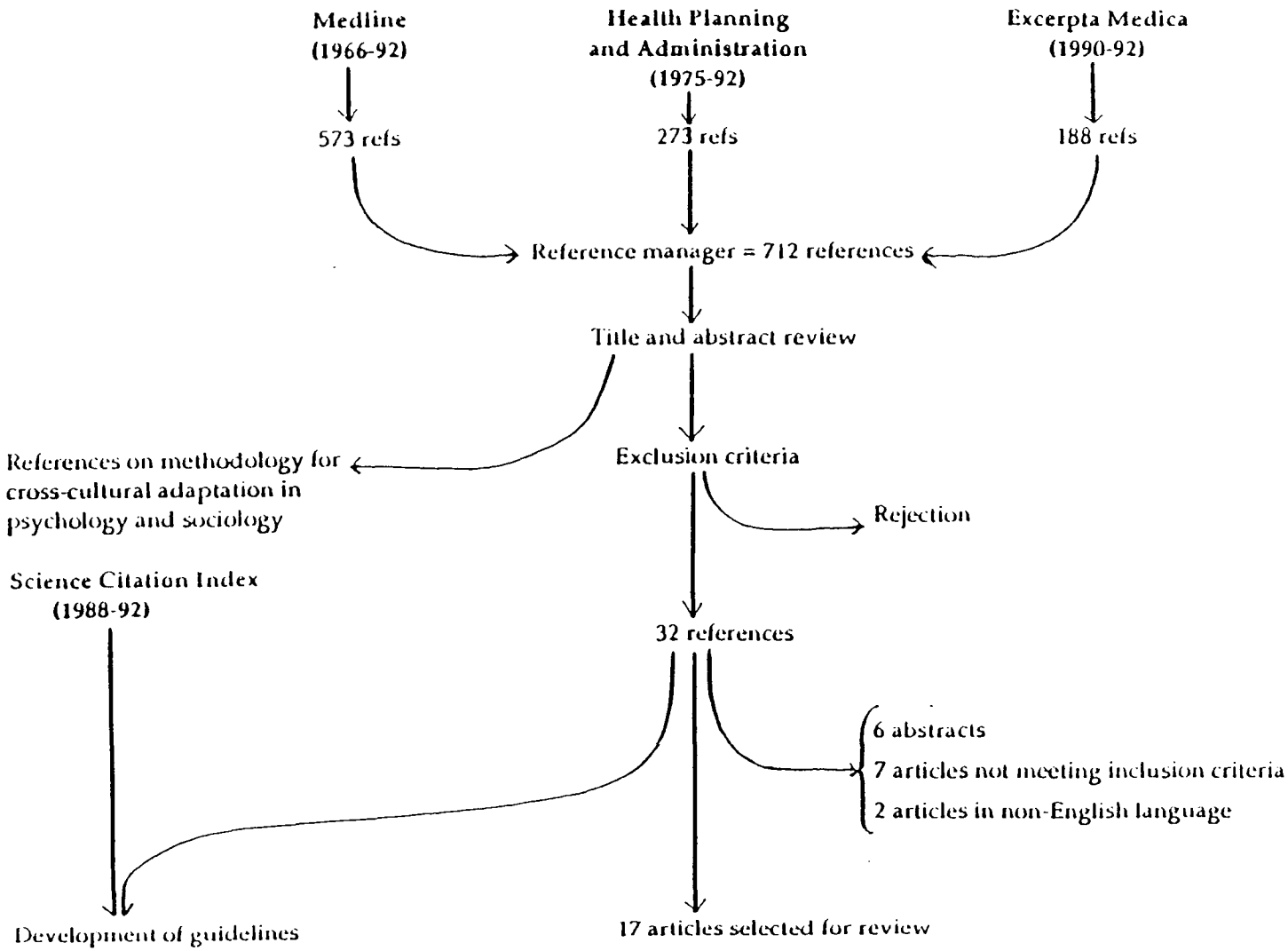


Table 2. Settings (example 2 to 5) for cross-cultural adaptation of a HRQOL measure originally developed in English in the USA (example 1).·

Examples	Culture of the target population		Language of the measure		Country of utilization		Translation required	Adaptation required
1	Same culture	*American*	Similar language	*English*	Same country	*USA*	-	-
2	Other culture	*Established immigrants (Hispanics)*	Similar language	*English*	Same country	*USA*	-	✓
3	Other culture	*British*	Similar language	*British English*	Other country	*UK*	-	✓
4	Other culture	*Recent immigrants (Hispanics)*	Other language	*Spanish*	Same country	*USA*	✓	✓
5	Other culture	*French*	Other language	*French*	Other country	*France*	✓	✓

Table 3. Guidelines to preserve equivalence in cross-cultural adaptation of HRQOL measures.*

1.Translation
Produce several translations
Use qualified translators

2.Back-translation
Produce as many back-translations as translations
Use appropriate back-translators

3.Committee review
Constitute a committee to compare source and final versions
Membership of the committee should be multidisciplinary
Use structured techniques to resolve discrepancies
Modify instructions or format, modify/reject inappropriate items, generate new items
Ensure that the translation is fully comprehensible
Verify cross-cultural equivalence of source and final versions

4.Pre-testing
Check for equivalence in source and final versions using a pre-test technique
Either use a probe technique
Or submit the source and final versions to bilingual lay people
Immigrants: Choose the language of administration or use a dual-format measure

5.Weighting of scores
Consider adapting the weights of scores to the cultural context

---

* It is not always possible to follow all the steps described here due to the design of the measure (e.g. no weighting score to be examined).

Table 4 : Reviewed articles with description of the methodology used for the cross-cultural adaptation of HRQOL measures

Original instrument	language	country	Translation	Back-translation	Committee	Pre-testing	Weighting score
Adaptation in similar language, other country							
AIMS, 1980 [70]	AIMS British, 1990 [37]	UK	N/A	N/A	-	30 RA patients Probe technique : questionnaire + interview	N/A
HAQ, 1980 [71]	HAQ British, 1986 [38]	UK	N/A	N/A	-	- 27 RA patients Check uderstandability of the original : 22 % pbs. - 33 RA patients Check understability of modified HAQ : 6 % pbs	N/A
SIP , 1981 [64]	SIP British, 1985 [39]	UK	N/A	N/A	-	-	135 consumers (stratified by age/sex) 54 health prof : 46 nurses + 8 physicians Thurstone method

Adaptation in another language, other country							
AIMS, 1980 [70]	AIMS French, 1990 [45]	Canada	1 translation 1 translator aware of objective and disease	-	2 bilingual researchers	-	N/A
HAQ,1980 [71]	HAQ Swedish, 1988 [46]	Sweden		-	-	64 RA patients Check relevance	N/A
-	HAQ portuguese, 1990 [44]	Brazil	1 translation by 2 translators aware of objectives and intent	1 back-translation by 2 translators	2 Rheumatologists 2 English professors 2 Physiotherapist 1 Occupational therapist 5 RA patients	31 RA patients with translation 22 RA Patients with substitute version Probe technique	N/A
-	HAQ French, 1992 [41]	France	3 translations (2 translators in each team, aware of objectives and disease) ?	3 back-translations (1 English-cultured)	Authors + 2 bilinguals	-	N/A

QOLS, 1978 [72]	QOLS Swedish, 1992 [48]	Sweden	1 translation by 2 translators aware of objective and disease (authors)	1 back-translation by 1 translator	-	In a pilot study	N/A
NHP, 1980 [65]	NHP Swedish, 1987 [47]	Sweden	-	-	-	-	259 patients + relatives Thurstone method
-	NHP French,1990 [43]	France	1 translation	1 back-translation	Bilinguals	-	355 Hosp patients 270 non patients Thurstone method
-	NPH Spanish + Catalan, 1990 [49]	Spain	2 translations (1 translator) by experts in NHP	1 back-translation undetailed	2 experts 10 unskilled workers	Cardiology + rheumatology patients Check acceptability	-

Adaptation in another language, same country (immigrants)							
AIMS, 1980 [70]	AIMS Spanish, 1989 [40]	USA	2 translations : 1 idiomatic + 1 literal 3 translators each 3 previous translations	2 back-translations (1) 2 previous back-translation	15 selected Hispanic American Selection of best version Iteration of back-translation	12 selected bilinguals Check comprehension and readability	N/A
-	AIMS Spanish, 1989[51]	USA	1 translation (1)	1 back-translation (1)	2 translators ?	-	N/A
SIP, 1981 [64]	SIP Spanish, 1980[50]	USA	several translations	-	-	-	-
-	SIP Spanish, 1984[52]	USA	4 translations : 3 by 1 translator and 1 by 2 translators	4 back-translations, each by 4 bilingual translators	3 English and 3 bilinguals	31 bilingual comparison to SIP English overall r = .95 (.65 to .95)	N/A
	SIP Spanish, 1992 [42]	USA	2 translations by 3 translators each : 1 literal + 1 idiomatic	1back-translation by other bilinguals ?	2 reviewers for 2 previous + 2 new translations	12 bilinguals choose the best version	29 health care consumers valuing statements
NHP, 1980 [65]	NHP Urdu, 1989 [53]	UK	2 translations (1 translator each)	1 back-translation (1 translator)	3 authors	-	-

N/A: Not applicable

Table 5. Mean scores and agreement between two judges in assessment of quality of cross-cultural adaptation studies.

	Number of studies	Mean score *	[range]	Intra-class correlation coefficient	Weighted kappa
Individual guidelines (5 sections)					
Translation	12	2.1	[1-3]		0.70
Back-translation	10	1.9	[1-3]		0.66
Committee	10	2.2	[1-3]		0.86
Pre-testing	9	2.1	[1-3]		0.88
Weighting scores	6	2.4	[2-3]		0.93
Overall guidelines					
Similar language, other country	3	0.8	[0.5-1]	1	
Other language, other country	8	1.3	[0.4-2.1]	0.96	
Other language, same country (immigrants)	6	1.6	[0.4-2.6]	0.87	
All articles	17	1.3	[0.4-2.6]	0.92	

* Mean score across studies calculated as the mean quality ratings assigned for each section with the following values:
Individual guidelines: good=3, moderate=2 and poor=1;
Overall guidelines: good=3, moderate=2, poor=1 and 0=not done.

DISABILITY AND REHABILITATION, 1992; VOL. 14, NO. 2, 71 – 77

Validity and discriminant ability of the HAQ Functional Index in early rheumatoid arthritis

FRANCIS GUILLEMIN, SERGE BRIANÇON and JACQUES POUREL

Accepted for publication: October 1991.

Keywords Rheumatoid arthritis, Stanford Health Assessment Questionnaire, functional abilities

Summary

A sensitive and valid instrument is needed in the earliest stages of rheumatoid arthritis (RA) for the measurement of functional disability. A French version of the Health Assessment Questionnaire Functional Disability Index (HAQ) has been developed, validated in an early RA sample (i.e. with disease duration < 5 yr; $n = 50$), and compared with longstanding RA ($n = 32$) and control subjects ($n = 59$). Five factors provided by principal components analysis accounted for 75% of the variability of the HAQ score (construct validity). It appeared to be significantly correlated with clinical and radiological variables and to be reproducible (r intraclass = 0·964). It proved to be discriminant in groups with various levels of disability (HAQ score = 1·332 in early RA, 1·745 in longstanding RA, and 0·152 in controls; $p < 10^{-5}$). Finally, the validity of the original scoring method of the HAQ, as compared with other scoring methods, was confirmed.

INTRODUCTION

Functional disability is one of the most important consequences of rheumatoid arthritis (RA) in the daily life of patients. While the erythrocyte sedimentation rate (ESR), haemoglobin or pain are signs indicating the activity of the disease, and radiographic damage is the result of irreversible organic lesions consequent to earlier inflammatory stages, functional disability reflects the actual clinical status of the patient.

A new strategy in RA therapy was recently put forward [1]. Unlike the classic approach which has been to give the less aggressive treatments at the onset of the disease, and to keep the more effective and more toxic treatments for later stages, this new strategy suggest using combinations of treatment from the onset of the disease in order to halt the pathological process before severe damage to joints takes place. This proposal is based on the fact that the progression of organic impairments and radiographic damage [2], as well as functional disability, is rapid in the early years [3] and slows down after 5 years [4]. Together with the development of new treatment such as low-dose oral methotrexate, this approach may provide some hope of limiting the severity of the disease. It has not yet been assessed.

Little work has been done reporting on functional disability in early RA, i.e. with disease duration under 5 years [5,6]. Since the target population for this strategy is somewhat different, any instrument intended to measure functional status needs to be tested for validity in early RA. A reliable instrument is required to assess changes occuring in patients coming under these new therapeutic schemes.

The quality of an instrument rests mainly on its validity [7,8,9]. Content validity concerns the appropriateness of the instrument for measuring what it is intended to measure and what it was designed for. It refers to the relevance and the scope of the instrument. Construct validity refers to consistency in structure within dimensions explored and from one dimension to another. Criterion validity is validity assessed by reference to a standard known as "gold standard", if one is available. The instrument should be reproducible, that is to say it provides the same results over several implementations with stable subjects, or if it is used several times by different observers. Responsiveness refers to the ability of the instrument to detect clinically important changes, even if they are small. This ability to differentiate between different levels of functional disability defines the discriminant ability of the instrument.

A number of indexes of health status [10-15] that focus in part or fully on functional status have been put forward in recent decades. Among these, the Stanford Health Assessment Questionnaire Disability Index (HAQ) (11) has been

widely used in its self-administered form, and has been altered and translated into several languages [16-19]. It has been used to measure outcome in therapeutic trials [20,5]. Several studies have been carried out on the properties of the instrument [21-23], which appears to be a genuine contribution to the measurement of characteristics of RA patients, being simple to use, inexpensive and not unduly time-consuming.

Because of the need for a sensitive, validated intrument for measuring functional disability in the earliest stages of the disease, we first developed a French version of the HAQ questionnaire, and then carried out a survey on its validity in early RA.

METHOD

The HAQ is a self-administered questionnaire investigating 8 dimensions in daily life activities (dressing and grooming, arising, eating, walking, hygiene, grip, reach, other activities). The item with the highest score within each dimension, weighted by the use of any aids or instruments, is selected and the final score is the average of the 8 dimension values.

1. General design of the study

The first stage was the development of the French version of the HAQ. The translation had to fit the concepts that guided its original elaboration, and give preference to conceptual rather than literal equivalence. The wording of some items could require adaptation to the current activities of daily living (ADL) and French lifestyle.

The second stage was to constitute three groups of patients. The construct validity and the clinical relevance of this questionnaire was tested in a group of early RA patients. Its reproducibility was assessed in early and long-standing RA patients. Its ability to discriminate between early RA, long-standing RA and controls was then assessed. Finally, the relevance of the scoring method was investigated.

2. Development of the French version of the HAQ

The original version of the HAQ was translated into French by three independent teams (each composed of two people of French culture) aware of the objectives and applications of the instrument. These translations were then submitted to medical staff for appraisal and to patients to test understanding and readability. Each version were then back-translated by three bilingual people of English culture who were unaware of the desired properties and underlying concepts of the instrument. Their back-translations were collated with the original

version. A stepwise procedure of adjustement of the intermediate French versions was used for each item, under the control of two bilingual individuals familiar with the concepts of ability and disability, and linguistic equivalence, in order to produce a final document conceptually equivalent to the original HAQ.

3. Patients and data

Three groups of patients were formed for the next stages of the study. In groups 1 and 2, clinical and biological variables recorded included age, sex, duration of the disease from the date of first symptoms, extra-articular manifestations, Ritchie index, ESR, C-reactive protein, rheumatoid factor titer and radiographic damage from X-ray of both hands scored with the Larsen method [24].

Group 1 was composed of 50 early RA patients. They were outpatients volunteering to take part in a prospective cohort study [25]. The sampling of this cohort is geographically based (Lorraine county) and aims to collect patients suffering from early RA (less than 5 years) between the ages of 20 to 70. All patients met the ARA criteria for RA [26]. They were 17 men and 33 women aged 53.3 years (17 to 89, sex ratio F/M=1.9). The mean duration of their disease, recorded from the appearance of first symptoms, was 2.1 years (range 0-4.4). They filled in the HAQ questionnaire during an initial home visit also used for recording clinical and biological data.

Group 2 covered 32 inpatients with long-standing RA from Nancy Rheumatology Clinic. They all met with ARA criteria, were aged 58.7 years (30 to 74, sex ratio F/M=2.5), and their mean disease duration was 8.6 years. They were examined, and filled in the HAQ questionnaire during hospitalization.

Group 3 was designed as a contrast group. Fifty nine healthy subjects were choosen for age and sex distribution close to that of group 1, were aged 46.7 years (30 to 62, sex ratio F/M=2.1). They were recruited in the Nancy Preventive Medicine Center which is where the population resident in Lorraine is invited to refer for health screening. They filled in the questionnaire during the visit. They were examined about possible locomotor impairment, and any chronic inflammatory disease (rheumatoid arthritis, ankylosing spondylitis) was obviously a motive for exclusion.

