# A comparison of utility measurements using EQ-5D and SF-6D preference-based generic instruments in patients with rheumatoid arthritis

F. Salaffi[1], M. Carotti[2], A. Ciapetti[1], S. Gasparini[1], W. Grassi[1]

*[1]Department of Molecular Pathology and Innovative Therapies, Division of Rheumatology, and [2]Department of Radiology, University of Ancona (Politecnica delle Marche), Jesi, Italy.*

## Abstract

### Objectives

*The purposes of this study were to analyse and compare aspects of validity (concurrent and discriminant) of the two widely used indirect utility instruments, the EuroQol-5D (EQ-5D) and the Short Form-6D (SF-6D) in a representative cohort of patients with rheumatoid arthritis (RA).*

### Methods

*Five hundred and eighty-three consecutive adult patients (435 women, 148 men) with RA and referred to the outpatient Clinic were evaluated. Patients were asked to complete EQ-5D and SF-36. SF-6D utility scores were calculated using the eight mean SF-36 scores, according to published algorithms. Disease-related characteristics included disease duration, co-morbidities, a measure for disease activity [Disease Activity Score-28 joint (DAS28)] and for radiographical damage (Sharp van der Heijde scoring method, SHS). The agreement between the utility instruments was evaluated by Bland-Altman analysis. Construct validity was assessed using the Kruskal-Wallis test, Mann-Whitney U-test, Spearman's correlations, and receiver operating characteristic (ROC) curves. Multivariate analyses were used to assess the relationship among HRQoL and disease-related characteristics and socio-demographic data.*

### Results

*A comparison of means showed that SF-6D values exceeded EQ-5D values ($p<0.0001$). Agreement between both measures was only moderate. Utility scores and domains and summary scores of the SF-36 were highly correlated. The EQ-5D and SF-6D both detected change in different health status (<0.0001). The discriminatory power of both indexes was good, without significant difference, with an AUC of 0.869 and 0.820, respectively for EQ-5D and SF-6D. The EQ-5D and SF-6D both detected change over different health status among RA patients (both al level of $p<0.0001$) although EQ-5D was more efficient in detecting differences between groups in almost all cases. Comparison of EQ-5D and SF-6D scores within VAS groups showed that, for less healthy individuals (VAS scores 0–50), the median EQ-5D score was significantly lower than the median SF-6D score. The multivariate regression models for EQ-5D and SF-6D included both SHS and DAS28 ($p=0.0001$). The relative contribution of these domains differed substantially between patients with short and long standing disease duration. The presence of multiple chronic conditions also appeared to contribute to reduce the levels of utility of both instruments.*

### Conclusion

*Although EQ-5D and SF-6D appeared to measure similar constructs, these instruments are quite different from each other in the assessment of RA. For worse health status the median EQ-5D scores were significantly lower than the median SF-6D scores. Moreover, EQ-5D and SF-6D appeared both significantly influenced by disease activity, radiological damage and co-morbidity. For that reason, we advise caution in the employment of these preference-based instruments, especially in RA patients with severe disease.*

**Key words**

rheumatoid arthritis, preference-based instruments, EQ-5D, SD-6D, health-related quality of life, cost-utility analyses

*Fausto Salaffi, MD, PhD*
*Marina Carotti, MD*
*Alessandro Ciapetti, MD*
*Stefania Gasparini, MD*
*Walter Grassi, MD*

*Please address correspondence and reprint requests to:*
*Fausto Salaffi, MD, PhD*
*Clinica Reumatologica*
*Università Politecnica delle Marche,*
*c/o Ospedale "C. Urbani",*
*Via dei Colli 52,*
*60035 Jesi (Ancona), Italy.*
*E-mail: fsalaff@tin.it*

*Received on December 7, 2010; accepted in revised form on April 7, 2011.*



*Competing interests: This study has been sponsored by UCB, Italy, but UCB has not influenced the manuscript content.*
*F. Salaffi has received a research grant from the Italian Ministry of Health, and for consultancy/speaker honoraria from Pfizer, Roche, Bristol-Myers Squibb, and Abbott;*
*A. Ciapetti is a member of the speakers' bureau of Bristol-Myers Squibb and Roche;*
*S. Gasparini is a member of the speakers' bureau of Roche;*
*W. Grassi has received consultancy/speaker honoraria from Pfizer, Roche, Bristol-Myers Squibb, Abbott, UCB Italy, Merck Sharp & Dohme, and Schering-Plough;*
*M. Carotti has declared no competing interests.*

## Introduction

Rheumatoid arthritis (RA) is a chronic disabling inflammatory joint disease affecting about 0.5% of the population (1) that is primarily characterised by persistent and progressive synovitis. Similarly to other chronic diseases, RA can affect quality of life, morbidity, mortality, ability to do paid or unpaid work and healthcare costs (2-4). Considering the functional and psychosocial impact of the disease, to obtain a holistic view of the health-related quality of life (HRQoL), study endpoints can be grouped into the following categories based on the source of the information: clinician-reported outcomes (CROs), that include outcomes either observed by a provider (*e.g.* joint count) or requiring interpretation (*e.g.* radiologic results, blood chemistry) and rheumatology-specific patient-reported outcomes (PROs) (5). The validity and usefulness of PRO data in evaluating and monitoring patients with RA have been well documented (6, 7).