4. Analysis

Construct validity. A study of the internal construct validity was performed using principal components analysis on group 1 data. Principal components analysis is a multivariate descriptive analysis that makes it possible to determine those variables that contributes the most to the overall variability of the data. These variables are best loaded on a limited number of independent factors characterized by their eigenvalue, i.e. the part of total variance explained by these factors. On the

basis of eigenvalues greater than 1 [27], a selection of the most representative factors can then be computed by orthogonal rotation making it possible to obtain a clustering of variables that is easier to interpret. Loadings after rotation are correlations of these variables with the factors. Factor score coefficients are then available for each factor included in the model. They can be applied to other sets of data (see further).

In a further analysis [28], the 8 items selected in the scoring method were entered in the principal components analysis as active variables and correlations with the original 20 items and the final score, i.e. illustrative variables (not contributing to the determination of the principal components) were sought. The degree of correlation between active and illustrative variables on the one hand and the main factors on the other measures the quality of the representation of the initial data.

Relevance of the HAQ. The clinical relevance of the instrument when applied to early RA was assessed in a stepwise multiple regression. Using HAQ scores obtained in group 1 as the dependent variable and several clinical and biological variables (age, sex, duration of the disease, presence of extra-articular symptoms, Ritchie index, ESR, C-reactive protein, Larsen score) as independent variables.

Reproducibility. The reproducibility of this questionnaire was assessed with 22 patients from group 1 and 20 patients from group 2 by administering it twice within a 10-day period (mean 4.85 days) and at the same time in the afternoon (3.2 pm vs 4.2 pm; NS). The intraclass correlation coefficient was computed from analysis of variance in a two crossed random factors model [29].

Discriminant validity. The discriminant validity of an instrument is its ability to detect differences between different levels of the variable in question (here, disability) in cross-sectional study or in longitudinal follow-up. It is assessed by measuring this variable in groups assumed to show differing levels. This was effected by comparing the mean scores obtained in the 3 groups with analysis of variance and pairwise Student t-test. Further to this, groups 1 and 2 were gathered together and analysis of variance was used to highlight the influence of the duration of the disease on the relationship existing between the HAQ score and the independent variables.

Validity of the scoring method. If data, however accurate, provided by 20 items is to be put to any practical use, the items should be aggregated to produce a single final score. The relevance of the scoring method put forward by the authors can be questioned. The method, which is clinically relevant, could lead to a loss of the information provided by the original 20 items to which answers were supplied by the patients, and possibly a loss of discriminant ability. Two methods of score

calculation were compared with the original method (8-item score). The first was a straightforward averaging of the 20 item responses (20-item score) [5]. The second consisted in applying the results of principal components analysis from group 1 to both groups 1 and 2. The factor score of a patient can be obtained by multiplying the value for each item for this patient by the factor score coefficients provided by principal components analysis of group 1. Each factor score and various mathematical combinations of those factor scores, as well as 8-item and 20-item scores were tested in successive stepwise multiple regression analyses in both group 1 and 2. The partial correlation coefficients for each score with the variables in the model was used to determine which model had the greatest sensitivity to changes in patients' condition.

Data analyses were performed on BMDP [29] and SPADN [30] software.

RESULTS

1. The French version of the HAQ

In order to adapt the questionnaire closely to the French target population the wording of some items had to be altered: "are you able to stand up from an armless straight chair" and "are you able to open a milk carton" were turned respectively into "are you able to stand up from a chair" and "are you able to open a milk or fruit juice carton". All items were maintained, as well as the self-administered form of the questionnaire* (*Available on request).

2. Validity of the HAQ in early RA

Construct validity. The principal component analysis of data from group 1 showed that data loaded on 5 factors for which the eigenvalues were greater than 1, accounted for 75% of the overall interperson variability (Table 1). Orthogonal varimax rotation of the principal factors provided 5 rotated factor loadings representing the 8 dimensions (Table 2). The interpretation of these factor loadings was as follows: factor 1 covered walking and other activities, factor 2 eating and grip, factor 3 dressing and grooming, and hygiene, while other dimensions (reach and arising) were loaded on several factors. The correlation of the HAQ score with the first factor was 0.95.

The partial correlation coefficient matrix of the 8 selected item values (Table 3) failed to provide any coefficient greater than 0.554. This suggests that no significant redundancy was present between the dimensions. A second principal components analysis on these data using 20 items as illustrative variables showed high

correlation of each item with the variable selected on the corresponding factor, although it was slightly lower for the variables arising and reach.

Relevance of the HAQ. Univariate correlations between scores and variables are shown in Table 4. Using stepwise multiple regression analysis of the scores in group 1 for age, sex, duration of the disease, presence of extra-articular symptoms, ESR, CRP, Ritchie index and Larsen score as independent variables, a model that included ESR and disease duration was obtained, providing a multiple correlation coefficient $R^2 = 0.22$. The same analysis performed on group 2 gave the Larsen score and disease duration with $R^2 = 0.36$.

Reproducibility. The intraclass correlation coefficient between scores at the time of the first administration (mean: 1.769) and the second (mean: 1.818) was very high ($r = 0.964$; $p < 0.00001$).

Discriminant validity. Table 5 shows that the mean disability indexes in the three groups were significantly different (overall and pairwise analyses). Further analysis on group 1 and 2 provided the following results: there was no interaction of the groups on the relation between HAQ score and disease duration (partial $r = 0.255$ in group 1 vs 0.400 in group 2; NS) or Larsen score (partial $r = 0.131$ in group 1 vs 0.453 in group 2; NS). The relation between the score and ESR was higher in early RA (partial $r = 0.399$ in group 1 vs 0. 015 in group 2; $p < 0.05$) indicating better sensitivity to changes in that variable in this group.

Validity of the scoring method. Table 6 presents partial correlation coefficients obtained with various scoring methods for the HAQ. It shows the performance of the original 8-item method, which has higher correlation with the variables of the model in the two groups than the 20-item method or other combinations . Only the method consisting in calculating the sum of the squares of 5 factor score coefficients derived from principal components analysis of group 1, i.e. the norm of individual HAQ scores in a 5 dimensions vectorial space, provided a moderate improvement of the correlations, but this was not completely consistent in early RA.

DISCUSSION

A French version of the HAQ disability index was validated in an early RA sample, and its discriminant ability was investigated.

Functional status is of major importance for patients suffering from RA. It is defined by individual performance in ADL. Functional ability, which differs slightly from performance in that it refers to an assessment by the individual himself or

herself of ability to perform ADL, is governed by a complex interaction of physiological variables with psycho-social and environmental factors [31].

The theoretical stages in validation of an instrument measuring health status have been adequately described in rheumatology [7,8]. Since no "gold standard" is available for measurement of functional status, the validation procedure must be meticulous and should be repeated over different population samples. The qualities of the HAQ disability index have been tested and analysed in several manners [21-23], and it has been shown to be reliable, reproducible and valid in long-standing RA.

Content validity is not discussed here, since it was established when the instrument was designed [11]. However, careful attention has been paid to the translation, in terms of linguistic and conceptual equivalences, in order to preserve the original purpose of the instrument. This was assessed by translators, medical staff and patient feed-back, and was judged to be satisfactory. This meant alteration of some items, and the same was true in British, Swedish, Dutch or Brazilian versions [16-19].

Construct validity proved satisfactory in this sample of early RA patients. Two dimensions appear to be less consistent (reach and arising) since they are loaded on several factors. However consistency within each dimension and overall consistency of the 8 dimensions is acceptable.

The HAQ is assumed to vary with the duration of the disease, the age of the patients, the level of activity of the disease which is characterized by acute inflammatory phases, and the degree of irreversible organic impairment defined by articular damage which increases progressively with time as a sequela of each flare. The level of disability in daily life is expected to be preferentially related to the activity of the disease in the early stages. In long-standing RA it is more likely to be connected with articular damage and a certain amount of persisting inflammatory activity. However, a moderate overlap is expected between HAQ scores in the disease in its early and later stages, since early RA in the active phases can show the same level of disability as later RA in non-active phases. As its discriminant ability has been established in long-standing RA [21], this validity is reinforced in early RA on account of its similar or improved sensitivity to change.

The scoring method suggested for this instrument is part of the conceptual framework in which it was elaborated. It can have the advantage of avoiding over-representation of dimensions the items of which are partially redundant. However, its validity has not been tested, and it has been pointed out that the use of an ordinal scale does not always make it possible to quantify the difference between persons in a given item or in the overall index [32]. Likert-type scales have in theory the

advantage of being based on judgment of a representative sample of the target population [33]. They appear to be more relevant for measuring perceived health status than ratings obtained through independent experts, as with the Thurstone method [34]. The possibility that the same final averaged score could be obtained from people with very different disease profile makes them less attractive. In addition, the validity of the scoring method was tested in patients with disease duration of 12 years on average [35] and its responsiveness assessed in patients with average disease duration of 9.5 years [21]. Whether the method retains relevance when applied to early RA is open to question [36]. The results of construct validity testing argue for it. With regard to the best way to aggregate the information provided by the questionnaire, which is desirable for clinical use, it is important to explore the different ways in which this information can be summarized to give a single score so as to improve its discriminant ability. As in a previous study [18], the 20-item score does not provide better sensitivity to changes in the inflammatory condition (ESR) and radiographic damage (Larsen score) over disease duration as a whole. A slight improvement is obtained by computing the vectorial norm of HAQ scores derived from 5 factor scores (vector coordinates) obtained in principal components analysis in group 1, but this is too sophisticated for practical use. In addition, the values of these factor score coefficients may differ in other groups of patients. It would require further application to other RA samples to test the actual improvement in discriminant ability that it provides. Consequently, the original scoring method appears to be simpler to use and to possess sufficient discriminant ability. It is thus applicable in clinical situations and routine surveys of early RA.

Its responsiveness, i.e. the sensitivity to changes in time [37], has been explored, and judged satisfactory in North American RA patients [21,38] and modest in English patients [39]. This is at present being examined in the follow-up of group 1.

The feasibility of the questionnaire, as for the original version, is very acceptable. The cost involved is low, it is very easy to understand, and acceptable to both patients and physicians. The duration of administration was about 2 minutes and never exceeded 5 minutes. Those qualities makes it possible to administer it in a reliable way at the same time as clinical examination, or by mail.

CONCLUSION

The large amount of literature devoted to the HAQ, and the numerous clinical applications that have already been tested, all contribute to demonstrating its

relevance and usefulness. Indeed, many applications are suggested: cross-sectional analysis where it appears to be sufficiently discriminant, longitudinal follow-up of RA cohorts, and clinical trials in which it can be used as an efficiency criterion or for patient stratification [40]. It is also suited to routine care surveys since the self-report questionnaires provide valid quantitative information on the functional status of individual patients, a major factor in RA, in an attractive manner [41]. This is now available for early RA patients.

ACKNOWLEDGEMENTS

This study was supported by Contract N°891102 from INSERM (France) and by Laboratoires SANDOZ (France). We are indebted to Dr Denis Bucquet (INSERM CJF 12 Montpellier) and Dr Claude Kuntz (Centre de Medecine Preventive, Nancy) for helpful contribution.

REFERENCES

1. Wilske KR, Healey LA. Remodeling the pyramid. A concept whose time has come. J Rheumatol 1989; 16: 565-567.

2. Fuchs HA, Kaye J, Callahan LF, Nance EP, Pincus T. Evidence of significant radiographic damage in rheumatoid arthritis within the first 2 years of disease. J Rheumatol 1989; 16: 585-591.

3. Sherrer YS, Bloch DA, Mitchell DM, Young DY, Fries JF. The development of disability in rheumatoid arthritis. Arthritis Rheum 1986; 29: 494-500.

4. Meenan RF, Kazis LE, Anderson JJ. The stability of health status in rheumatoid arthritis: a five-year study of patients with established disease. Am J Public Health 1988; 78: 1484-1487.

5. Borg G, Allander E, Lund B, Berg E, Brodin U, Pettersson H, Trang L. Auranofin improves outcome in early rheumatoid arthrtis. Results from a 2-year, double blind, placebo controlled study. J Rheumatol 1988; 15: 1747-1754.

6. Eberhardt KB, Svensson B, Moritz U. Functional assessment of early rheumatoid arthritis. Br J Rheumatol 1988; 27: 364-371.

7. Bombardier C, Tugwell P. Methodological considerations in functional assessment. J Rheumatol 1987; 14 (suppl 15): 6-10.

8. Liang MH, Jette AM. Measuring functional ability in chronic arthritis. A critical review. Arthritis Rheum 1981; 24: 80-86.

9. Kirshner B, Guyatt G. A methodological framework for assessing health indices. J Chronic Dis 1985; 38: 27-36.

10. Meenan RF, Gertman PM, Mason JM. Measuring health status in arthritis: the arthritis impact measurement scale. ArthritisRheum 1980; 23: 146-152.

11. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. ArthritisRheum 1980; 23: 137-145.

12. Kaplan RM, Bush JW, Berry CC. Health status: types of validity for an index of well-being. Health Serv Res 1976; 11: 478-507.

13. Bergner M, Bobbitt RA, Pollard WE, Martin DP, Gilson BS. The sickness impact profile: validation of a health status measure. Med Care 1976; 14: 57-67.

14. Jette AM. The functional status index: reliability and validity of a self-report functional disability measure. J Rheumatol 1987; 14(suppl 15): 15-19.

15. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged: the index of ADL - a standardized measure of biological and psychological function. JAMA 1963; 185: 914-919.

16. Kirwan JR, Reeback JS. Stanford health assessment questionnaire modified to assess disability in British patients with rheumatoid arthritis. Br J Rheumatol 1986; 25: 206-209.

17. Siegert CEH, Vleming LJ, Vandenbroucke JP, Cats A. Measurement of disability in Dutch rheumatoid arthritis patients. Clin Rheumatol 1984; 3: 305-309.

18. Ekdahl C, Eberhardt K, Andersson SI, Svensson B. Assessing disability in patients with rheumatoid arthritis. Scand J Rheumatol 1988; 17: 263-271.

19. Perez EC, Mackenzie CR, Ryan C. Development of a Spanish version of the modified health assessment questionnaire. ArthritisRheum 1989; 32(suppl 4): S100.

20. Bombardier C, Ware J, Russel IJ, Larson M, Chalmers A, Read JL. Auranofin therapy and quality of life in patients with rheumatoid arthritis. Am J Med 1986; 81: 565-578.

21. Wolfe F, Kleinheksel SM, Cathey MA, Hawley DJ, Spitz P, Fries J. The clinical value of the stanford health assessment questionnaire functional disability index in patients with rheumatoid arthritis. J Rheumatol 1988; 15: 1480-1488.

22. Ferraz MB, Atra E. Rheumatoid arthritis and the measurement properties of the physical ability dimension of the Stanford health assessment questionnaire. Clin Exp Rheumatol 1989; 7: 341-344.

23. Bell MJ, Bombardier C, Tugwell P. Measurement of functional status, quality of life and utility in rheumatoid arthritis. ArthritisRheum 1990; 33: 591-601.

24. Larsen A, Dale K, Eek M. Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films. Acta Radiol Diag 1977; 18: 481-491.

25. Guillemin F. Impact of social aspects and care on prognosis of early RA. Hypotheses and preliminary results of EURIDISS project. Clin Rheum 1990; 9: 27-28.

26. Arnett C, Edworthy SM, Bloch DA, *et al.* The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. ArthritisRheum 1988; 31: 315-324.

27. Frane JW, Hill M. Factor analysis as a tool for data analysis. Commun Statist Theor Meth 1976; A5: 487-506.

28. Benzecri JP. L'analyse des données, Vol 2. Paris: Dunod Ed; 1973.

29. Dixon WJ, Brown MB, Engelman L. BMDP statistical software. University of California Press, 1989.

30. SPAD Systeme portable d'analyse de Données. CESIA 82 rue de Sévres 75007 Paris FRANCE.

31. Liang MH. The historical and conceptual framework for functional assessment in rheumatic disease. J Rheumatol 1987; 14(suppl 15): 2-5.

32. Ferraz MB, Oliveira LM, Araujo PMP, Atra E, Tugwell P. Crosscultural reliability of the physical ability dimension of the health assessment questionnaire. J Rheumatol 1990; 17: 813-817.

33. Likert R. A technique for the measurement of attitudes. Arch Psychol 1932; 140.

34. Thurstone LL. The measurement of social attitudes. J Abnorm and Soc Psychol 1931; 26: 249-260.

35. Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales. J Rheumatol 1982; 9: 789-793.

36. Feinstein AR, Josephy BR, Wells CK. Scientific and clinical problems in indexes of functional disability. Ann Intern Med 1986; 105: 413-420.

37. Guyatt GH, Deyo RA, Charlson M, Levine MN, Mitchell A. Responsiveness and validity in health status measurement: a clarification. J Clin Epidemiol 1989; 42: 403-408.

38. Liang MH, Larson MG, Cullen KE, Schwartz JA. Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. ArthritisRheum 1985; 28: 542-547.

39. Fitzpatrick R, Newman S, Lamb R, Shipley M. A comparison of measures of health status in rheumatoid arthrtis. Br J Rheumatol 1989; 28: 201-206.