Consideration of HRQoL has become progressively more important in decisions regarding resource allocation, intervention design and pharmacological treatment with biologic agents of individuals with chronic inflammatory disabling conditions (8). A review (9) have reported the usefulness of the quality-adjusted life year (QALY), based on patient measurements of HRQoL identified preference-based instruments such as the EuroQol (EQ-5D) and the most recent Short-Form-6D (SF-6D) (10, 11). The EQ-5D is a generic preference-based instrument that provides a comprehensive framework within which to determine health status and measure HRQoL. The SF-6D was calculated from SF-36 by using a definite scoring function (12). This algorithm has the potential to extend the scope for undertaking economic evaluations in health care using exiting and future publications of SF-36 data sets.

To date, in Italy, the use of these instruments is still limited. In two studies, the EQ-5D was found to be applicable and adaptable to the Italy environment (13, 14). However, its construct validity was demonstrated only recently using a large sample of the Italian general population (15). The SF-36 have been translated into Italian (16) and its reliability and validity were established in patients with rheumatic conditions (2, 17, 19). Over the years, utility became accepted as the theoretically most defendable approach for weighting life-years to calculate QALY (20, 21). Preference-based instruments are indirect, individual or combined, valuation methods, whereas other instruments are direct valuation techniques (*i.e.* standard gamble (SG), time trade-off (TTO), visual analogue scale (VAS), or rating scale (21). Utilities obtained by indirect methods are recommended by the US Panel on Cost-Effectiveness in Health and Medicine and the Outcome Measures in Rheumatology Clinical Trials (OMERACT) Consensus-Based Reference Case for Economic Evaluation in Rheumatoid Arthritis (22). Their ease of administration has contributed to their increased use as a source of quality weightings in economic evaluations and in clinical trials. In spite of agencies, such as National Institute for Clinical Excellence (NICE), recommends using choice-based measures in technology appraisals (23), a widely accepted key concern is that utility scores can vary according to the choice of instrument. In fact, many researchers found significant differences in global utility scores obtained by different instruments (24).

The purpose of this study were to contribute to the ongoing discussion on the choice of instrument in cost-utility analyses (CUA) by studying and comparing psycometric properties of the two widely used indirect utility instruments, the EQ-5D and the SF-6D, in a representative cohort of patients with RA and to define the variables that can influence the utility measurement of their health status.

## Patients and methods

*Study population*

Five hundred and eighty-three consecutive adult patients (435 women, 148 men) fulfilling the American College of Rheumatology (ACR) classification criteria for RA (25) and referred to the Rheumatology Department of the Università Politecnica delle Marche

(Ancona, Italy) were evaluated by two rheumatologists. All patients gave their informed consent for anonymous analysis of data.

*Demographics, disease-related characteristics and radiographic evaluation*

A comprehensive questionnaire package including socio-demographic data, quality of life items and disease-related variables was administered to the patients. The following socio-demographic variables were included: age, gender and highest attained level of education (primary and secondary school, university). The presence of the following comorbidities was assessed: hypertension, myocardial infarction, lower extremity arterial disease, major neurological problems, diabetes, gastrointestinal disease, chronic respiratory disease, kidney disease, and poor vision. The Disease Activity Score 28-joint count (DAS28) was used to evaluate disease activity (26). The disease activity was interpreted as low (DAS28 ≤3.2), moderate (3.2 <DAS28 ≤5.1) or as high (DAS28 >5.1), according to the European League Against Rheumatism (EULAR) criteria (27-29) and in remission (DAS28 <2.6) according to the OMERACT criteria (30). Radiographical damage was assessed, by a musculoskeletal radiologist (MC) who was unaware of patient identity, according to Sharp's method as modified by the Sharp-van der Heijde Score (SHS) (31). Interobserver agreement was tested also by a second investigator (FS) on 42 consecutive pairs of radiographs and the intraclass correlation coefficient (ICC) between the two investigators was more than 0.89. By then, considering the high interobserver agreements between the two readers, the radiographs have been read by only one reader (MC).

*Health-related quality of life assessment*

By completing the EQ-5D and SF-36 questionnaires three measures of HR-QoL can be estimated: EQ-5D, EQ-VAS and SF-6D indices. Utility scores are provided by the EQ-5D and SF-6D, whereas the EQ-VAS summarises HR-QoL on a 0–100 scale.

*EQ-5D index*

The EQ-5D explores the following five domains: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each domain has one item and each item has three levels: one denoting no problems and three denoting severe problems (32). A utility index score was calculated for each subject's EQ-5D health status by applying the time trade-off-based valuations from a general UK population sample to the observed EQ-5D profile, as data from an Italian norm are not available at the present time. The EQ-5D utility scores range from -0.59 (0=being dead; negative values represent health status considered worse than "dead") to 1.00 (good health status). Values close to zero indicate worse conditions and 1.00 represents perfect health status. In addition, patients were asked to rate their current health status on a vertical, graduated 20-cm visual analogue scale (EQ-5D VAS) ranging from 0 (worst possible health state) to 100 (best possible health status).

*SF-6D index*

The SF-6D is derived from the standard 4-week recall validated Italian translation of the self-administered SF-36 (IQOLA SF-36 Italian Version 1.6) (16). The SF-6D is focused on six of the eight health domains covered by the SF-36 Health Survey: physical functioning, role participation (combined role-physical and role-emotional), social functioning, bodily pain, mental health, and vitality. The SF-36 global measure of physical (PCS) and mental functioning (MCS) were calculated as well as descriptive intent (33). The SF-6D was calculated from SF-36 by using a definite scoring function (12) in order to create a weighted index score ranging from 1.0 [no difficulty in any dimensions (or perfect health)] to 0.296 (severely impaired levels in all dimensions).