40. Liang MH, Cullen KE, Larson MG. Measuring function and health status in rheumatic disease clinical trials. Clin Rheum Dis 1983; 9: 531-539.

41. Pincus T, Callahan LF, Brooks RH, Fuchs HA, Olsen NJ, Kaye JJ. Self-report questionnaires scores in rheumatoid arthritis compared with traditional physical, radiographic, and laboratory measures. Ann Int Med 1989; 110: 259-266.

Table 1. Principal components analysis of the HAQ score: contribution of main factors to the overall variability

	Factor				
	1	2	3	4	5
Eigenvalue	8.88	2.46	1.47	1.12	1.07
Factor variance (%)	44.4	12.3	7.3	5.6	5.3
Cumulative variance (%)	44.4	56.7	64.0	69.6	74.9

Table 2. Factor loadings of each item of the HAQ on the 5 first factors evaluated using orthogonal rotation

	Factor				
	1	2	3	4	5
Dressing					
Shoelaces,buttons	0.098	0.149	0.553	0.653	0.161
Shampoo your hair	0.663	0.274	0.371	0.085	0.074
Arising					
Stand up from chair	0.624	0.392	-0.351	0.174	0.278
Get in and out of bed	0.267	0.055	0.167	0.013	0.896
Eating					
Cut your meat	0.098	0.792	0.119	-0.033	0.230
Lift a full cup or glass	0.084	0.819	0.378	0.020	-0.004
Open a milk carton	0.357	0.827	-0.125	0.124	0.038
Walking					
Walk outdoors	0.801	0.097	-0.018	0.149	0.078
Climb up five steps	0.791	0.132	0.197	0.131	0.105
Hygiene					
Wash and dry your body	0.315	0.344	0.718	0.282	0.094
Take a tub bath	0.650	0.312	0.043	0.371	0.072
Get on and off the toilet	0.716	0.246	0.036	0.294	-0.013
Reaching					
Reach a 5 pound object	0.638	0.387	0.255	-0.067	0.169
Bend down	0.325	-0.090	-0.004	0.814	-0.044
Gripping					
Open car doors	0.214	0.769	0.352	0.150	-0.062
Open jars	0.322	0.653	0.170	-0.212	-0.378
Turn faucets on and off	0.283	0.398	0.604	-0.072	0.111
Other activities					
Run errands and shop	0.813	0.037	0.395	-0.017	0.115
Get in and out of a car	0.788	0.124	0.043	0.240	0.173
Do chores	0.814	0.190	0.279	-0.073	-0.147
Eigenvalue	6.044	3.926	2.144	1.627	1.252

Table 3. Partial correlation coefficient matrix[@] of items with highest value selected in each dimension of the HAQ*

	Dressing	Arising	Eating	Walking	Hygiene	Reaching	Gripping	Other act
Dressing	1.000							
Arising	0.058	1.000						
Eating	-0.242	0.266	1.000					
Walking	-0.113	0.314	-0.073	1.000				
Hygiene	0.554	0.097	0.268	0.167	1.000			
Reaching	0.220	0.173	0.041	0.110	0.103	1.000		
Gripping	0.346	-0.109	0.544	-0.052	-0.335	0.165	1.000	
Other act	0.055	-0.028	-0.011	0.521	0.145	0.150	0.267	1.000

[@] Correlation between two dimensions removing the effect of the other 6 dimensions (e.g. correlation between Dressing and Arising removing the effect of Eating, Walking, Hygiene, Reaching, Gripping and Other activities dimensions is 0.058).

* HAQ score is the average of these 8 selected items

Table 4. Univariate and multivariate (stepwise multiple regression) correlations between the HAQ score and features of the patients in early RA

	Univariate	*Multivariate*
Age	0.264	0.226
Disease duration	0.154	0.281*
Ritchie Index	0.142	0.127
ESR	0.399**	0.399***
CRP	0.368	0.019
Larsen score	0.157	0.098

* $p < 0.05$
** $p < 0.01$
*** $p < 0.001$

Table 5 Mean HAQ scores in the 3 groups of patients

	HAQ score mean (SEM)	*p pairwise*	*p overall*
Controls (n = 59)	0.152 (0.04)	< 0.01	
Early RA (n = 50)	1.332 (0.106)		< 0.00001
Longstanding RA (n = 32)	1.745 (0.026)	< 0.01	

Table 6. Partial correlation coefficients[@] of the HAQ score with main features of the patients in early RA and longstanding RA using various scoring methods

	Early RA			Longstanding RA		
	Disease duration	Larsen score	ESR	Disease duration	Larsen score	ESR
8-item	0.255	0.131	0.399*	0.400*	0.453*	-0.015
20-item	0.226	0.176	0.382*	0.303	0.452*	0.012
Factor1	0.095	-0.008	0.365*	0.301	0.344	-0.021
F1+F2+F3+F4+F5	0.201	0.238	0.354*	0.257	0.459*	0.007
$F1^2+F2^2+F3^2+F4^2+F5^2$	0.122	0.039	0.412*	0.405*	0.477*	-0.224

[@] Correlation of each feature with the HAQ removing the effect of the other two features.

* $p < 0.05$

Factor1: Linear combination of 20 items weighted by factor score coefficients of the first factor (principal components analysis in early RA)

F1+F2+F3+F4+F5: Linear combination of the 5 factor score coefficients

$F1^2+F2^2+F3^2+F4^2+F5^2$: Vectorial norm of HAQ scores (Quadratic combination of the 5 factor score coefficients)

Functional Disability in Rheumatoid Arthritis: Two Different Models in Early and Established Disease

FRANCIS GUILLEMIN, SERGE BRIANÇON, and JACQUES POUREL

Abstract. **A hypothesis that the development of functional disability in rheumatoid arthritis (RA) would be more rapid at the beginning than in established disease was studied in a group of 82 patients with RA referred to the Rheumatology Unit of Nancy, France. The relationship between disease activity, articular destruction and functional disability indicated a significant interaction (moderating effect) of disease duration. Two different models are proposed to explain functional disability. A model for early RA (less than 5 years duration) includes biological activity, and a model for established disease (more than 5 years) includes disease duration, extraarticular lesions and radiographic damage. These models could be useful in designs of therapeutic trials.** (*J Rheumatol 1992;***19:366–9**)

Key Indexing Terms:
DISABILITY RHEUMATOID ARTHRITIS MODELS INTERACTION

INTRODUCTION

The measurement of functional disability (FD) is frequently used in rheumatoid arthritis (RA) patients. It is a central aspect of the measurement of health status and several instruments have been elaborated to assess it separately or combined with other dimensions of health status (1,2,3). These instruments are validated and widely used in various languages. Cross-sectional and longitudinal studies (4,5,6,7) have clearly demonstrated that the level of FD is related to the duration of the disease.
The assessment of clinical status usually relies on variables measuring disease activity, disease severity and bony erosions or articular destructions following inflammatory flares. These variables are also related to FD. But several studies (8,9,10,11) have suggested than those variables probably do not progress linearly. The FD would worsen rapidly at the beginning of the disease and become stable after a few years of duration. This suggests that the development of FD is not related to the same factors at various stages of RA, or that the effect of those factors changes with the disease duration. So, we studied a group of patients with RA to investigate this interaction effect of the disease duration. Our hypothesis specifies that disability relates differently to disease activity or sequelae depending on the duration of the disease.

PATIENTS AND METHODS

Patients

All the subjects of the study were recruited in 1990 among in and outpatients referred to the Rheumatology Unit of Nancy hospital. Eighty two RA patients satisfied the 1987 ARA criteria (12), 52 females and 30 males, with a mean age of 55.3. The date of onset of the disease was determined as the date of first symptoms related to the disease.

Each patient was examined by a single physician and the following variables were recorded: age, sex, disease duration, number of tender or swollen joints, Ritchie index, extraarticular lesions, erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), articular damage assessed on front X-rays of both hands on the same film, scored according with the Larsen method (13) by two independent readers blinded to name and clinical status.

At the time of examination, each patient filled in a self-administered Health Assessment Questionnaire Functional Disability Index (HAQ). A French version of this questionnaire was developed and its validation has been reported elsewhere (14). Briefly, a translation from the original was carefully performed, respecting rather conceptual than literal equivalences, even adapting some items to the French

cultural context, in order to preserve its content validity. Construct validity and reliability of the instrument were then assessed and found satisfactory in RA patients. The scoring method was kept identical, providing a figure on a continuous scale with a 0-3 range.

Analysis

The relationship between FD measured by the HAQ score and the disease duration was tested using the Pearson correlation statistic. To examine the relationship between clinical variables and FD, disease duration was taken into account with other variables in a forward stepwise multiple regression analysis. This analysis assumed a model in which FD (dependent variable) was linearly related to clinical variables (independent variables). Our hypothesis of an interaction effect assumed that FD relates to independent variables depending on the disease duration. This moderating effect was sought by entering one or more interaction terms (cross-product terms of disease duration with other independent variables) in the model. Deviation-from-the-mean scores were used to estimate main effects in the presence of a significant interaction effect (15). This method has been recommended as a remedy for the multicollinearity problem - the high correlation between the cross-product term and one or more of its constituents - that often arises in the use of raw scores and leads to errors in computation (16). If such interaction effects were significant, the relationship between HAQ and independent variables would be conditional on disease duration in a multiplicative model of tricky interpretation. This has been taken into account by dividing the sample in subgroups in which the same stepwise multiple regression analyses were conducted. The subgroups were constituted in accordance with results of previous studies suggesting an evolution of the disease differing before and after 5 years following disease onset (8,9,10,11). The models were considered interpretable if the interaction effects were not significant. The quality of each model was examined by plotting the normal unweighted probability of residuals. The analysis was performed using BMDP software (17).

RESULTS

The characteristics of the patients are reported in Table 1. The mean level of FD assessed by the HAQ was 1.5 (SE=0.09). The HAQ score significantly increased with the disease duration ($r=0.396$; $p<0.001$).

In a stepwise multiple regression analysis with the age forced in the equation, FD was significantly related to the age, disease duration, ESR, Ritchie index and extraarticular lesions (multiple $R=0.64$). During the stepwise procedure, the relationship of the Larsen score to FD appeared to be highly significant (F to enter

=17.15), but this variable was removed at the step disease duration entered the equation. In a multiplicative model, FD was related to the same variables, and to significant interaction terms duration x Larsen and duration x ESR (multiple R=0.667) (Table 2).

Two groups were constituted according to the disease duration. A group of early RA (less than 5 years of disease duration) included 32 patients (mean age: 53.3; sex-ratio: 1.9) with 2.3 years average disease duration. The group of established RA (more than 5 years) included 50 patients (mean age 58.6; sex-ratio: 2.6) with 14.7 years average disease duration.

Multiple regression analyses were conducted with the same variables (Table 3). In early RA, the FD was significantly related to the ESR, Ritchie index and extraarticular lesions with no significant interaction (multiple R=0.603). In established disease, FD was significantly related to the disease duration, extraarticular lesions and Larsen score with no significant interaction (multiple R=0.696). The residuals in these two models were normally distributed.

DISCUSSION

This study evidences that the FD relates differently to activity and articular damage depending on the duration of RA. This significant moderating effect of the disease duration was taken into account by proposing two models according to the disease duration. In a model fitting for early RA, FD was mainly related to the level of biological inflammation. In a second model appropriate to established RA, FD depended linearly - and independently - on disease duration, articular destructions (assessed by radiographic damage) and extraarticular manifestations which is a characteristic of severe forms of disease.

These results are in agreement with the usual and intuitive clinical experience. At the beginning of the disease the FD is recognized to be mainly affected by inflammatory flares and to decrease moderately by the end of each flares. In established disease the FD increases more regularly and slowly, being less influenced by acute phases (18,19).

The concept of an interaction of the disease duration reflects the timed-variation of the slope of development of disability since disease onset. One might suppose that the interaction terms became non significant in the subgroup analysis because of a lower number of patients. But this is not supported by the low values of the statistics calculated in both models, as a F value of 3.9 is required for a variable to enter the model (F=1.08 and 2.86 in early and established RA respectively). Furthermore, in the established RA group, disease duration and Larsen score remained significant.

However, if an interaction were present with a larger sample of RA patients in each group, the same subgroup analysis strategy according to intervals of disease duration could be proposed, although there is no *a priori* hypothesis available to further separate subgroups. Such an interaction would reinforce the moderating effect of disease duration on the development of FD that is evidenced in this study.

Our sample may be considered representative of RA cases referred to this Rheumatology Unit in a year. However, the generalization of our conclusions is possibly limited by the fact that FD was measured with one instrument only. In such a field without reference measurement, it would be wise to confirm these results with other instruments, which are not yet validated in French language. However, a good concordance has been evidenced in the various measurement of FD (20) and the HAQ has been already used in many studies such as therapeutic trials or routine follow-up of patients. In addition, the validity of the observations might also be suggested by evidence that the HAQ score in its modified version provides data quite similar, which are correlated with ESR, radiographic score, and physical examination of the joints, suggesting that the finding may be quite generalizable.

It is well known that health status of RA patients is also related to other variables than clinical or biological ones. The role of psychological and social variables has been found independently significant (21,22). These variables would be also important to consider in further studies.

Our conclusions are directly applicable to studies interested in FD, a major component of health status in RA. To assess an improvement consecutive to any medical intervention with FD as judgment criteria (23), it will be appropriate to focus on disease activity variables in early RA, and articular destructions and sequelae in established RA. These models are reliable to better assess changes captured with instruments measuring disability. So, the actual alteration in FD following any therapeutic intervention will be more accurately assessed in early and established RA.

This recommendation is particularly important regarding with the current revision of therapeutic strategies in RA. Indeed, second-line drugs tend to be administered at earlier stages of the disease. Among the various advanced strategies, early combination of second-line drugs or sequential administration of those treatments are proposed (24,25). Their evaluation should take into account this moderating effect of the disease duration and clinical trials should preferably focus on earlier disease. In the future FD assessment should be used as a criteria among a series of quality of life measures, currently being extensively developed (10,26), to evaluate second-line therapy. Other investigations like prognosis study or supervision of RA patients in clinical practice could fruitfully use such models.

REFERENCES

1. Fries JF, Spitz P, Kraines RG, Holman HR: Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137-145.

2. Meenan RF, Gertman PM, Mason JM: Measuring health status in arthritis: the arthritis impact measurement scale. Arthritis Rheum 1980;23:146-152.

3. Jette AM: The functional status index: reliability and validity of a self-report functional disability measure. J Rheumatol 1987;14(suppl 15):15-19.

4. Liang MH, Jette AM: Measuring functional ability in chronic arthritis. A critical review. Arthritis Rheum 1981;24:80-86.

5. Wolfe F, Kleinheksel SM, Cathey MA, Hawley DJ, Spitz P, Fries J: The clinical value of the stanford health assessment questionnaire functional disability index in patients with rheumatoid arthritis. J Rheumatol 1988;15:1480-1488.

6. Fries JF, Spitz PW, Young DY: The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales. J Rheumatol 1982;9:789-793.

7. Pincus T, Callahan LF, Brooks RH, Fuchs HA, Olsen NJ, Kaye JJ: Self-report questionnaires scores in rheumatoid arthritis compared with traditional physical, radio-graphic, and laboratory measures. Ann Int Med 1989;110:259-266.

8. Sherrer YS, Bloch DA, Mitchell DM, Young DY, Fries JF: The development of disability in rheumatoid arthritis. Arthritis Rheum 1986;29:494-500.

9. Fuchs HA, Kaye J, Callahan LF, Nance EP, Pincus T: Evidence of significant radiographic damage in rheumatoid arthritis within the first 2 years of disease. J Rheumatol 1989;16:585-591.

10. Meenan RF, Kazis LE, Anderson JJ: The stability of health status in rheumatoid arthritis: a five-year study of patients with established disease. Am J Public Health 1988;78:1484-1487.

11. Weisman MH: Use of radiographs to measure outcome in rheumatoid arthritis. Am J Medicine 1987;83(suppl 4B):96-100.

12. Arnett C, Edworthy SM, Bloch DA, et al: The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315-324.

13. Larsen A, Dale K, Eek M: Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films. Acta Radiol Diag 1977;18:481-491.

14. Guillemin F, Briançon S, J Pourel: Mesure de la capacité fonctionnelle dans la polyarthrite rhumatoide: adaptation française du Health Assessment Questionnaire. Rev Rhum Osteo Art 1991;58:459-465.

15. Finney JW, Mitchell RE, Cronkite RC, Moos RH: Methodological issues in estimating main and interactive effects: examples from coping/social support and stress field. J of Health and Soc Behav 1984;25:85-98.

16. Southwood KE: Substantive theory and statistical interactions: five models. Am J Sociol 1978;83:1154-1203.

17. Dixon WJ, Brown MB, Engelman L: BMDP statistical software. University of California Press, 1990.

18. Harris ED: The clinical features of rheumatoid arthritis. In: Kelley WN, Harris ED, Ruddy S, Sledge CB, eds. Textbook of Rheumatology. Philadelphia: WB Saunders. 1989.

19. Rasker JJ, Cosh JA: Course and prognosis of early rheumatoid arthritis. Scand J Rheumatol 1989;suppl 79:45-56.

20. Liang MH, Larson MG, Cullen KE, Schwartz JA: Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. Arthritis Rheum 1985;28:542-547.

21. Deyo RA, Inui TS, Leininger J, Overman S: Physical and psychosocial function in rheumatoid arthritis. Clinical use of a self-administered health status instrument. Arch Intern Med 1982;142:879-882.

22. Hagglund KJ, Haley WE, Reveille JD, Alarcon GS: Predicting individual differences in pain and functional impairment among patients with rheumatoid arthritis. Arthritis Rheum 1989;32:851-858.

23. Liang MH, Cullen KE, Larson MG: Measuring function and health status in rheumatic disease clinical trials. Clin Rheum Dis 1983;9:531-539.

24. Wilske KR, Healey LA: Remodeling the pyramid. A concept whose time has come. J Rheumatol 1989;16:565-567.

25. Fries JF: Reevaluating the therapeutic approach to rheumatoid arthritis: the "sawtooth" strategy. J Rheumatol 1990;17(suppl 22):12-15.

26. Bell MJ, Bombardier C, Tugwell P: Measurement of functional status, quality of life, and utility in rheumatoid arthritis. Arthritis rheum 1990;33:591-601.

Table 1. Characteristics of RA patients (n = 82)

	mean	(range) *
Age (years)	55.3	(17.8 - 89.8)
Females	69.5 %	
Disease duration (years)	7.2	(0 - 31)
ESR (mm at first hour)	37	(0 - 112)
C - Reactive protein (UI/ml)	18.8	(0 - 81.1)
Ritchie index	12.7	(0 - 47)
Larsen score	22.9	(0 - 110)
Extraarticular lesions	58 %	
HAQ score	1.5	(0 - 3)

* : Percentage in categorical variables

Table 2. Stepwise multiple regression analyses with HAQ score as dependent variable in additive and multiplicative models for 82 RA patients.

	Additive model		Multiplicative model	
	b *	*p* **	*b* *	*p* **
Intercept	1.128		0.984	
Age	0.007	< 0.05	0.008	< 0.05
Sex	- 0.028	NS	0.009	NS
Disease duration	0.003	< 0.05	0.002	< 0.05
ESR	0.008	< 0.05	0.007	< 0.05
Ritchie index	0.026	< 0.05	0.024	< 0.001
Larsen score	0.008	< 0.001	0.004	NS
Extraarticular lesions	- 0.293	< 0.01	- 0.262	< 0.05
Cross-product terms				
Duration x Larsen	------		0.0001	< 0.001
Duration x ESR	------		- 0.0001	0.05
R^2	0.409		0.445	

* Regression coefficients for variables in the equation of the model.

** *p* value for variables at entry in the model.

Table 3. Stepwise multiple regression analyses using HAQ score as dependent variable in additive and multiplicative models in early RA (less than 5 years of disease duration;n = 50) and established RA (more than 5 years;n = 32)

	Early RA				Established RA			
	Additive model		Multiplicative model		Additive model		Multiplicative model	
	b *	*p* **	*b* *	*p* **	*b* *	*p* **	*b* *	*p* **
Intercept	0.844		0.844		2.235		2.235	
Age	0.009	NS	0.009	NS	- 0.008	NS	- 0.008	NS
Sex	0.249	NS	0.249	NS	- 0.371	NS	- 0.371	NS
Disease duration	0.007	NS	0.007	NS	0.003	< 0.05	0.003	< 0.05
ESR	0.009	< 0.05	0.009	< 0.05	0.001	NS	0.001	NS
Ritchie index	0.028	< 0.01	0.028	< 0.01	0.013	NS	0.013	NS
Larsen score	- 0.004	NS	- 0.004	NS	0.012	< 0.001	0.012	< 0.001
Extraart. lesions	- 0.257	< 0.05	- 0.257	< 0.05	- 0.354	< 0.05	- 0.354	< 0.05
Cross-product terms								
Duration x Larsen	------		-------		------		- 0.0003	NS
Duration x ESR	------		- 0.0002	NS	------		--------	
R^2	0.364		0.364		0.484		0.484	

* Regression coefficients for variables in the equation of the model.
** *p* value for variables at entry in the model.

FUNCTIONAL DISABILITY IN EARLY RHEUMATOID ARTHRITIS: DESCRIPTION AND RISK FACTORS.

Francis Guillemin, Theo Suurmeijer, Boudien Krol, Claire Bombardier, Serge Briançon, Dirk Doeglas, Robert Sanderman, Wim van den Heuvel.

Authors: *F Guillemin,* MD, Maître de Conference and *S Briançon,* MD, Professor, University of Nancy, France. *C Bombardier,* MD, FRCPC, Professor, University of Toronto, Canada. *ThPBM Suurmeijer,* Associate Professor, Department of Sociology and Department of Health Sciences, *B Krol,* Researcher psychologist, Department of Health Sciences, *DM Doeglas,* Researcher sociologist, Northern Center for Health Care Research, *R Sanderman,* Senior researcher psychologist, Department of Health Sciences and *WJA van den Heuvel,* Professor, Department of Sociology, Department of Health Sciences and Northern Center for Health Care Research, University of Gronigen, The Netherlands.

Reprints requests and correspondence address:

Francis Guillemin, School of Public Health
Faculty of Medicine BP 184
54505 Vandoeuvre-les-Nancy FRANCE

ABSTRACT:

Objective: To provide a description and identify risk factors of functional disability in early rheumatoid arthritis (RA).
Methods: A cohort of 337 early RA patients with disease duration under 5 years was constituted in two areas in France and The Netherlands. Clinical examination included the Ritchie index, the presence of nodes and other extra-articular manifestations, the erythrocyte sedimentation rate (ESR). The Health Assessment Questionnaire, adapted and validated in the French and Dutch languages, was used to assess functional disability.
Results: The results allowed for the cross-sectional description of a marked early functional disability, with a score of 1 (adjusted for disease activity variables) from the first year of the disease. Functional disability was increasing non-linearly with the disease duration in a quadratic model. Disease activity variables, namely ESR and Ritchie index, were identified as other important components of functional disability.
Conclusion: Consequences for the early management of RA are underlined.

Word-count: 140

Key-words: Rheumatoid Arthritis - Functional disability - Risk factor

Running head: Functional disability in early RA

INTRODUCTION

Functional disability is a major outcome for patients suffering from rheumatoid arthritis (RA). It has been described as a progressive deterioration of functional abilities along the course of the disease (1-3). However, the functional disability in early RA is not clearly documented because few studies have focused on physical functional disability at this stage (4-11). There is controversy on either a rapid progression (4,5) or a stable level of functional disability (6,7) in the early years of the disease. Methodological differences in designs as well as observation bias in selected populations may account for these discrepancies. Patients attending hospital or included in clinical trials do not represent the whole spectrum of the disease. Moreover, the definition of early RA varies from the 0-2 years up to the 0-5 years period after the onset of the disease. Also, the components of functional disability in early RA have not been much documented in the literature (12,13).

The Health Assessment Questionnaire (HAQ) (14) is a measure of functional disability, that has been developed and proved valid in RA (3,15). It is widely used and provides a useful measure of a major component of health status (16). It has been translated and validated in at least 10 languages (17).

We report on the initial data from an observational cohort study of 337 early RA patients, as part of on ongoing multicenter European study, allowing for the description of functional disability and the analysis of factors of functional disability at this early stage.

METHOD

The European Research on Incapacitating Diseases and Social Support (EURIDISS) (18) is a multicenter European study set up to explore the impact of social support on the quality of life of patients with chronic diseases. Rheumatoid arthritis has been chosen as a model in a 3-year longitudinal study. This paper reports on preliminary cross-sectional data from the Dutch and French samples.

Sampling. RA patients resident in the sampling areas were identified and asked to participate in this cohort study from October 1990 to November 1991. In France, patients in the Lorraine county were identified from several sources: inpatients and outpatients at the Rheumatology Department of the University Hospital, Nancy were listed. The 30 private rheumatologists and nearly 1400 general practitioners of the

county were contacted either personally or by medical publications that informed them about the survey and encouraged them to refer their RA patients to the study centre. Several announcements were also issued in the local and regional press, radio and television that called all newly diagnosed patients to contact their physician or the study centre.

In the Netherlands, all RA patients satisfying the inclusion criteria in the Groningen county were listed from inpatients and outpatients attending rheumatology practices in community hospitals of this county, and from the hospital registration system (inpatients) and the Rheumatology Department (outpatients) at the University Hospital, Groningen. All rheumatologists in the area stimulated their patients to participate in the study and checked for eligibility criteria.

The inclusion criteria were as follows: resident in the study areas, aged 20 to 70 years, diagnosis of RA according to the 1987 ACR criteria (19), disease duration of less than 5 years. Exclusion criteria were: other incapacitating disease, stage IV in the ARA functional classification, or probable loss to follow up (e.g. foreseeing moving outside the area). Of the patients identified and eligible for the study, 30 % and 12 % in France and the Netherlands respectively did not volunteer to participate in the study. Volunteers and non volunteers did not differ significantly neither by age (mean 54.5 (SE=0.9) and 55.7 (SE=0.6) years) nor by sex (67.3 % and 67.5 % females).

Demographic data. Age, gender and education level according to the ISCED categories (years of education) (20) were recorded.

Medical data. Medical data were recorded by a rheumatologist or a rheumatology research nurse on a special appointment either in a home visit (France) or in the outpatient department (Netherlands). The date of diagnosis was defined as the first time 4 out of 7 ACR criteria were met, consecutively or successively, as assessed by a physician or documented from medical files. This criteria-based definition for the date of diagnosis was chosen as a rule in the EURIDISS study for several reasons. It was intended to reduce inter-area or inter-physician variations in clinical ascertainment of diagnosis, and to reduce recall bias that would have been induced by using the date of first symptoms recorded from patients. Thus, it was used for the sake of comparative sampling across countries, international definition and generalizability of the results. The time elapsed between the date of diagnosis, as defined above, and the date of examination was used as the disease duration in the analysis.

Clinical variables at examination included the Ritchie index, a measure of pain (=1), wince (=2) and withdrew (=3) reaction to pressure on 24 joints (score 0-72) (21), the

presence and number of rheumatoid nodules, the presence of extra-articular manifestations of RA, erythrocyte sedimentation rate (ESR). In French patients, radiographic damage was assessed by two readers blinded to patient's status on antero-posterior hand X-rays by the Larsen score (22).
Functional disability was assessed by the Health Assessment Questionnaire Functional Functional disability Index (HAQ)(14). This instrument has been adapted and validated in the Dutch (12) and French (23) languages.

Statistical analysis. Chi-square test, Cochran-Mantel-Haenszel test for stratified analyses and Mantel-Haenszel test for trend were used for categorical data. Continuous data were analyzed using Student t test and Pearson correlation coefficients. Differences between countries were taken into account by stratifying categorical variables and analyzing continuous variables with adjustment for country.
To describe the functional disability, patients were separated in 5 cohorts by disease duration, namely [0-1[, [1-2[, [2-3[, [3-4[, [4-5] years of disease duration. Mean adjusted values of the HAQ score in each cohort were calculated by the least square method.
To determine factors of functional disability, linear regression analyses were used with the HAQ score as the dependent variable and the country, age, sex, level of education, disease activity variables, i.e. Ritchie index and ESR, and disease duration or duration squared as independent variables. The term duration squared entered into the regression analysis allowed for the potential description of a non linear relationship between functional disability and disease duration.
All analyses were conducted with SAS/PC 6.04 package (24).

RESULTS

A total of 337 patients (116 in France and 221 in The Netherlands) were enroled in the study (Table 1). The patients in the cohort were 227 females and 110 males, with a mean age of 54.5 years. The average age at first symptoms was 49.8 years and the mean age at diagnosis was 52.0 years. Their clinical characteristics included an average Ritchie index of 11.1, subcutaneous nodules in 14.4 % and extra-articular lesions in 10.5 % of patients (mainly muco-cutaneous (3.6 %) and haematological (3.0 %) abnormalities). The mean erythrocyte sedimentation rate was 25.9 mm at first hour. A variety of second-line drugs were used by 70.8 % of patients. The Dutch and French samples did

not differ significantly by clinical variables except for extra-articular features which were significantly less frequent in Dutch patients (7.7% vs 18.3%; $p<0.01$).
The clinical variables at examination were not related to the age, but differed by gender. The females reported first symptoms at a younger age (48.7 yrs vs 52.2; $p<0.05$) and were diagnosed at a younger age (50.7 vs 54.6; $p<0.01$). They were also younger at examination (53.4 vs 56.7; $p<0.05$) and had a higher Ritchie index (12.0 vs 9.3; $p<0.05$).
Dividing the sample into cohorts by disease duration (less than 1 year: n=51; 1 to 2 years: n=94; 2 to 3 years: n=88; 3 to 4 years: n=72; 4 to 5 years: n=32) and controlling for country, most of the demographic and clinical variables did not differ significantly between cohorts. In the first year, more patients received hydroxychloroquine (50.0 %) and less received methotrexate (7.1 %) than in the fith year (13.3 % and 26.7 % respectively). But the overall analysis for second-line drugs did not show statistically significant trend by disease duration across cohorts. Nor did the analysis for nonsteroidal and steroidal anti-inflammatory drugs used by 79.7 % and 14.9 % of the patients respectively.

In this cross-sectional analysis, the mean level of functional disability as measured by the HAQ score was 1.12. It was independent of age, slightly higher in Dutch patients (1.18 vs 1.01; $p=0.053$) and significantly higher in females than males (HAQ score=1.18 vs 0.99; $p<0.05$). The HAQ score was significantly related with the disease duration ($r=0.22$; $p<0.0001$) and disease activity variables (ESR and Ritchie index: $r=0.23$; $p<0.0001$ and $r=0.61$; $p<0.0001$ respectively) (Table 2). These relationships persisted in a multivariate regression analysis using age, sex, education level and country as confounding factors ($R^2=0.46$ in that model) (Table 3). In the French sub-sample, the analysis failed to include the Larsen score as an additional significant covariate ($p=0.1$) in the model.
The values of the HAQ score adjusted for disease activity variables (ESR, Ritchie index) were recalculated in each cohort. Crude values were 0.99, 0.83, 1.22, 1.29, 1.55 and adjusted values were 1.04, 0.94, 1.16, 1.23 and 1.43 in the 0-1, 1-2, 2-3, 3-4 and 4-5 cohorts respectively (Figure 1).
To better identify the components of functional disability which did show a non linear pattern across cohorts, a regression analysis using a quadratic model for disease duration successfully fitted the data with similar significant covariates, namely Ritchie index, ESR, disease duration squared and country ($R^2=0.47$) (Table 3). The meaning of the positive regression coefficient for disease duration squared in this model was that of a nonlinearly increasing functional disability across cohorts with initially slow and

subsequent faster deterioration of patients'status. No interaction term was found significant in that model.

DISCUSSION

This study describes functional disability in early RA in an observational European cohort of patients treated in routine practice. It shows that 1) functional disability is marked at the early stage of the disease, and 2) disease activity is a major contributor to functional disability in the first five years of the disease.

An important finding is the existence of a marked level of functional disability from the beginning of the disease. Indeed, after controlling for confounding factors, the HAQ score was 1 or more (on a 0-3 scale) in every cohort including the cohort with earliest disease (less than one year duration). Cases within the first year of RA are more likely to attend a physician during a flare-up than later RA cases, who are seen in more routine visits, particularly in hospital. Thus, this high HAQ score in the first cohort might merely have been related to flare-up at the time of first referral. In our study, such a bias was avoided by planning the patient's clinical examination to be independent of their routine visit with physicians. This procedure probably also accounts for the relatively low level of disease activity. The occurence of physical functional disability in the life of previously healthy people is likely to be a strong disturbance. Whether this marked score of 1 is a true reflection of patients' perceptions or is confounded by the scoring method is questionable (25). This findings has to be emphasized and is consistent with other studies of early RA conducted in various settings in North-American and European countries (7,10).
The second finding of interest is the profile of the functional disability status suggested by the data within the first 5 years of the disease in this cross-sectional observation. A non linear relationship between HAQ and disease duration was recently suggested (26) in a cohort of 209 RA patients with mean disease duration of 12 years (+/- 9 years). Authors proposed several non linear models to fit their data that did not include observations in the early stages of the disease. Our study, although cross-sectional, adds to their findings for this early period. Longitudinal follow up of this European sample is currently being conducted for further confirmation of our results. Similarly, non linear functional disability has been recently reported in a cohort of 89 RA Swedish patients (6) and in a cohort of 264 American RA patients recorded in the ARAMIS

database (7) in a longitudinal follow up. This modifies the theoretical model proposed by Rasker (4) and the findings of Sherrer (5) which were probably lacking observations of patients at the early stage of the disease. From our data after 2 years disease duration, the worsening of functional disability is substantial (12.9% yearly increase of HAQ [0.16 score unit] on average), also in accordance with recent data from the literature (7). A possible explanation of this non linear evolution might be linked with the treatment effect. Therapeutics may be initially effective and thereafter have decreasing effectiveness, leaving the disease progress towards a higher functional disability.