*Statistical analysis*

The statistical analyses were performed using the SPSS version 15.0 (SPSS Inc, Chicago, USA), and the MedCalc® version 10.0 (MedCalc Software, Mariakerke, Belgium). Continuous data were presented as means with standard deviations (SDs) or medians with 95% confidence interval (95% CI), depending on the distribution of the data (tested with the Kolmogorov-Smirnov test). Categorical data were presented as proportions. Demographic and clinical measures were compared using Mann-Whitney U-test or Kruskal-Wallis test for continuous variables, and chi-square analysis for discontinuous variables. *P*-values below 0.05 were regarded as statistically significant.

*Concurrent validity*

To measure convergent validity a corresponding new instrument is usually compared with an established questionnaire. To define the convergent validity of the EQ-5D health status classifier, corresponding dimensions of the EQ-5D and domains of the SF-36 were compared using Spearman's rho correlation. A correlation from 0.00 to 0.20 was interpreted as no correlation; 0.21–0.40 as low correlation; 0.41–0.60 as moderate correlation; 0.61–0.80 as marked correlation; and 0.81–1.00 as high correlation. Since the SF-6D is based on the same measure of the SF-36 and would naturally be expected to correlate with the domains of that measure, tests of convergent and divergent validity were performed only for the EQ-5D. Agreement was assessed by the Bland-Altman plot (34).

*Discriminant validity*

Discriminant validity evaluates whether a measure is able to identify differences between patient subgroups stratified based on an external anchor of health (35). The second component of the EQ-5D, the EQ-VAS was used to classify individuals into health status groups, covering the range from very poor to very good health status (36). Specifically, each subject was included in one of eight groups according to EQ-VAS score: 0–20, 21–30, 31–40, 41–50, 51–60, 61–70, 71–80, and 81–100, with each group containing approximately a comparable number of respondents. To test the validity of the EQ-VAS as a discriminator of health status, the eight groups were used in the RA patients to assess the EQ-5D and SF-6D mean differences. Mann-Whitney U-tests were

used to compare differences associated with health status groups for the EQ-5D and SF-6D. In addition, the relationship between the econometric indexes and different levels of activity scores in individual patients were evaluated. Therefore, we have created 4 patient groups based on the patients' DAS28 ranks within the cohort. The patients were grouped in the same way based on their EQ-5D and SF-6D scores and by using non parametric Kruskal-Wallis test to assess the level of significance of different disease activity categories on individual patients. Finally, the discriminative properties of both EQ-5D and SF-6D were compared using receiver operating characteristic (ROC) curves (37). In this analysis, the performance of the EQ-5D index and the SF-6D index was evaluated against two external indicators of disease activity status: "low disease activity", and "high disease activity". The DAS28 EULAR response criteria were applied as external criterion (26). Since ROC analysis requires external criteria to be dichotomous, remission and low activity were grouped together as "overall" "low disease activity", whereas moderate and high activity were clustered in "high disease activity". The area under the ROC curve (AUC) was calculated to quantify the discriminative accuracy. The non-parametric Wilcoxon's signed ranks test is used for calculation and comparison of the areas under the ROC curves, as suggested by Hanley and McNeil (38).

*Mutivariate analysis*

A set of multivariable analyses were constructed to adjust for factors potentially associated with poor HRQoL in the RA patients. Covariates chosen by *a priori* analysis were the following: gender (as a dichotomous variable; 0=male; 1=female); age (as a continuous variable); educational level (years of education as a continuous variable); number of co-morbidities (as a continuous variable), DAS28 and SHS scores (both as a continuous variable). All these factors were then introduced as covariates in multiple regression models in which EQ-5D and SF-6D scores were dependent variables. All variables were entered simultaneously. Because of a relevant statistical interaction between disease duration and radiographic damage, a sub-analysis was performed for patients with low (≤3 years) *versus* high (>3 years) disease duration. The level of statistical significance was set at 0.01 in order to reduce increasing risk of reporting errors due to multiple comparisons.

## Results

*Demographic and clinical data*

The majority of the sample were women (74.6%), married or co-habiting (62.5%) and with low educational level (73.5%). The respondents' age ranged from 19 to 85 years, with a mean of 58 years (SD=14.2 years). They were most frequently retired or manual workers and living in urban areas. Of the 583 subjects enrolled, 313 (53.7%) reported one or more medical co-morbidities, mostly cardiovascular (33.2%), respiratory (15.5%), and metabolic (14.9%) disorders. The large majority was classified as having moderate (284 subjects, 48.7%) or high (186 subjects, 31.9%) disease activity. The proportion of patients in low disease activity or remission were 15.4 (90 subjects) and 4.0% (23 subjects), respectively. Table I summarises the scores for the EQ-5D, EQ-VAS, SF-6D and PCS and MCS summary component of the SF-36 and disease activity (DAS28) of all patients with RA.