A model for functional disability in early RA has been recently suggested (13). Conceptually, the current disease activity status and the sequelae of previous episodes and deteriorations are accounting for functional disability. In early RA, it was mainly relying on disease activity. This study provides consistent results in a larger independent sample. It shows that the disease activity variables in the first 5 years, namely ESR and Ritchie index, and the disease duration are important factors of functional disability, while radiographic damage did not in an analysis restricted to 116 patients. This is also consistent with recent reports relating functional disability measured by self-administered questionnaire with the disease activity as measured by the number of tender or swollen joints, ESR and general health in early RA in Dutch as well as in American patients (27,28). Conversely, the number of years of education, previously proposed as a factor predicting morbidity (29), did not predict functional disability in the multivariate analysis in this study, possibly because it was lower in these patients than observed in other settings. In comparison with the general population of corresponding age groups, the number of years of education is roughly similar in RA patients in France (30). In the Netherlands, RA patients were more frequently out of the 5 to 11 years categories than in the general population of corresponding age (31). Cross-national differences in formal education would deserve more attention in future investigations.

The methodology used to constitute the EURIDISS cohort avoids common methodological biases. In clinical trials, inclusion criteria are very specific and tend to homogenize the type of patients by demographic characteristics, disease severity or disease activity level, thus introducing a selection bias. Patients selected do not represent the whole spectrum of the disease. On the other hand, the constitution of an observational cohort from in- and out-patient attending hospital only may constitute an important referral bias, since patients with more severe disease are more likely to

attend hospital. The recruitment of patients from a definite geographical area, as used in the EURIDISS study, was designed to prevent such a referral bias. Although it is not an inception cohort - the prospective inclusion of all incident cases in a given period of time would gather the ideal representative sample of the disease - the sampling method provides a good representation of the disease spectrum.

The main consequences of these findings concern the choice of an adequate measure of functional disability, the therapeutic decisions and the assessment of medical intervention in early RA. Health status measures have been developed for specific research objectives (32), or routine follow up of patients (33) and their responsiveness or sensitivity to change is considered an important measurement property (34). If the deterioration of health status toward functional disability is not linear, an adequate instrument is expected to detect such a non linear deterioration. The choice of an adequate measure of functional disability should especially take into account the high level of functional disability in the first years of the disease. Consideration of the effectiveness of a new therapy should go beyond lowering deterioration of functional disability status. Stable functional disability scores should not be considered satisfactory results and a treatment should be accepted as effective only if it offers a real improvement in the patient's daily life activities.
Functional disability appears already substantial in early RA supporting arguments to implement more aggressive therapy early in the disease course. Waiting for more severe lesions to make the decision of an agressive therapy is no longer justified and the early stage of the disease may well represent the adequate period for such treatment.

ACKNOWLEDGEMENTS. The authors are grateful to Pr A Gaucher, Pr J Pourel (Rheumatology Department, University Hospital, Nancy) and all other rheumatologists in Lorraine, France, for their collaboration. Also, the authors would like to thank the rheumatologists S van der Burg (Martini Hospital, Groningen), H Lim (Scheper Hospital, Emmen), Pr M van Rijswijk and M van Leeuwen (Academic Hospital, Groningen) and F Speerstra (Wilhelmina Hospital, Assen and Refaja Hospital, Stadskanaal) for their cooperation and for stimulating their patients to participate in this study in The Netherlands.

SUPPORT. The EURIDISS study is supported nationally by a INSERM contract CRE 891102 and a grant from the Société Française de Rhumatologie in France, and by the Ministry of Welfare, Health and Cultural Affairs and Het Nationaal Reumafonds in The Netherlands. Internationally, the study is supported by the COMAC-Health Service Research contract MR4*-0344-NL from the European Economic Community.
FG was supported by a grant from the Fondation pour la Recherche Therapeutique and the Association de Recherche sur la Polyarthrite, France.

REFERENCES

1. Harris ED. The clinical features of rheumatoid arthritis. In : Kelley WN, Harris ED, Ruddy S, Sledge CB, eds. Textbook of Rheumatology. Philadelphia: WB Saunders, 1989.

2. Pincus T, Callahan LF, Sale WG, Brooks AL, Payne LE, Vaughn WK. Severe functional declines, work functional disability, and increased mortality in seventy-five rheumatoid arthritis patients studied over nine years. Arthritis Rheum 1984;27:864-872.

3. Wolfe F, Kleinheksel SM, Cathey MA, Hawley DJ, Spitz P, Fries JF. The clinical value of the Stanford Health Assessment Questionnaire functional disability index in patients with rheumatoid arthritis. J Rheumatol 1988;15:1480-1488.

4. Rasker, JJ, Cosh JA. Course and prognosis of early rheumatoid arthritis. Scand J Rheumatol 1989;79 (Suppl):45-56.

5. Sherrer YS, Bloch DA, Mitchell DM, Young DY, Fries JF. The development of disability in rheumatoid arthritis. Arthritis Rheum 1986;29:494-500.

6. Eberhardt KB, Rydgren LC, Petterson H, Wollheim FA. Early rheumatoid arthritis. Onset, course, and outcome over 2 years. Rheumatol Int 1990;10:135-142.

7. Wolfe F, Hawley D, Cathey MA. Clinical and health status measures over time: prognosis and outcome assessment in rheumatoid arthritis. J Rheumatol 1991;18:1290-7.

8. Reeback J, Silman A. Predictors at two years in patients with rheumatoid arthritis. J Royal Soc Med 1984;77:1002-1005.

9. Caruzo I, Santandrea S, Puttini PS, Boccassini L, Montrone F, Cazzola M, Azzolini V, Segre D. Clinical, laboratory and radiographic features in early rheumatoid arthritis. J Rheumatol 1990;17:1263-1267.

10. Borg G, Allander E, Lund B, et al. Auranofin improves outcome in early rheumatoid arthritis. Results from a 2-year, double blind, placebo controlled study. J Rheumatol 1988;15:1747-1754.

11. Eberhardt KB, Svensson B, Moritz U. Functional assessment of early rheumatoid arthritis. Br J Rheumatol 1988;27:364-371.

12. van der Heidje DM, van Riel PL, van de Putte LB. Sensitivity of a Dutch Health Assessment Questionnaire in a trial comparing hydroxychloroquine vs sulphazalasopyrine. Scand J Rheumatol 1990;19:407-412.

13. Guillemin F, Briancon S, Pourel J. Functional disability in rheumatoid arthritis: two models in early and established disease. J Rheumatol 1992;19:366-9.

14. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137-145.

15. Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales. J Rheumatol 1982;9:789-793.

16. Bell MJ, Bombardier C, Tugwell P. Measurement of functional status, quality of life and utility in rheumatoid arthritis, Arthritis Rheum 1990;33:591-601.

17. Ramey DR, Raynauld JP, Fries JF. The Health Assessment Questionnaire 1992: Status and Review. Arthritis Care Res 1992;5:119-129.

18. EURIDISS group. European Research on Incapacitating Disease and Social Support. Int J Health Sciences 1990;1:217-228.

19. Arnett C, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1897 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315-324.

20. International Standard Classification of Education (ISCED), COM/ST/ISCED. Paris: Unesco, 1976.

21. Ritchie DM, Boyle JA, McInnes JM, Jasani MK, Dalakos TG, Grieveson P, Buchanan WW. Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis. Quarterly J Med 1968;37:393-406.

22. Larsen A, Dale K, Eck M. Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films. Acta Radiologica 1977;18:481-491.

23. Guillemin F, Briancon S, Pourel J. Validity and discriminant ability of the HAQ Functional Index in early rheumatoid arthritis. Disability and Rehab 1992;14:71-77.

24. SAS Users' guide. Cary: SAS Institute Inc, NC 27513, USA

25. Feinstein AR. Basic principles in the structure of clinimetric indexes. In: Feinstein AR, Ed. Clinimetrics. New Haven and London: Yale University Press; 1987:22-43.

26. Leigh JP, Fries JF, Parikh N. Severity of disability and duration of disease in rheumatoid arthritis. J Rheumatol 1992;19:1906-1911.

27. van der Heidje DMF, van Riel PLCM, van Leeuwen MA, van't HOF MA, van Rijswijk MH, van de Putte LBA. Prognostic factors for radiographic damage and physical disability in early rheumatoid arthritis. A prospective follow-up study of 147 patients. Br J Rheumatol 1992;31:519-525.

28. Pincus T, Callahan LF, Brooks RH, Fuchs HA, Olsen NJ, Kaye JJ. Self-report questionnaire scores in rheumatoid arthritis compared with traditional physical, radiographic, and laboratory measures. Ann Intern Med 1989;110:259-266.

29. Pincus T, Callahan L. Formal education as a marker for increased mortality and morbidity in rheumatoid arthritis. J Chron Dis 1985;42:449-457.

30. INSEE. National Institute of Statistic and Econmic Studies 75675 Paris, France.

31. Doeglas DM, van der Ploeg SW. Social reproduction in Czechoslovakia and the Netherlands: a comparative analysis. In: Peschar JL (Ed), Social reproduction in Eastern and Western Europe. Comparative analysis in Czechoslovakia, Hungary,The Netherlands and Poland. OOMO-Series, Department of Sociology, University of Groningen, Groningen, The Netherlands.

32. Meenan RF, Gertman PM, Mason JM. Measuring health status in arthritis: the arthritis impact measurement scale. Arthritis Rheum 1980;23:146-152.

33. Pincus T, Summey JA, Soraci SA, Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford health assessment Questionnaire. Arthritis Rheum 1983;26:1346-1353.

34. Liang MH, Larson MG, Cullen KE, Schwartz JA. Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. Arthritis Rheum 1985;28:542-547.

Table 1. Patients characteristics by country (n=337)

	The Netherlands n = 221	France n = 116	p
Age (years)	54.8	53.9	ns
Sex ratio (M/F)	1.9	2.3	ns
Education level ≤ 5 yrs	40.9	48.3	ns
5 to 8 yrs	31.4	12.1	
8 to 11 yrs	19.5	29.3	
11 to 15 yrs	7.7	5.2	
> 15 yrs	0.5	5.2	
Age at first symptoms (years)	49.7	50.2	ns
Age at diagnosis (years)	52.6	50.9	ns
Ritchie Index	10.8	11.7	ns
ESR (mm)	27.1	23.6	ns
Nodules	15 %	14 %	ns
Extra-articular lesions	7.7 %	18.3 %	<0.01
HAQ score	1.18	1.05	0.05
Drug treatment	88.6 %	86.2 %	
Second-line drugs	78.3 %	56.4 %	ns
Gold	36	17	ns
Hydroxychloroquine	58	18	
D-peni, acadione	7	8	
Salazopyrin	26	3	
Methotrexate, CyA	12	10	
NSAIDS	80.4 %	78.2 %	ns
Steroids	9.8 %	24.7 %	ns
Daily dose (mg prednisone-equivalent)	5.8	8.9	ns

Table 2. Correlations of functional disability (HAQ score) with demographic and clinical variables.

	Age	Education level	Disease duration	Ritchie index	ESR
Education level	- 0.29***				
Disease duration	- 0.001	- 0.08			
Ritchie index	0.04	- 0.1	0.12*		
ESR	0.16**	0.01	0.02	0.06	
HAQ score	0.09	- 0.12*	0.22***	0.61***	0.23***

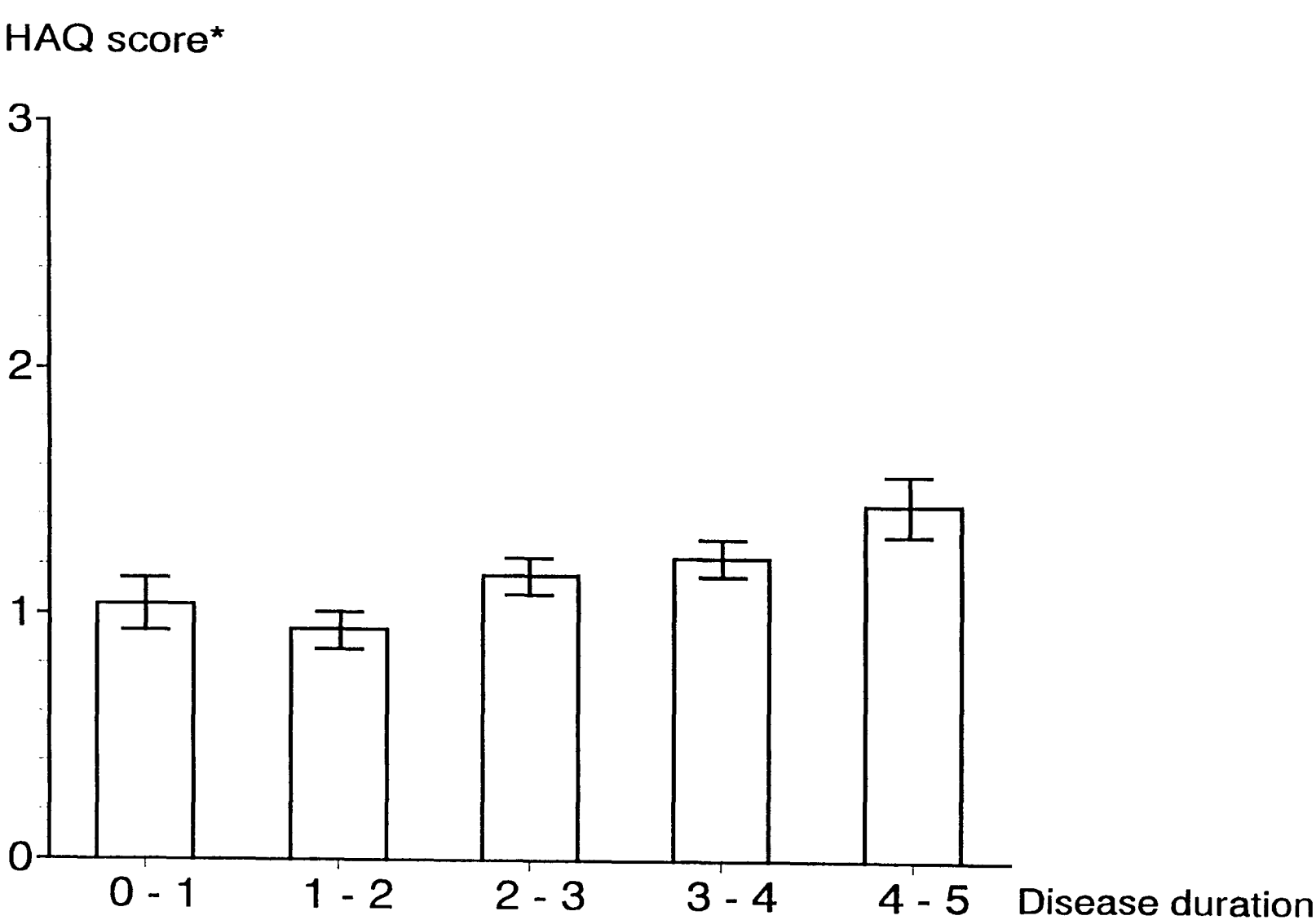
*** $p < 0.0001$
** $p < 0.01$
* $p < 0.05$

Table 3. Factors of functional disability (HAQ) by regression analysis in two models (linear and non linear for disease duration) (n=337).

Variables	Linear			Non linear		
	β	Partial R^2	p	β	Partial R^2	p
Intercept	- 0.29	--	--	- 0.15	--	--
Ritchie index	0.05	0.38	0.0001	0.05	0.38	0.0001
ESR	0.006	0.04	0.0001	0.006	0.04	0.0001
Country	0.23	0.02	0.001	0.25	0.02	0.0002
Disease duration	0.10	0.02	0.001	- 0.13	--	0.2
Disease duration^2				0.022	0.03	0.0004
Education level	- 0.02	--	0.3	- 0.02	--	0.3
Sex	0.06	--	0.4	0.07	--	0.3
Age	0.001	--	0.7	0.001	--	0.6
Extra-articular lesions	- 0.03	--	0.7	- 0.03	--	0.8
Total R^2		0.46			0.47	

LEGEND

Figure 1. Description of functional disability in early RA (n=337).



* Mean scores (± SE) adjusted (using the least square method) for Ritchie index, ESR and country.

STRESSFUL LIFE EVENTS AND DISABILITY

IN EARLY RHEUMATOID ARTHRITIS

F Guillemin, B Krol, S Briançon, R. Sanderman,
T. Suurmeijer, D Doeglas, W. van den Heuvel.

Authors: *F Guillemin*, MD, Maître de Conférence and *S Briançon*, MD, Professor, University of Nancy, France. *B Krol*, Researcher psychologist, Department of Health Sciences, *R Sanderman*, Senior researcher psychologist, Department of Health Sciences, *ThPBM Suurmeijer*, Associate Professor, Department of Sociology and Department of Health Sciences, *DM Doeglas*, Researcher sociologist, Northern Center for Health Care Research, and *WJA van den Heuvel*, Professor, Department of Sociology, Department of Health Sciences and Northern Center for Health Care Research, University of Gronigen, The Netherlands.