*Score distributions and agreement between utility measures*

Figure 1 presents estimates of central tendency and distribution of score for EQ-5D (Fig. 1A), SF-6D (Fig. 1B), and scatter plots of EQ-5D with SF-6D (Fig. 1C) in the RA patient cohort. EQ-5D and SF-6D values were not-normally distributed (Kolmogorov-Smirnov test) and the median baseline values have different locations in their respective scoring ranges (Fig. 1 A-B). Comparisons of means showed that SF-6D values exceeded EQ-5D values. The mean difference is 0.096 [95% confidence interval (CI) of difference 0.084 to 0.108, $p<0.0001$]. There was a marked correlation between the 2 utility measures (Spearman's coefficient of rank correlation, rho=0.619, $p<0.001$) (Fig. 1C). Figure 2 reported the percentage of subjects with RA reporting problems, by EQ-5D dimension. To note that the more frequently severe problems referred by the patients were related to the following domains: pain and disability (44.3%), anxiety and depression (29.7%) and inability to do usual activities (14.4%). Subjects who did not report any problems represent the minority, ranging between 1% (related to the pain and disability domains) and 34.8% (related to the self-care domain). Agreement between both meas-

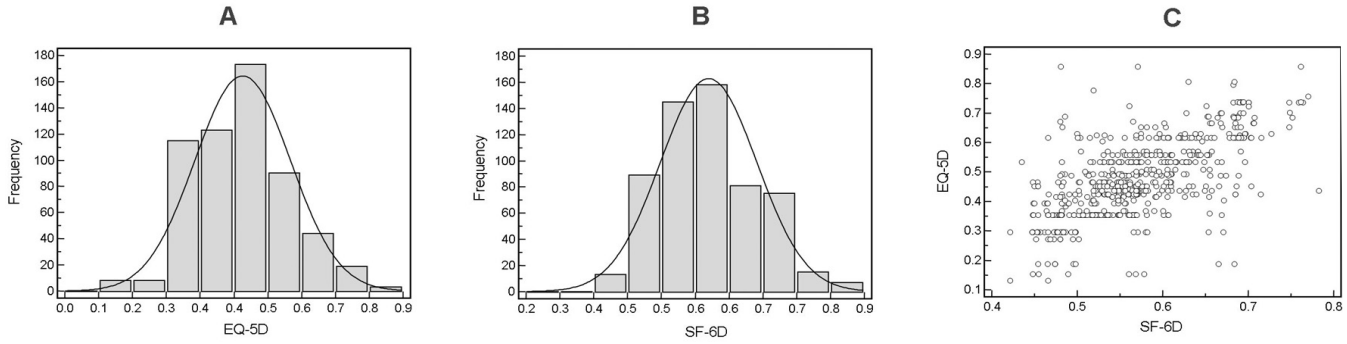



**Fig. 1.** Distribution of EQ-5D (A), SF-6D (B) and scatter plots (C) of EQ-5D with SF-6D.

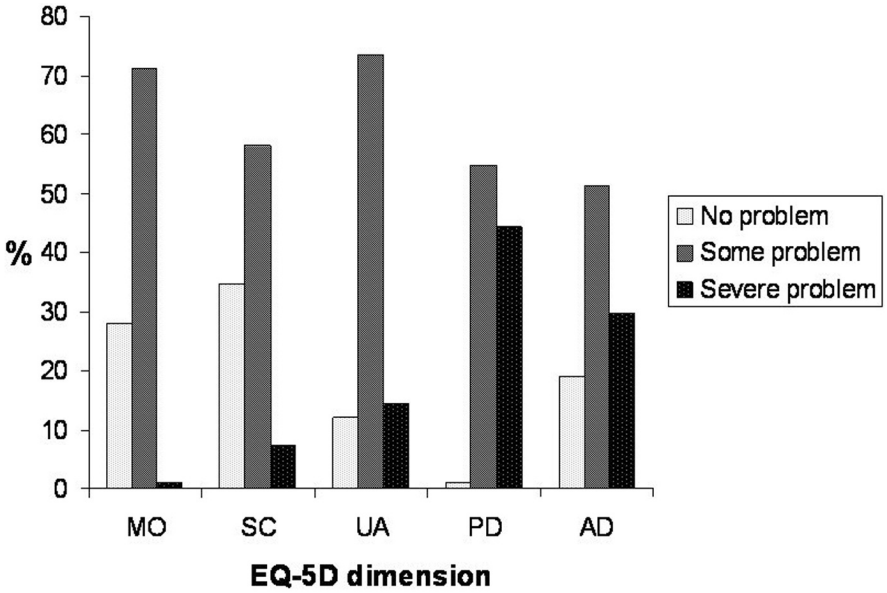



**Fig. 2.** Percentage of subjects with rheumatoid arthritis reporting problems by EQ-5D dimension. MO: mobility; SC: self-care; UA: usual activity; PD: pain and disability; AD: anxiety and depression.

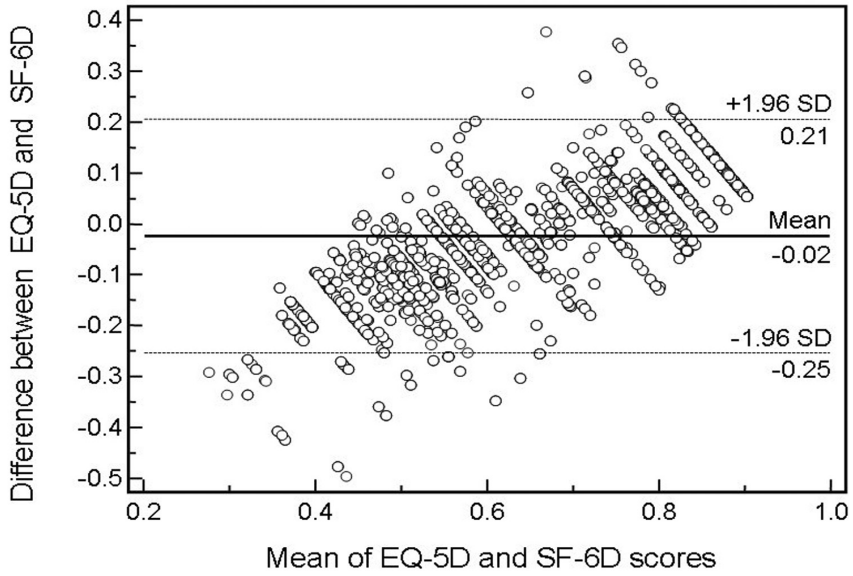



**Fig. 3.** Bland and Altman plot of differences between EQ-5D and SF-6D for patients with rheumatoid arthritis.

ures was moderate. The Bland-Altman plot showed proportional error and wide limits of agreement (Fig. 3).