Reprints requests and correspondence address:
Francis Guillemin, School of Public Health
Faculty of Medicine BP 184
54505 Vandoeuvre-les-Nancy FRANCE

ABSTRACT

Objective: To investigate the relationship of stressful life events (LE) with disability in early rheumatoid arthritis (RA) in taking into account a possible stress-buffering effect of the social network.

Methods: As part of a European study (EURIDISS), 337 RA patients with disease duration under 5 years (mean=2.3) were interviewed for LE occured in the past year in France and the Netherlands. The social network (SN) composition was documented through a structured format scale and described by density and proximity characteristics. Disability was assessed with the HAQ after careful adaptation to the French and Dutch languages and cultures.

Results: Subjects reported an average 2.3 significant LE over the past year. They had an average 20 important persons in their SN. The disability, assessed by the HAQ score, did correlate negatively with the number of LE experienced ($r=-0.11;p<0.05$) and with the number of SN members in monthly contact with the subjects ($r=-0.13;p<0.05$). Controlling for country, the HAQ score was related to the disease duration, the disease activity variables, the number of LE related to health and the number of SN members in monthly contact with the patient. A significant LE times SN cross-product term entered the model better fitting the data ($R^2=0.51$).

Conclusions: Significant factors associated with disability are: disease duration, disease activity, number of LE experienced by the patients and number of SN members with monthly contacts. The results suggest the existence of a stress-buffering effect of the social network to cope with stressful life events in early RA.

INTRODUCTION

Stressful life events (LE) have been shown to increase the risk for mortality in the general population (1,2) and are significant factors favouring the onset of illness (3,4). Stressors from psychological, social and occupational nature have also been shown to aggravate the issue of a variety of chronic diseases like hypertension and other cardiovascular diseases or Grave's disease (5,6,7). On the other hand, recent reports on mortality factors of diseases reported no effect of LE neither on mortality from cardiovascular diseases (8), nor on survival, length of hospital stay and functional outcome after stroke (9) nor on relapse of breast cancer (10). In rheumatoid arthritis, a chronic condition resulting in disablement and handicap, LE have been examined as possible triggering factors for the disease onset (11,12) or as factor preceding flare-up of the disease (13). It has been found possibly associated with disease severity (14).
Disability is a central issue in the quality of daily life of RA patients. The role of LE on the development of disability has never been addressed. The recent development of valid and reliable health status measures specific for RA (15) allows for studying this issue. The impact of stressors on the development of disability has to be evaluated with consideration of the social network, i.e. the family members, partners, friends and acquaintances that provide social support and may help subjects cope with these LE and with their disease (16).
The European Research on Incapacitating Diseases and Social Support (EURIDISS) (17) is a multicenter European study set up to explore the impact of LE and social support on the quality of life of patients with chronic diseases. Rheumatoid arthritis has been chosen as a model in a 3-year longitudinal study. This paper reports on preliminary cross-sectional data from 337 Dutch and French early RA patients investigated with the objective of 1) determining the relation between LE occured in the past year and current disability status and 2) appraising the role of social network members involved in each reported LE and taking into account the support provided by these network members in assessing LE as risk factors for disability.

PATIENTS AND METHODS

Sampling. RA patients resident in the sampling areas were identified and asked to participate in this cohort study from October 1990 to November 1991. In France, patients in the Lorraine county were identified from several sources: in- and out-

patients at the Rheumatology Department of the University Hospital, Nancy, were listed. The 30 private rheumatologists and nearly 1400 general practitioners of the county were contacted either personally or by medical publications that informed them about the survey and encouraged them to refer their RA patients to the study centre. Several announcements were also issued in the local and regional press, radio and television that called all newly diagnosed patients to contact their physician or the study centre. In The Netherlands, all RA patients in the Groningen county satisfying the inclusion criteria were listed from in- and out-patients attending rheumatology practices in community hospitals of the Groningen county, and from the hospital registration system (inpatients) and the Rheumatology Department (outpatients) at the University Hospital, Groningen. All rheumatologists in the area stimulated their patients to participate in the study and checked for eligibility criteria. The inclusion criteria were as follows: resident in the study areas, aged 20 to 70 years, diagnosis of RA according to the 1987 ACR criteria (18), disease duration of less than 5 years. Exclusion criteria were: other incapacitating disease, stage IV in the ARA functional classification, or probable loss to follow up (e.g. foreseeing moving outside the area). Of the patients identified and eligible for the study, 30 % and 12 % in France and the Netherlands respectively did not volunteer to participate in the study. Volunteers and non volunteers did not differ either by age (mean 54.5 (SD=18.5) and 55.7 (SD=11.0) years) or sex (67.3 % and 67.5 % females respectively).

Demographic data. Age, gender and education level according to the ISCED categories (years of education) (19) were recorded.

Medical data. Medical data were recorded by a rheumatologist or a rheumatology research nurse on a special appointment either in a home visit (France) or in the outpatient department (Netherlands). Clinical variables at examination included the disease duration, the Ritchie index - a measure of pain (=1), wince (=2) and withdrew (=3) reaction to pressure on 24 joints (score 0-72) (20) -, the presence and number of rheumatoid nodules, the presence of extra-articular manifestations of RA and erythrocyte sedimentation rate (ESR).

Disability was assessed by the Health Assessment Questionnaire (HAQ)(15). This instrument has been translated and validated in the Dutch (21) and French (22) languages. It is a 20-item measure of disability in daily life activities incorporating the use of aids or the help of another person to perform activities, assessing the level of difficulty (no difficulty=0, some difficulty=1, much difficulty=2, unable to do=3) in 8

domains (dressing, rising, walking, eating, washing, reaching, grip strength, other activities), resulting in a score of 0 (no disability) to 3 (complete disability).
Stressful life events (LE) occured in the past year were documented in a structured format scale using a 26-item standard list derived from the Holmes and Rahe scale (23), plus one open question to elicit additional LE. Each LE reported was further investigated to assess who it was experienced by (the patient her/himself, somebody else inside or ouside the household), and how far the subject was perceiving the event as a consequence of RA or of another independent health problem.

The social network composition was recorded through a structured format scale previously validated (24). Its extent was broken down into 10 categories of network members (partner, parents, parents-in-law, children, brothers and sisters, other family members, friends and acquaintances, neighbours, other important persons). The support network (SN) was further characterized by the network density, i.e. the number of contact of the subject with each category of SN member (by visit, phone or mail), and by the network proximity, i.e. the distance of residency.

Statistical analysis. Pearson correlation coefficients were computed to investigate the relationship between the HAQ score and demographic variables, clinical variables, the number of LE and the number of social network members. Using the HAQ score as dependent variable, multiple regression analyses were conducted with the objective of assessing the relation between the disability and the number of LE by characteristics and by category of people having experienced these LE on one hand, and the number of SN members by density (number of contact) and proximity (distance of residency) characteristics on the other hand, controlling for country, disease duration and clinical disease activity variables, i.e. the age, sex, Ritchie index, ESR, that have been previously shown to be risk factors for disability in early RA (25). A stress-buffering effect of the SN moderating the impact of LE on disability was sought by introducing a LE*SN cross-product (interaction) term in the analysis after correction for possible colinearity using deviation-from-the-mean scores (26). The selection of the best model for prediction of disability was based on the highest adjusted R^2 value, which represents the percentage of the variation in the dependent variable that can be explained by variations in the independent variables.
All analyses were conducted with SAS/PC 6.04 package (27).

RESULTS

The 337 subjects included in France and The Netherlands were 227 females and 110 males, with a mean age of 54.5 years and a mean disease duration of 2.3 years (Table 1). The average number of LE occured in the past year reported was 2.34 with no difference between country (Table 2). The number of LE was correlated with the age ($r = -0.22; p<0.0001$) but not with the disease duration ($r = -0.1; p>0.5$). Sixteen percent of LE reported were occuring in the specific RA area (admission to or discharge from hospital, surgical operation). These LE occuring in the RA-related health area were removed from further analyses. The remaining mean number of LE was 1.96 per subject. When subjects were asked to characterize these LE, they attributed 31.5 % of these LE to their RA and 9.2 % LE to another disease. The proportion of LE experienced by the subjects themselves was 67.5 %, while 16.1 % were experienced by somebody else in the household and 16.4 % by a person outside the household.

The disability assessed by the HAQ score did correlate with the number of LE restricted to those experienced by the subject themselves ($r=0.11;p=0.034$) and with those attributed to health ($r=0.18;p<0.001$) in univariate analysis. LE were not correlated with clinical variables (number of nodes, of extra-articular lesions, Ritchie index, ESR).

The total number of SN members reported by the subjects in 10 categories of members (Table 3) was similar in both countries (n=20). Dutch subjects had a slightly higher number of friends and French had a slightly higher number of family members in their network. In average, there was 1.5 person living in the household. Subjects had daily contact with an average 4.3 members, weekly with 4.2, monthly with 6.1 and yearly contacts with 5.1 members of their network. A higher HAQ disability score was significantly related with a lower number of members in monthly and yearly contact with the subject ($r=-0.13;p=0.02$ and $r=-0.11;p=0.036$ respectively), while it was not with the number of members in weekly or daily contact ($r=0.05;p=0.3$ and $r=0.09;p=0.1$ respectively) in univariate analysis.

In multivariate analysis adjusted for country, age and sex, 47.7 % of the variance of the HAQ was explained by clinical disease activity (Ritchie, ESR), disease duration, the number of SN members with monthly contact (either by phone or visit) and the number or LE as a consequence of health problems (Table 4). LE related to RA and to another disease were each significantly related to HAQ in two separate analyses providing models with $R^2=0.48$ and 0.47 respectively. They have been grouped in a best fitting regression model. A higher number of LE attributed to health was related to a higher

disability. A higher number of network members in monthly contact was related with a decreasing HAQ (negative regression coefficient). These effect of the frequency of contact with SN members and of LE were significant independently each other. Finally, the adjunction of the (number of LE)*(number of SN members) interaction term in the analysis was significantly better fitting the model to the data (R^2=0.51).

DISCUSSION

This study shows that LE experienced by a subject suffering RA are significantly related to the disability in the early stage of the disease. Moreover, the results indicate that the number of contact with SN members decreases, i.e. has a moderating effect on the positive relationship between LE and the HAQ score. Because SN members provide social support to cope with LE, it suggests that the SN acts as a stress-buffer in RA subjects experiencing LE. Although the amount of variance explained by LE, SN and LE*SN interaction variables is moderate, our observation conducted in subjects with RA of less than 5 years may be important to consider since this early period of the disease is a key period for coping process and adjustment to the disease.

Most of the previous studies of LE in RA have been conducted in subjects with established disease and were interested in disease severity or disease activity as outcome variables. Few studies investigated the role of LE as a factor for disease onset. No such studies used the disability as the outcome variable.

The relation of LE with a variety of characteristics of RA like onset, perpetuation and modulation of the disease has been investigated with few significant results. From case-control studies (11,12), LE seem to have a triggering role on the onset of the disease, compared to a matched period in controls. In one study (11), families with low cohesion and expressiveness and high conflicts were also at increased risk for RA. Rimon was able to identify a sub-group of RA female with high severity that was exposed to major conflict (14). A comparison of active and inactive RA has shown a higher number of LE in years before measurement in active versus inactive disease (13).
A possible mediation of such phenomenon through immunological mechanisms has recently gained interest in the field of rheumatology and is being currently investigated by psycho-immunologists (28,29).

The measurement of LE over the past year may be subjected to recall bias or selection by subject's memory filter. In a recent study (30), an average 5 significant LE within 8 months was reported by monthly assessment while our patients reported only 2 significant LE in the past year. More than different disease duration between samples, the difference mainly relies in the methodology of this study that used monthly interviews to elicit LE from subjects. Some temporarily significant LE may have been later underestimated, thus underreported, although they might be involved in the disability process. The fall-off in the number of LE reported with time is a common phenomenon (31).

Health status as determined by current disability is a major issue in RA. It is worsening along time (32). The role of LE was not investigated so far in the determination of the level of disability. From our results, a person with early RA experiencing LE reports higher level of disability (HAQ score) than a person reporting no significant LE in the past year.

Both measures of LE and disability are self-reported, describing the personal recall of past year experience and the current perception of daily life disability, hence are free of any observer bias. The differences in demographic characteristics of the Dutch and French samples (33) have been taken into account by controlling for country. Cultural differences between countries have also been controlled for by carefully translating and adapting all self-administered questionnaires using a translation-backtranslation process under the control of an expert committee (34).

Of utmost importance in such a chronic disease is the SN of patients. A recent review of the coping process in RA presented the variation of coping strategies as depending on the support provided by SN (35). A positive support in interpersonal relationship improves the psychological condition of patients, leading to more information seeking and cognitive reappraisal coping behaviors. Using the network density (number of contact by visit, phone or mail) and proximity (distance to SN members) as estimators of the SN, our findings evidence that they are associated with a lower disability. Furthermore, this is inversely related to the number of SN members. This is in favour of a stress-buffering effect, previously described in a number of situations (6,7,36-38).

The theory of the stress-buffering effect has been formerly developed by several authors (39,40) and validated in a variety of conditions. It suggests that the SN may act as a buffer to decrease the impact of stressful LE on health. Our study offers additional elements for validity to this theory in a particular chronic disease. One may

hypothesize this is part of a behavioural strategy developed by subjects to cope with the disability from RA. It is thus probable that RA subjects in the early years of their disease can benefit in contact with higher number of members of their SN.
But a direct effect of LE on health has also been reported without interaction of the SN on this relationship (41). Actually, these two effects, direct or stress-buffered, are probably not mutually exclusive and both occuring validly in different situations (41-43).

Since our study design does not allow for causal inference, the longitudinal follow-up of this cohort is now in progress. Using these data, we will address how LE actually alter person's disability in interaction with the SN, by checking for the number of LE over the years, the stability or the transformation of the SN and the support provided. The measurement of within-subject disability variations over time will be examined in the light of these considerations.

SUPPORT. The EURIDISS study is supported nationally by a INSERM contract CRE 891102 and a grant from the Société Française de Rhumatologie in France, and by the Ministry of Welfare, Health and Cultural Affairs and Het Nationaal Reumafonds in The Netherlands. Internationally, the study is supported by the COMAC-Health Service Research contract MR4*-0344-NL from the European Economic Community.
FG was supported by a grant from the Fondation pour la Recherche Therapeutique and the Association de Recherche sur la Polyarthrite, France.

ACKNOWLEDGMENTS: The authors would like to thank the rheumatologists S van der Burg (Martini Hospital, Groningen), H Lim (Scheper Hospital, Emmen), M van Rijswijk and M van Leeuwen (Academic Hospital, Groningen) and F Speerstra (Wilhelmina Hospital, Assen and Refaja Hospital, Stadskanaal) from The Netherlands. Also, the authors are grateful to A Gaucher, J Pourel (Rheumatology Department, University Hospital, Nancy) and all other rheumatologists for their collaboration in Lorraine, France. Authors are particularly grateful to all patients for their cooperation in this study.

REFERENCES

1. Berkman LF, Syme SL. Social networks, host resistance and mortality: a nine-year follow-up study in the Alameda county residents.
Am J Epidemiol 1979;109:186-204.

2. Blazer D. Social support and mortality in an elderly community population.
Am J Epidemiol 1982;115:684-694.

3. Sarason IG, Sarason BR, Potter EH, Antoni MH. Life events, social support, and illness.
Psychosom Medicine 1985;47:156-163.

4. Rabkin JG, Struening EL. Life events, stress, and illness.
Science 1976;194:1013-1020.

5. Winsa B, Adami HO, Bergstrom R, Gamstedt A, Dahlberg PA, Adamson U, Jansson R, Karlsson A. Stressful life events and Graves'disease.
Lancet 1991;338:1475-1479.

6. Briançon S, Jolly D, Baumann M, Feard S, Blanchard F. Stressful life events and family cardiovascular risk factors.
Int J Health Sciences 1990;1:171-176.

7. Linder W, Fenerstein M. Essential hypertension and social coping behavior experimental finding.
J Human Stress 1983;7:22-31.

8. Hollis JF, Connett JE, Stevens VJ, Greenlick MR. Stressful life events, type A behavior, and the prediction of cardiovascular and total mortality over six years.
J Behav Med 1990;13:263-280.

9. Birket-Smith M, Knudsen HC, Nissen J, Blegvad N, Kohler O, Rasmussen D, Worm-Petersen S.
Psychoter Psychosom 1989;52:146-150.

10. Barraclough J, Pinder P, Cruddas M, Osmond C, Taylor I, Perry M. Life events and breast cancer prognosis.
BMJ 1992;304:1078-1081.

11. Baker GH. Life events before the onset of rheumatoid arthritis.
Psychother Psychosom 1982; 38: 173-177.

12. Darwish MJ, Armenian HK. A case-control study of rheumatoid arthritis in Lebanon.
Int J Epidemiol 1987; 16: 420-424.

13. Euller-Ziegler L, Zeigler G, Ramaioli A, Martin H. Life events and disease activity in rheumatoid arthritis.
Arthritis Rheum 1991; 34 (suppl 9): D92.

14. Rimon R, Laakso RL. Life stress and rheumatoid arthritis. A 15-year follow-up study.
Psychother Psychosom 1985; 43: 38-43.

15. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis.
Arthritis Rheum 1980;23:137-145.

16. McFarlane AH, Norman GR, Streiner DL, Roy R, Scott DJ. A longitudinal study of the influence of the psychological environment on health status: a preliminary report.
J Health and Soc Behav 1980;21:124-133.

17. EURIDISS. European Research on Incapacitating Diseases and Social Support.
Int J Health Sciences 1990;1:217-228.

18. Arnett C, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1897 revised criteria for the classification of rheumatoid arthritis.
Arthritis Rheum 1988;31:315-324.

19. International Standard Classification of Education (ISCED), COM/ST/ISCED. Paris: Unesco, 1976.

20. Ritchie DM, Boyle JA, McInnes JM, Jasani MK, Dalakos TG, Grieveson P, Buchanan WW. Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis.
Quarterly J Med 1968;37:393-406.

21. van der Heidje DM, van Riel PL, van de Putte LB. Sensitivity of a Dutch Health Assessment Questionnaire in a trial comparing hydroxychloroquine vs sulphazala-sopyrine.
Scand J Rheumatol 1990;19:407-412.

22. Guillemin F, Briancon S, Pourel J. Validity and discriminant ability of the HAQ Functional Index in early rheumatoid arthritis.
Disability and Rehab 1992;14:71-77.

23. Holmes TH, Rahe RH. The social readjustment rating scale.
J Psychosom Res 1967;11:213-218.

24. van Sonderen E. Measurement of social network and social support. Empirical results in relation to the EURIDISS instruments.
Int J Health Sciences 1990;1:203-216.

25. Guillemin F, Suurmeijer T, Krol B, Bombardier C, Briançon S, Doeglas D, Sanderman R, van den Heuvel W. Disability in early rheumatoid arthritis: Description and risk factors.
J Rheumatol (In press)

26. Finney JW, Mitchell RE, Cronkite RC, Moos RH. Methodological issues in estimating main and interactive effects: examples from coping/social support and stress field.
J Health Soc Behav 1984;25:85-98.

27. SAS Users' guide. Cary: SAS Institute Inc, NC 27513, USA

28. Sternberg EM, moderator. The stress response and the regulation of inflammatory disease.
Ann Int Med 1992;117:854-866.

29. Rubin LA, Hawker GA. Stress and the immune system: preliminary observations in rheumatoid arthritis using an in vivo marker of immune activity.
Arthritis Rheum 1993;36:204-207.

30. Raspe HH, Ensslen F, Mattussek S. Significant life events: do they modify the short- and medium-term course of rheumatoid arthritis ? In: Balint G et al, eds. Rheumatology, State of the art. ICS 984, Amsterdam, Elsevier, 1992.

31. Funch DP, Marshall JR. Measuring life stress: factors affecting fall-off in the reporting of life events.
J Health Soc Behav 1984;25:453-464.

32. Wolfe F, Hawley D, Cathey MA. Clinical and health status measures over time: prognosis and outcome assessment in rheumatoid arthritis.
J Rheumatol 1991;18:1290-7.

33. Goldberg EL, Comstock GW. Epidemiology of life events: frequency in general populations.
Am J Epidemiol 1980;111:736-752.

34. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines.
J Clin Epidemiol (In press)

35. Manne SL, Zautra AJ. Coping with arthritis. Current status and critique.
Arthritis Rheum 1992;35:1273.

36. Gore S. The effect of social support in moderating the effect og health consequences on unemployment.
J Health Soc Behav 1978;19:157-165.

37. Myers JK, Lindenthal JJ, Pepper MP, Ostrander DR. Life events and mental status: a longitudinal study.
J Health Soc Behav 1972;13:398-406.

38 Berkman LF, Leo-Summers L, Horwitz RI. Emotional support and survival after myocardial infarction. A prospective, population-based study of the elderly.
Ann Intern Med 1992;117:1003-1009.

39. Thoits PA. Conceptual, methodological nad theoretical problems in studying support as a buffer against life stress.
J Health Soc Behav 1982;23:145-159.

40. Sarason IG, Levine HM, Basham RB, Sarason BR. Assessing social support: the social support questionnaire.
J Pers Soc Psycho 1983;44:127-139.

41. Bell RA, LeRoy JB, Stephenson JJ. Evaluating the mediating effects of social support upon life events and depressive symptoms.
J Commun Psychology 1982;10:325-340.

42. Broadhead WE, Kaplan BH, James SA, Wagner EH, Schoenbach VJ, Grimson R, Heyden S, Tibblin G, Gehlbach SH. The epidemiologic evidence for a relationship between social support and health.
Am J Epidemiol 1983;117:521-537.

43. Cohen S, Wills TA. Stress, social support, and the buffering hypothesis.
Psychol Bull 1985;98:310-357.

Table 1. Total sample of RA patients (n=337)

	Mean (SE)	%
Demographic data		
Age (years)	54.5 (0.9)	
Females		67.3
Education level ≤ 5 yrs		43.5
5 to 8 yrs		24.7
8 to 11 yrs		22.9
11 to 15 yrs		1.8
>15 yrs		7.1
Medical data		
Disease duration (years)	2.3 (0.1)	
Ritchie index (0-72)	11.1 (0.7)	
Nodules		14.4
Extra-articular lesions		10.5
ESR (mm at first hour)	25.9 (1.7)	
Disability status		
HAQ score (0-3)	1.12 (0.08)	

Table 2. Distribution of stressful life events (LE) in the past year by type of event (Each subject may have experienced more than one LE over the year)

Type of LE	Number of subject experiencing each event*	Type of LE	Number of subject experiencing each event*
Admission/discharge from hospital	198	Loss, damage or theft of property	17
Death	87	Accidents	16
Illness	75	Examinations	15
Surgical operation	62	Starting or finishing school	11
Time off work (illness)	55	Pregnancy	10
Other important LE	38	Trouble at work (arguments, strikes)	8
Birth	31	Contact with police or court	7
Return to work	27	Burglaries	6
Loss of job or resignation	24	Divorce	4
Making important decisions	21	Separation	3
Marriage	20	Miscarriage	2
Retirement	19	Engagement	1
Moving home	18		

Total number of LE per subject		Category of people experiencing LE	Percent of total LE
Mean number	2.34	Respondents	67.5
Mean number excluding hospital	1.96	People in the household	16.1
		People outside the household	16.4

Table 3. Social network members by category, and mean number of network member by proximity (place of residence) and density (frequency of contact) of the network.

Network category	France Average %	The Netherlands Average %
Partner	4.5	4.4
Parents	3.8	3.6
Children	12.3	12.6
Brothers/sisters	29.3	28
Parents-in-law	2.9	3
Children-in-law	6.1	6
Other relatives	9.4	7
Friends	17.5	19
Neighbours	10	11
Other important persons	4.2	5
Mean number (SD) of SN members per subject		19.7 (6.7)

Place of residence	Mean (SE)
In the household	1.5 (0.1)
Outside the household	18.2 (0.4)

Frequency of contact	Mean (SE)
Once a year	5.1 (0.2)
2 to 11 times-a-year	6.1 (0.2)
1 to 3 times-a-month	4.2 (0.2)
1 to 7 times-a-week	4.3 (0.2)

Table 4. Number of stressful life events (LE) and number of social network (SN) members in regression models for disability by HAQ score (n=337).

		Model without interaction		Model with interaction	
	Beta	Partial R^2	p	Partial R^2	p
Intercept	- 0.153		0.0001		
Ritchie index	0.05	0.38	0.0001	0.385	0.0001
ESR	0.005	0.04	0.001	0.036	0.0001
Disease duration	0.116	0.02	0.0005	0.024	0.0001
Country	0.293	0.005	0.09	0.02	0.0005
Number of health-related LE	0.465	0.01	0.008	0.016	0.001
Number of SN members in monthly contact	- 0.028			0.02	0.0007
Contact SN*Health LE	- 0.046			0.014	0.003
Total R^2			0.477		0.516

LONGITUDINAL DETERMINANTS OF FUNCTIONAL DISABILITY IN EARLY RA.

Francis Guillemin, Jean-Marc Virion, Serge Briancon

AVERTISSEMENT

Le travail qui suit est un manuscrit en préparation au moment de la présentation de cette thèse.

Les études longitudinales de l'incapacité fonctionnelle au cours de la polyarthrite rhumatoïde publiées dans la littérature à ce jour n'utilisent que des méthodes de régression simple, au mieux ajustée sur le degré d'incapacité initiale. Des méthodes d'analyse de données tenant compte du fait que des mesures longitudinales se font chez un même sujet et ne sont donc pas statistiquement indépendantes ont été récemment développées. Il est nécessaire de les utiliser pour produire des résultats pertinents et il faut souligner le risque d'obtenir des résultats erronnés par des méthodes inadéquates.

Il s'agit donc d'un travail préliminaire portant sur une période de suivi courte (un an), montrant que l'emploi des méthodes classiques de régression ferait conclure à tort à l'existence d'un changement significatif au cours du temps, tandis que l'emploi de méthodes plus précises montre qu'il n'en est rien.

Cependant, l'objectif de ce travail n'est pas de faire une revue méthodologique spécialisée de toutes les techniques d'analyse de données longitudinales. Il existe en effet d'autres méthodes plus récentes non présentées ici, qui sont plus complexes et sortiraient du cadre de travail fixé. Toutefois le lecteur intéressé par ces méthodes peut utilement consulter un mise au point récente de Liang KY, Zeger SL. Regression analysis for correlated data. Ann Rev Public Health 1993;14:43-68.

Correspondence address: School of Public Health
Faculty of Medicine BP 184
54505 Vandoeuvre-les-Nancy
France
Fax: (33) 83 59 26 90

ABSTRACT

Objectives: To identify determinants of functional disability in a cohort of early RA patients. To compare respective advantages of various statistical techniques available for longitudinal data analysis.

Patients and methods: A cohort of 111 RA patients satisfying ACR criteria, with disease duration of less than 5 years (mean=2.6) was followed up for one year. Functional disability, measured by the HAQ, and several clinical and biological disease activity and severity variables were assessed at baseline and final measurement time. Several analyses were conducted using linear conditional and non conditional regression with baseline and final covariates. Further analysis of variance with repeated measures was used to take into account within-subject correlation over time.

Results: Baseline average HAQ score was equal to 1. Simple conditional linear regression identified baseline HAQ score as a strong predictor of final HAQ, but did not when using a model with summary measure of change as dependent variable. Analysis of variance with repeated measure allowed to identify that most of the difference observed stemed from between-subject origin, while there was no significant within-subject change.

Conclusion: Determinants of disability at 1 year in early RA are the disease duration at entry in the study, Ritchie index and ESR changes. ANOVA with repeated measure seems a better model to identify these predictors and their between-subject origin. This highlights the importance of individual management of the patients.

Key-words: Functional disability - Rheumatoid arthritis - Longitudinal analysis - Repeated measures - Risk factors

INTRODUCTION

The long-term evolution of rheumatoid arthritis (RA) is resulting in an overall deterioration of joints, functional disability, health status and quality of life. Some studies followed patients prospectively to identify the pattern of the disability along time (1-7). In these studies, the duration of follow-up ranged from 2 to 20 years and they mainly focused on established RA. A linear deterioration of disability with time has been frequently reported. More seldom, non linear models have been used to describe disability (3,5). They also attempted to determine factors contributing to this deterioration, like joint damage (on radiographs), age, sex or immunological characteristics.

Early RA has not yet been much documented in the longitudinal way (2,3,6). A reason might be the difficulty to incorporate patients in cohort studies at a time when the diagnosis frequently remains uncertain. The recent revision of the ACR 1987 criteria, together with a clarification of their use, i.e. using it for classifying patients for epidemiologic studies after they have been diagnosed as RA by clinicians rather than to use it for ascertaining the diagnosis, partly helped overcome this problem (8).

The recent emphasis put on longitudinal studies focuses the interest on a crucial period of the disease: the early years where a lot of psychological disturbances may occur, and where sometimes major social coping processes have to be implemented by the patient and her/his social network (9).

The objective of this study is to investigate the early evolution and the determinants of functional disability in RA in the longitudinal analysis of a European cohort of early RA patients with one-year follow-up. The advantages of various statistical methods used to analyze such longitudinal data are depicted.

METHODS

Sample

One hundred and sixteen patients have been included in a cohort study in the Lorraine area, France. They are part of a European (EURIDISS) cohort study set up to explore the impact of social support on the quality of life of patients with rheumatoid arthritis (10). All patients are aged between 20 to 70 years, with RA according to ACR criteria of less than 5 years disease duration recruited within the Lorraine county (11) entered the

study in 1991. At the end of the one-year follow-up period, 5 patients were lost to follow-up (3 refused to further participate and 2 left the study area).

Measurements

All 111 other patients had complete assessment of their condition at initial and final measurement time after one year recorded by a rheumatology research nurse in a home visit on a special appointment.

Outcome measures at examination included the Ritchie index, a measure of pain (=1), wince (=2) and withdrew (=3) reaction to pressure on 24 joints (score 0-72) (12), the presence of extra-articular lesions (subcutaneous nodes, muco-cutaneous, hematological, renal or neurological lesions) of RA, erythrocyte sedimentation rate (ESR). Radiographic damage was assessed by two readers blinded to patient's status on antero-posterior hand X-rays by the Larsen score (13).

Functional disability was assessed by the Health Assessment Questionnaire Functional Functional disability Index (HAQ)(14). This instrument has been adapted and validated in the French language (15).

Analysis

Based on previous cross-sectional analysis performed on baseline data, and on data available from the baseline and final measurements, multivariate analyses were conducted to identify and to compare the results from various statistical methods applied to longitudinal data.

An non conditional linear stepwise regression analysis with final HAQ score as dependent variable and a set of independent variable including patients characteristics at baseline, with and without adjustment to at final measurement time variables, was first performed (Model 1).

Model 2 was a conditional model using the same dependent and independent variables as in model 1, with the addition of the baseline HAQ score in the independent set, so as to obtain an estimate of predictors of the final HAQ score conditioned on the baseline HAQ score.

Since such model do not take into account the fact that measurements at different time points are from the same subjects and that successive observations on a given subject are likely to be correlated, alternative approaches are recommended. One of those is the use of summary measures (16). A summary measure simple to use when a monotonic change is hypothesized is the difference between final and baseline measures. So we

used Δ HAQ (Final HAQ - Baseline HAQ) as a dependent variable, with the same sets of covariates in non conditional (Model 3) and conditional analyses (Model 4).
Another way of taking the within-subject correlation into account is to consider repeated measures in all subjects at different time points as the phenomenon of interest, here the average change in HAQ score over time. An analysis of variance with repeated HAQ score measures was run (Model 5) that allowed for separate between-subject and within-subject analyses of variance. This provided insights into the between-subject average variations across time on one hand, and the within-subject time effect on the other hand. Additionally, BMDP 2V allowed for including series of repeated measures in independent variables, like Ritchie index at baseline and final time points, so as to avoid similar statistical problem (same subjects, within-subject correlation across time).
All analyses were conducted using BMDP statistical software (17).

RESULTS

The cohort included 77 females and 34 males, aged 53 years at entry in average, with a mean duration of RA of 2.6 years. Table 1 contains a description of the patients, with an average 5 percent increase in HAQ score over 1 year, and of potentially determinant variables of HAQ score at baseline and final measurement times.
Multivariate analyses were conducted with these covariates as possible predictors for disability at one year. Simple linear regression analysis (Model 1) showed a significant linear relationship of the fianl HAQ score with the disease duration at entry in the study, baselie presence of extra-articular lesions and Ritchie index (R^2=0.42) (Table 2). Considering also final measures allowed for the adjustment to final presence of extra-articular lesions, without alteration of the overall fit of the model.
A similar analysis conditioned on baseline HAQ score (Model 2) showed an important contribution of this baseline to the final HAQ score. Only baseline Ritchie index was an independent variable related to the final HAQ score, explaining a significant part of the variance (R^2=0.7). The adjustment to final measurement covariates showed an improvement of the model prediction with final presence of extra-articular lesions as additional significant covariate (R^2=0.83).
As to overcome the correlation between baseline and final measures of a given variable (baseline and final HAQ score: r=0.83)), a model with summary measure (Model 3) failed to provide any significant predictor (none covariate entered). When the final measures were also considered, only the presence of extra-articular lesions entered this

model which fitted poorly to the data (R^2=0.04). The conditional analysis provided similar results.

Finally, an analysis of variance with repeated measures incorporated repeated measures of HAQ scores as dependent variables, and baseline measures of age, sex, disease duration and Larsen score, plus repeated measures of covaribles (Ritchie index, presence of extra-articular lesions, ESR) (Model 5) (Table 3). The result of between-subject analysis of variance showed that there was a significant linear relationship between HAQ scores averaged across time and 3 covariables, i.e. disease duration, Ritchie index and ESR with estimated regression coefficient of 0.11, 0.05 and 0.006 respectively. The within-subject analysis of variance showed neither significant within-subject change of HAQ score over time, nor significant relationship between within-subject covariable change and HAQ change.