*Convergent validity*

EQ-5D utility scores and domains and summary scores of the SF-36 were all moderately correlated (Table I). With respect to the socio-demographic data both the EQ-5D and SF-6D were inversely correlated to the age and disease duration, but they did not showed any significant correlation with the educational level. The disease-related characteristics (number of co-morbidities, DAS28 scores and radiographic damage score) have been resulted to correlate with both the utility measures. Moderate correlations were also found between both EQ-5D and SF-6D and DAS28 score (rho=-0.478 and -0.440, respectively) (Table II).

*Discriminant validity*

To compare the mean EQ-5D and SF-6D scores across groups of differing health, the subjects were categorised into eight health status groups on the basis of their responses to the EQ-VAS. The EQ-5D and SF-6D both detected change over different health status among RA patients (Kruskal-Wallis test, Ht=178.42 and 79.63, respectively; both al level of $p<0.0001$) although EQ-5D was more efficient in detecting differences between groups in almost all cases (Fig. 4). This validity exercise provided evidence that the consecutive EQ-VAS based groups did differ in health status and simulated the full range of health. Comparison of EQ-5D and SF-6D scores within VAS groups showed that, for less healthy individuals (VAS scores 0–50), the median EQ-5D score was significantly lower than the median SF-6D score. The patients with VAS scores over than 50 had higher score for the SF-6D as well. The "crossover" appeared to occur somewhere (VAS scores 61-70) in which the difference between the EQ-5D and SF-6D mean scores was minimal (Fig. 4). From that point and on, as health status improved (higher VAS scores), the SF-6D scores were higher (but not significant) than EQ-5D scores. On categorising patients into those with remission, low activity, moderate activity and high activity, with respect to the DAS28, EQ-5D and SF-6D were highly significantly different and showed a similar magnitude for the four categories (Kruskal-Wallis test, Ht=146.37 and 114.95, respectively; both at a level of $p<0.0001$) (Fig. 5A, 5B). Figure 6 shows the AUC-ROC for the EQ-5D and SF-6D indexes when detecting differences in disease activity. The discriminatory power of EQ-5D and SF-6D was good, without significant difference, with an AUC of 0.869 (95% CI 0.839±0.896), and 0.820 (95% CI 0.786±0.850), respectively (differences between areas =0.049±0.028 with 95% C.I. from 0.005 to 0.105; $p=0.079$) (Fig. 6).

*Factors associated with poor health-related quality of life*

Multiple regression models were constructed to adjust for factors potentially associated with poor HRQL measured by the two utility measures. The EQ-5D was influenced by a high disease activity, followed by radiographic damage (both at a $p$-level of <0.0001), and

**Table I.** Summary scores for the preference-based instruments (EQ-5D, EQ VAS and SF-6D), physical (PCS) and mental functioning (MCS) summary component of the SF-36, and disease activity (DAS28) of all patients with rheumatoid arthritis (n=583).

| | EQ-5D (score 0.59–1) | SF-6D (score 0.296–1) | EQ-5D VAS (score 0–100) | SF-36 PCS (score 0–100) | SF-36 MCS (score 0–100) | DAS28 (score 0–9.4) |
|---|---|---|---|---|---|---|
| Lowest value | 0.1305 | 0.14214 | 13.0000 | 15.6300 | 16.7600 | 2.1500 |
| Highest value | 0.8584 | 0.7829 | 90.0000 | 52.3300 | 76.7100 | 8.1800 |
| Arithmetic mean | 0.4734 | 0.5699 | 53.6827 | 31.3920 | 42.5589 | 4.6476 |
| 95% CI for the mean | 0.4622 – 0.4833 | 0.5640 – 0.5757 | 52.4487 – 54.9166 | 30.8891 – 31.8949 | 41.5772 – 43.5405 | 4.5626 – 4.7326 |
| Median | 0.4650 | 0.5618 | 55.0000 | 30.8100 | 41.2600 | 4.6300 |
| 95% CI for the median | 0.4513 – 0.4651 | 0.5559 – 0.5691 | 55.0000 – 55.0000 | 30.4370 – 31.4400 | 39.2987 – 42.1600 | 4.5300 – 4.7296 |
| Variance | 0.0162 | 0.0051 | 230.1277 | 38.2250 | 145.6399 | 1.0915 |
| Standard deviation | 0.1279 | 0.0714 | 15.1700 | 6.1826 | 12.0681 | 1.0448 |
| Relative standard deviation | 0.269 (26.91%) | 0.1254 (12.54%) | 0.2826 (28.26%) | 0.1969 (19.69%) | 0.2836 (28.36%) | 0.2248 (22.48%) |
| Mean standard error | 0.0052 | 0.0029 | 0.6283 | 0.2561 | 0.4998 | 0.0432 |
| Coefficient of Skewness | 0.1783 ($p$=0.0782) | 0.4408 ($p$<0.0001) | -0.5132 ($p$<0.0001) | 0.6130 ($p$<0.0001) | 0.4311 ($p$<0.0001) | 0.3628 ($p$=0.0005) |
| Coefficient of Kurtosis | 0.1277 ($p$=0.4838) | -0.3062 ($p$=0.1500) | 0.0048 ($p$=0.9061) | 0.8208 ($p$=0.0020) | -0.5509 ($p$=0.0214) | 0.3905 ($p$=0.0795) |
| Kolmogorov-Smirnov test for Normal distribution | reject Normality ($p$=0.001) | reject Normality ($p$=0.001) | reject Normality ($p$<0.001) | reject Normality ($p$=0.023) | reject Normality ($p$=0.006) | accept Normality ($p$=0.236) |

**Table II.** Spearman rank correlation between EuroQol-5D (EQ-5D) and Short Form-6D (SF-6D) and socio-demographic data and disease-related characteristics.

| | | Age | Disease duration | Educational level | Number of comorbidities | DAS28 score | Sharp score (SHS) |
|---|---|---|---|---|---|---|---|
| EQ-5D | Correlation Coefficient | -0.089 | -0.123 | 0.026 | -0.161 | -0.478 | -0.220 |
| | Significance Level p | 0.031 | 0.003 | 0.536 | <0.001 | <0.001 | <0.001 |
| SF-6D | Correlation Coefficient | -0.146 | -0.120 | 0.024 | -0.144 | -0.440 | -0.205 |
| | Significance Level p | <0.001 | 0.004 | 0.556 | <0.001 | <0.001 | <0.001 |
| Age | Correlation Coefficient | | -0.015 | -0.053 | 0.157 | 0.059 | -0.031 |
| | Significance Level p | | 0.723 | 0.204 | <0.001 | 0.155 | 0.450 |
| Disease duration | Correlation Coefficient | | | -0.037 | 0.046 | 0.009 | 0.717 |
| | Significance Level p | | | 0.372 | 0.270 | 0.829 | <0.001 |
| Educational level | Correlation Coefficient | | | | -0.034 | 0.053 | 0.032 |
| | Significance Level p | | | | 0.415 | 0.199 | 0.437 |
| Number of co-morbidities | Correlation Coefficient | | | | | 0.088 | 0.044 |
| | Significance Level p | | | | | 0.034 | 0.293 |
| DAS28 score | Correlation Coefficient | | | | | | 0.084 |
| | Significance Level p | | | | | | 0.042 |

number of co-morbidities conditions ($p$=0.0013) (Table III). A similar association between high disease activity, severe radiographic damage, chronic co-morbidities and SF-6D were also found (Table IV). Because of a relevant statistical interaction between disease duration and radiographic damage, a sub-analysis was performed for patients with low (≤3 years) *versus* high (>3 years) disease duration. Concerning the EQ-5D, an association was found with the disease activity among both patients with early (≤3 years) and late RA onset (>3 years) (both at a $p$-level of $p$<0.0001). The radiographical damage appeared to be an influential variable only in the group of patients with long standing disease (t=-5.294; $p$<0.0001). The number of co-morbidities showed a moderate correlation with poor HRQoL among both groups of patients stratified by disease duration (Table III). Similar results have been obtained by using SF-6D index as dependent variable in the multivariate statistical analysis (Table IV).

## Discussion

The diverging scores between utility measures in patients with RA and the consequent impact on the results of cost-utility analyses have been well documented (39, 40). The differences between utility instruments seems due to the selected instrument rather than differences in treatment efficacy and may lead to different utility values and, therefore, to different resource allocation decisions (41, 42).

The original contribution of this study is that for the first time the discriminative features of the SF-6D have been compared to the EQ-5D index in an Italian population with RA. To our knowledge, no other study has published utility values derived from clinically diagnosed cases of RA drawn from Italy, based upon both reference case measures, the EQ-5D and SF-6D. Moreover, no study so far has directly compared the performance of these two measures in an Italian population receiving care for RA and it remains unclear if these two "reference case" measures are sensitive to change in the underlying activity and severity disease.

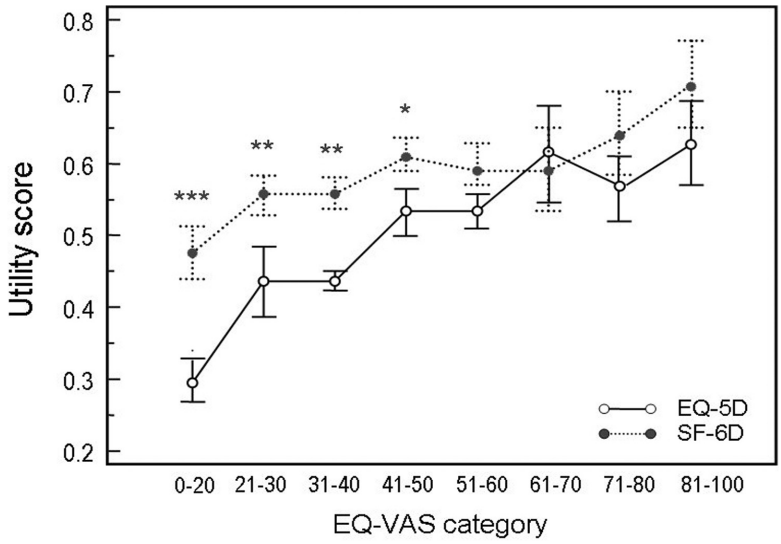



**Fig. 4.** Comparison of EQ-5D and SF-6D utilities across groups of differing health by EQ-VAS in RA patients. $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$ according to Mann-Whitney U-tests.