In summary, this model showed no significant average HAQ score change and no relation with covariates change over time, and indicated that most of the variation of HAQ observed longitudinally was related to between-subject differences with patients of higher RA disease duration, Ritchie index or ESR having higher HAQ score.

DISCUSSION

The results of the one-year follow-up of this cohort of early RA patients indicate that predictors of disability may vary a lot depending on the statistical method employed for their identification. Based on repeated measures analysis, which has several advantages and better validity, these predictors are disease duration at entry, Ritchie index and ESR. It also clarifies the issue of the origin of the variations observed, i.e. between-subject differences rather than within-subject changes.

Analysis using simple linear regression yielded results close to than those observed in cross-sectional analysis (11). In addition, the conditional model elicited baseline HAQ score as a strong predictor, as previously observed in established RA (5). But such approaches can provide misleading impression of the way individual subjects respond over time and gives no information about variation among subjects in their response over time. Multiple statistical tests are not independent at each time, thus not valid (16). Indeed, there is commonly a strong correlation between baseline and final HAQ score (r=0.83 in our study, 0.77 in (5)), that is not taken into account in these models.

The use of a summary measure of HAQ change over time coped with this problem and showed a lower prediction ability of the model (Model 3) where presence of extra-articular lesions at final measurement time weakly explained the variance of the

outcome. A conditional model has also been tested, because subjects may have similar change, but starting from different baseline HAQ level. It did not improve the prediction of one-year HAQ score change. However, the negative (though not significant) HAQ regression coefficient indicates that patients with high baseline score had a small HAQ change, while those with lower baseline score had a greater change over time, probably correponding to a regression-to-the-mean phenomenon (18).
Discrepancies between these regression models can also be partly attributed to some instability of the statistical stepwise computation process and to the moderate amount of score change.

The advantages of the repeated measure analysis of variance are its ability to decompose the total variance (σ_t) of the average HAQ change across time into 2 components, i.e. between-subject variance (σ_b) and within-subject variance (σ_w) where $\sigma_b + \sigma_t = \sigma_t$, and then to perform separate analyses of each variance component. It indicates the significant covariates in each source of variance. The longitudinal analysis of data with this method has also been shown to improve statistical power in treatments comparison (19).
In the particular case where there was no significant HAQ change over time, it showed that most of the difference observed in the HAQ score is explained by disease duration at entry, Ritchie index and ESR. These variables are determinants of between-subject overall HAQ score, but not of within-subject differences in HAQ change pattern.

This is an interesting finding since the study is quite short-term, compared to the lifetime duration of this chronic disease starting aroun the age of 50 in average. The average change in HAQ score over one year (+ 5 percent) in similar to the change reported in other series (4,20).
This contributes to the current knowledge of the disease at the early stage, particularly of disability, an important patient self-assessed outcome. Functional disability measures refer to both subjective and objective assessment of its own condition. Subjective perception of disability can be partly coped with by psychosocial adjustment, self-help and support received from the social network. Objective deterioration of the functional ability can be fighten against by various means, including job counselling, aids and accessories use in daily life activities, or other adjustments that could be fruitfully oriented by social workers. This study emphasizes that it has to be assessed and counselled on basis of individual needs.

Conclusion: Statistical techniques employed to analyze longitudinal data may strongly influence the determination of predictors of disability. Methods taking into account the correlation between repeated measures across time and/or identifying sources of between- and within-subject differences avoid misleading conclusions about the independent significance of potential determinants. Analysis of variance with repeated measures showed that change in HAQ score over one-year significantly vary across patients in early RA, depending on disease activity (Ritchie index, ESR) and severity (presence of extra-articular lesions), although there was no significant average HAQ change across such a short period of time.

REFERENCES

1. Wolfe F, Hawley D, Cathey MA. Clinical and health status measures over time: prognosis and outcome assessment in rheumatoid arthritis. J Rheumatol 1991;18:1290-7.

2. Eberhardt KB, Rydgren LC, Petterson H, Wollheim FA. Early rheumatoid arthritis. Onset, course, and outcome over 2 years. Rheumatol Int 1990;10:135-142.

3. Rasker JJ, Cosh JA. Course and prognosis of early rheumatoid arthritis. Scand J Rheumatol 1989;79:45-56.

4. Wolfe F, Cathey MA. The assessment and prediction of functional disability in rheumatoid arthritis. J Rheumatol 1991;18:1298-1306.

5. Leigh JP, Fries JF. Predictors of disability in a longitudinal sample of patients with rheumatoid arthritis.
Ann Rheum Dis 1992;51:581-7.

6. van der Heidje DMFM, van Riel PLCM, van Leeuwen MA, van't Hof MA, van Rijswijk MH, van de Putte LBA. Prognostic factors for radiographic damage and physical disability in early rheumatoid arthritis. A prospective follow-up study of 147 patients. Br J Rheumatol 1992;31:519-525.

7. Reeback J, Silman A. Predictors at two years in patients with rheumatoid arthritis. J Royal Soc Med 1984;77:1002-1005.

8. Arnett C, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1897 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315-324.

9. Parker JC, Bradley LA, DeVellis RM et al. Biopsychosocial contributions to the management of arthritis disability.
Arthritis Rheum 1993;36:885-889.

10. EURIDISS group. European Research on Incapacitating Disease and Social Support. Int J Health Sciences 1990;1:217-228.

11. Guillemin F, Suurmeijer T, Krol B, Bombardier C, Briançon S, Doeglas D, Sanderman R, van den Heuvel W. Functional disability in early rheumatoid arthritis: description and risk factors. J Rheumatol (In press)

12. Ritchie DM, Boyle JA, McInnes JM, Jasani MK, Dalakos TG, Grieveson P, Buchanan WW. Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis. Quarterly J Med 1968;37:393-406.

13. Larsen A, Dale K, Eck M. Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films. Acta Radiologica 1977;18:481-491.

14. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137-145.

15. Guillemin F, Briancon S, Pourel J. Validity and discriminant ability of the HAQ Functional Index in early rheumatoid arthritis. Disability and Rehab 1992;14:71-77.

16. Matthews JNS, Altman DG, Campbell MJ, Royston P. Analysis of serial measurements in medical research. Br Med J 1990;300:230-236.

17. BMDP Statistical software. Dixon WJ, Ed, University of California Press, 1990.

18. Davis CE. The effect of regression to the mean in epidemiologic and clinical studies. Am J Epidemiol 1976;104:493-498.

19. Edworthy SM, Bloch DA, Brant RF, Fries JF. Detecting treatment effects in patients with rheumatoid arthritis: the advantage of longitudinal data. J Rheumatol 1993;20:40-44.

20. Leigh JP, Fries JF, Parikh N. Severity of disability and disease duration in rheumatoid arthritis. J Rheumatol 1992;19:1906-1911.

Table 1. Patients characteristics at baseline and one-year measures (n = 111)

Characteristics		m (SD)
Age (years)		53.9 (10.9)
Sex	Female	69.4 %
	Male	30.6 %
Baseline measures		
	Disease duration (years)	2.6 (1.4)
	Ritchie index	11.7 (7.9)
	Presence of extra-articular lesions	18.2 %
	ESR	26.5 (25.2)
	Larsen score	14.4 (15.8)
	HAQ score	1.01 (0.7)
Final measures (one year)		
	Ritchie index	12.1 (10.2)
	Presence of extra-articular lesions	24.3 %
	ESR	27.2 (21.9)
	HAQ score	1.06 (0.78)

Table 2. Results of linear regression analyses in various models fitted to one-year longitudinal data in 111 early RA patients (significant regression coefficients of variables entered in the model).

	Model 1 Non conditional		Model 2 Conditional		Model 3 Non conditional		Model 4 Conditional	
Dependent variable	HAQ final		HAQ final		Δ HAQ#		Δ HAQ#	
Independent variables	Baseline	Baseline + Final	Baseline	Baseline + Final	Baseline	Baseline + Final	Baseline	Baseline + Final
Baseline								
HAQ	NU	NU	0.83*	0.8*	NU	NU	--	--
Age	--	--.	--	--	--	--	--	--
Sex	--	--	--	--	--	--	--	--
Disease duration	0.14*	0.11*	--	--	--	--	--	--
Ritchie index	0.05*	0.05*	0.15*	0.01*	--	--	--	--
Presence of extra-articular lesions	0.36*	0.30*	--	--	--	--	--	--
ESR	--	--	--	--	--	--	--	--
Larsen score	--	--	--	--	--	--	--	--
Final								
Ritchie index	NU	--	NU	--	NU	--	NU	--
Presence of extra-articular lesions	NU	0.3*	NU	0.23*	NU	0.17*	NU	0.17*
ESR	NU	--	NU	--	NU	--	NU	--
R^2	0.421	0.425	0.703	0.719	--	0.038	--	0.038

$ Intercept coefficients, and coefficients of non significant variables are not reported in this table.
# Δ HAQ = Final HAQ - Baseline HAQ
* $p < 0.05$
NU = Not used in the model

Table 3. Results of the analysis of variance with repeated measures of HAQ score (Model 5).

	Between-subject ANOVA		Within-subject ANOVA	
	Coeff #	p	Coeff #	p
Dependent repeated variable: Baseline, Final HAQ				
Independent variables				
Time effect			--	0.24
Age	0.005	0.29 .		
Sex	--	0.56 .	--	0.15
Disease duration	0.11*	0.005 .		
Ritchie index (a)	0.05*	0.00001	-0.008 .	0.41
Presence of extra-articular lesions (a)	--	0.34 .	--	0.23
ESR (a)	0.006*	0.01 .	-0.0005	0.88
Larsen score	0.004	0.31 .		

(a) Covariate repeatedly assessed at each measurement time point
Coefficients are regression of covariates to the overall HAQ scores in the between-subject analysis, and to the HAQ change across time in the within-subject analysis. They are not available for categorical variables entered as grouping factors.
* $p<0.05$

ANNEXE

HEALTH ASSESSMENT QUESTIONNAIRE
ADAPTATION FRANÇAISE

Veuillez indiquer d'une croix la réponse qui décrit le mieux vos capacités au cours des 8 derniers jours.

	Sans AUCUNE difficulté	Avec QUELQUE difficulté	Avec BEAUCOUP de difficulté	Incapable de le faire
S'HABILLER ET SE PREPARER				
Etes-vous capable de :				
- Vous habiller, y compris nouer vos lacets et boutonner vos vêtements ?	|___|	|___|	|___|	|___|
- Vous laver les cheveux ?	|___|	|___|	|___|	|___|
SE LEVER				
Etes-vous capable de :				
- Vous lever d'une chaise ?	|___|	|___|	|___|	|___|
- Vous mettre au lit et vous lever du lit ?	|___|	|___|	|___|	|___|
MANGER				
Etes-vous capable de :				
- Couper votre viande ?	|___|	|___|	|___|	|___|
- Porter à votre bouche une tasse ou un verre bien plein ?	|___|	|___|	|___|	|___|
- Ouvrir une brique de lait ou de jus de fruit ?	|___|	|___|	|___|	|___|
MARCHER				
Etes-vous capable de :				
- Marcher en terrain plat à l'extérieur ?	|___|	|___|	|___|	|___|
- Monter 5 marches ?	|___|	|___|	|___|	|___|

Veuillez indiquer d'une croix si vous utilisez habituellement un de ces appareils ou accessoires pour effectuer ces activités :

Canne	|___|	Accessoires pour s'habiller (crochet à bouton ou à fermeture-éclair, chausse-pied à long manche...)	|___|
Déambulateur	|___|	Ustensile spécialement adapté	|___|
Béquilles	|___|	Chaise spécialement adaptée	|___|
Chaise roulante	|___|	Autre(s) (préciser) :	

Veuillez indiquer les activités pour lesquelles vous avez besoin de l'aide de quelqu'un :

S'habiller et se préparer	|___|	Manger	|___|
Se lever	|___|	Marcher	|___|

Adaptation française du HAQ Disability Index - CHRU Nancy - 1990

Veuillez indiquer d'une croix la réponse qui décrit le mieux vos capacités au cours des 8 derniers jours :

	Sans AUCUNE diffi-culté	Avec QUELQUE diffi-culté	Avec BEAUCOUP de dif-ficulté	Incapable de le faire
HYGIENE				
Etes-vous capable de :				
- Vous laver et vous sécher entièrement?	|___|	|___|	|___|	|___|
- Prendre un bain ?	|___|	|___|	|___|	|___|
- Vous asseoir et vous relever des toilettes ?	|___|	|___|	|___|	|___|
ATTRAPER				
Etes-vous capable de :				
- Prendre un objet pesant 2,5 kg situé au-dessus de votre tête ?	|___|	|___|	|___|	|___|
- Vous baisser pour ramasser un vêtement par terre ?	|___|	|___|	|___|	|___|
PREHENSION				
Etes-vous capable de :				
- Ouvrir une porte de voiture ?	|___|	|___|	|___|	|___|
- Dévisser le couvercle d'un pot déjà ouvert une fois ?	|___|	|___|	|___|	|___|
- Ouvrir et fermer un robinet ?	|___|	|___|	|___|	|___|
AUTRES ACTIVITES				
Etes-vous capable de :				
- Faire vos courses ?	|___|	|___|	|___|	|___|
- Monter et descendre de voiture ?	|___|	|___|	|___|	|___|
- Faire des travaux ménagers tels que passer l'aspirateur ou faire du petit jardinage ?	|___|	|___|	|___|	|___|

Veuillez indiquer d'une croix si vous utilisez habituellement un de ces accessoires ou appareils pour effectuer ces activités :

Siège de WC surélevé	|___|	Poignée ou barre de baignoire	|___|
Siège de baignoire	|___|	Instrument à long manche pour attraper les objets	|___|
Ouvre-pots (pour les pots déjà ouverts)	|___|	Instrument à long manche dans la salle de bain	|___|

Autres (préciser) : .....

Veuillez indiquer les activités pour lesquelles vous avez besoin de l'aide de quelqu'un :

Hygiène	|___|	Saisir et ouvrir des objets	|___|
Atteindre et attraper	|___|	Courses et tâches ménagères	|___|

UNIVERSITE DE NANCY I
FACULTE DE MEDECINE

THESE DE DOCTORAT DE L'UNIVERSITE DE NANCY I

de **Monsieur GUILLEMIN Francis**

Mention : RECHERCHE EN SANTE PUBLIQUE

AUTORISE A IMPRIMER LA THESE

Nancy, le 24 DEC. 1993 n° 439

LE PRESIDENT DE L'UNIVERSITE DE NANCY I

UNIVERSITE NANCY I
LE PRESIDENT

Professeur M. BOULANGÉ

AUTEUR : Francis GUILLEMIN

TITRE : Mesure et déterminants de l'incapacité fonctionnelle dans la polyarthrite rhumatoïde récente. La cohorte EURIDISS.

THESE : Doctorat d'Université de Nancy I

ANNEE : 1993

MOTS-CLEFS : Polyarthrite rhumatoide - Incapacite fonctionnelle - Qualite de vie- Culture - Health Assessment Questionnaire - Facteurs de risque - Modèles d'analyse.

RESUME : L'objectif est d'améliorer la connaissance de l'incapacité fonctionnelle dans la polyarthrite rhumatoïde (PR) à son début en vue d'une meilleure prise en charge.

La première partie présente le développement d'une méthodologie d'adaptation culturelle des instruments de qualité de vie pour une mesure équivalente dans des cultures différentes. Les recommandations formulées sont appliquées à l'adaptation en français du HAQ, mesure de l'incapacité fonctionnelle (composante majeure de la qualité de vie dans les maladies rhumatismales), qui est ensuite validé dans la PR. Dans une deuxième partie, la comparaison d'un échantillon de PR récentes (durée inférieure à 5 ans) et anciennes (durée moyenne de 8 ans) met en evidence la spécificité de l'incapacité et de ses composantes dans la PR récente. Dans une troisième partie, le HAQ est utilisé dans une étude européenne de PR évoluant depuis moins de 5 ans, la cohorte EURIDISS. Les determinants identifiés par l'analyse transversale sont : durée de la maladie, indice de Ritchie, vitesse de sédimentation. Elle montre également le lien qui existe avec les événements de vie, liés à une incapacité plus élevée, et avec la densité du réseau social, liée à une incapacité plus modérée, suggérant l'existence d'un effet-tampon du réseau social. L'approche longitudinale préliminaire (après un an) rappelle l'hétérogénéité de l'incapacité et des facteurs qui la déterminent, montrant un faible changement de cette incapacité au cours du temps.

Cette connaissance des déterminants de l'incapacité appréciée par le patient est un outil précieux pour le clinicien et les autres acteurs de santé publique qui organisent la prise en charge du malade chronique dans son environnement médical et social.

JURY :

Pr M MANCIAUX	Président
Pr J-F VIEL, Besançon	Rapporteur
Pr W v den HEUVEL, Groningen (Pays-Bas)	Rapporteur
Pr S BRIANCON	Directeur de recherche
Pr J POUREL	Juge
Pr J LELLOUCH, Paris	Juge