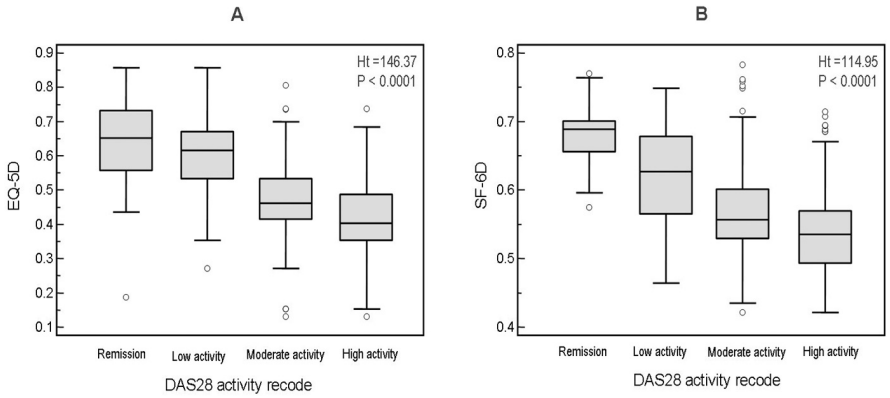
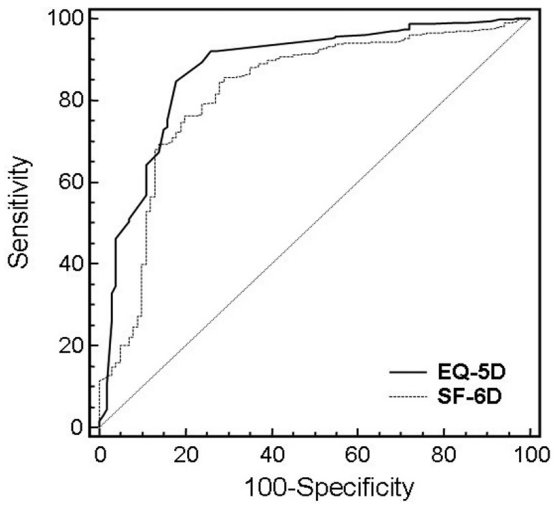



**Fig. 5.** The box-plots present the median, quartiles and extreme values for the EQ-5D (A) and SF-6D (B) utility scores for each DAS28 activity state.




**Pairwise comparison of ROC curves**

| | AUC | SE | 95% CI |
|---|---|---|---|
| EQ-5D | 0.869 | 0.0236 | 0.839 to 0.896 |
| SF-6D | 0.820 | 0.0267 | 0.786 to 0.850 |

| EQ-5D *vs* SF-6D | |
|---|---|
| Difference between areas | 0.0496 |
| Standard Error | 0.0282 |
| 95% Confidence Interval | -0.005 to 0.105 |
| z statistic | 1.758 |
| Significance level | P = 0.079 |

**Fig. 6.** ROC curves illustrating the relationship between sensitivity and complement of specificity (100-specificity) in rheumatoid arthritis for the EQ-5D and SF-6D using disease activity as an external indicator. The area under the ROC curve (AUC) in this setting can be interpreted as the probability of correctly identifying the ''low disease activity'', from the ''high disease activity''. A line that runs diagonally across the figure from lower left to upper right will have an area of 0.5; this represents an instrument that does not discriminate.

For this purpose, we compare the performance of the instruments in socio-demographic and disease health status groups, in a representative sample of patients with RA. RA, which was previously shown to strongly affect the Italian population (2, 17, 18), was chosen on the basis of the fact that impaired HRQoL is associated not only with the disease itself, but largely with the existing co-morbidities conditions (1, 2). The results of this study indicate substantial discrepancies between instruments. Although the EQ-5D and SF-6D both detected changes over different health status and disease activity groups, the data demonstrate disparities between these instruments. For worse health status the SF-6D provides a value that is consistently higher than the EQ-5D result whilst, in healthier status the SF-6D and EQ-5D shows similar scores. This discrepancy, as expected, is due to the fact that the SF-6D scoring algorithm not assigns a value close to or below zero in any health status condition (the most severe health state defined by SF-6D has a utility score of 0.296). Other explanations might include the different evaluation systems used for each instrument. EQ-5D health status was evaluated by using the TTO method, whereas SF-6D was derived from the SF-36 by using a definite scoring algorithm (10-12). Brazier *et al*. (43) have suggested adding more intermediate levels to the EQ-5D or lower levels to the SF-6D, at least for the physical functioning and role limitations.

A comparison between the the EQ-5D and the SF-6D have been widely reported in literature (40, 44, 45). Jorstad and collegues (46) reported that despite a good relative correlation between different utilty values obtained from four indirect questionnaires (15-D, EQ-5D, SF-6D and EQ-VAS) administered to a RA population, the utility values appeared significantly different for the same health status. Moreover, the authors underlined the fact that when these differences are incorporated in a cost utility analysis they can lead to diverging results and, therefore, have consequences in terms of potential reimbursement decisions related to RA

**Table III.** Regression models with the EQ-5D (utility values) as dependent variable. Results are related to the entire RA population and disease duration.

| Independent variables | Coefficient | Std. Error | t | *p*-value |
|---|---|---|---|---|
| *Entire rheumatoid arthritis population (n=583)* | | | | |
| (Constant) | 0.7728 | | | |
| Age, years | -0.0003135 | 0.0003128 | -1.002 | 0.3167 |
| Comorbidity | -0.009550 | 0.002953 | -3.234 | 0.0013 |
| Educational level | 0.001423 | 0.0007761 | 1.834 | 0.0672 |
| Gender | 0.009330 | 0.007553 | 1.235 | 0.2172 |
| DAS28 | -0.05481 | 0.003904 | -14.038 | <0.0001 |
| Sharp SHS | -0.0005842 | 0.0001122 | -5.207 | <0.0001 |
| F-ratio | | | | 44.5544 |
| Significance level | | | | $p<0.001$ |
| *Disease duration ≤3 years (n=198)* | | | | |
| (Constant) | 0.7407 | | | |
| Age, years | -0.00003245 | 0.0006416 | -0.0506 | 0.9597 |
| Comorbidity | -0.01322 | 0.005785 | -2.286 | 0.0235 |
| Educational level | 0.0009165 | 0.002266 | 0.404 | 0.6864 |
| Gender | 0.03687 | 0.01935 | 1.906 | 0.0583 |
| DAS28 | -0.05876 | 0.007144 | -8.225 | <0.0001 |
| Sharp SHS | -0.0001185 | 0.0003799 | -0.312 | 0.7554 |
| F-ratio | | | | 13.4118 |
| Significance level | | | | $p<0.001$ |
| *Disease duration >3 years (n=385)* | | | | |
| (Constant) | 0.7750 | | | |
| Age, years | -0.0003046 | 0.0003647 | -0.835 | 0.4043 |
| Comorbidity | -0.008828 | 0.003496 | -2.525 | 0.0120 |
| Educational level | 0.001659 | 0.0008237 | 2.014 | 0.0447 |
| Gender | 0.004704 | 0.008205 | 0.573 | 0.5668 |
| DAS28 | -0.05286 | 0.004835 | -10.935 | <0.0001 |
| Sharp SHS | -0.0007284 | 0.0001376 | -5.294 | <0.0001 |
| F-ratio | | | | 29.7931 |
| Significance level | | | | $p<0.001$ |

treatments. Kontodimopoulos *et al*. (36) confirmed the hypotheses that EQ-5D generates higher scores than the SF-6D in healthier subjects. Lillegraven *et al*. (40) have observed conflicting scores of these utility instruments, especially in RA patients with higher Health Assessment Questionnaire (HAQ) scores. In particular, lower utility scores were found for EQ-5D in patients with severe disabilities, and these differences are reproduced when comparing mean utility scores across groups, according the rate of overall health (first item of SF-36). Moreover, Barton *et al*. (35) have shown that the mean deficit of HRQoL, collected from UK general practitioner cross-sectional survey, associated with many clinical conditions, was estimated to be higher according to the EQ-5D than the SF-6D. Brazier *et al*. (43) have found overall similar results in patients with mild diseases, but other authors (46, 47) showed that at baseline values were clearly different in patients with severe co-morbidities conditions (*e.g*. liver transplantation or stroke). Marra *et al*. (48), as well and similarly to our results, have reported a moderate correlations between SF-6D and PRO measures.

Overall, these studies confirmed some of the discrepancies we found in our study, mainly due to different descriptive content and adopted range of scoring (49) and do not support the construct validity. In our study, multiple regression models were constructed to adjust for factors potentially associated with poor HRQoL measured by the two utility measures. Both the instruments were influenced by disease activity, radiographic joint damage and comorbidity in decreasing order of strength, but not by age, gender and educational level. Because of a relevant statistical interaction between disease duration and radiographic damage, a sub-analysis was performed for patients with low (≤3 years) *versus* high (>3 years) disease duration. The radiographical damage appeared to be an influential variable only in the group of patients with long standing disease. There are many studies demonstrating the relation between the general physical status of the RA patients and the radiological damage or disease activity (50-53). These studies have been performed in an attempt to correlate commonly used clinical indices, such as the HAQ, grip strength, and the Ritchie index, with radiographic findings. The results had suggested that functional capacity was influenced largely by disease activity in early RA and by joint destruction in established RA (50-53). In a cohort of active early RA patients, Knijff-Dutmer and Cohen *et al*. (54) found a linear relationship between time integrated disease activity parameters and progression of radiological damage. Similar results were reported by Molenaar *et al*. (55) and Welsing *et al*. (56). However, literature is lacking in studies investigating the relationship of the utility of the health status with disease activity and radiological damage. The implication of our findings is that, as in active RA, the goal of treatment in patients with low or inactive RA should be to both suppress joint inflammation to the lowest level possible and to retard radiographic progression, in order to maintain functional capacity. Several co-morbidities were found in our patient's group. In particular, 53.7% of RA patients reported at least one comorbidity condition. Similar results have been described by Rupp *et al*. (56%) (57) Berkanovic *et al*. (54%) (58) and Gabriel *et al*. (49.3%) (59). Potential additional effect of co-morbidities on health outcomes in RA have been the subject of several studies. It is increasingly known that RA-related co-morbidities, including cardiovascular disease, infection, osteoporosis, lymphoproliferative malignancy and peptic ulcer disease, serve as major determinants of disease-associated outcome (60). As opposed to a linear relationship between arthritis-related co-morbidity and disability development, the frequency of disability rises exponentially

**Table IV.** Regression models with the SF-6D (utility values) as dependent variable. Results are shown for the entire RA population and according to disease duration.

| Independent variables | Coefficient | Std. Error | t | *p*-value |
|---|---|---|---|---|
| *Entire rheumatoid arthritis population (n=583)* | | | | |
| (Constant) | 0.7394 | | | |
| Age, years | -0.0003958 | 0.0001821 | -1.974 | 0.0701 |
| Comorbidity | -0.004137 | 0.001719 | -2.407 | 0.0164 |
| Educational level | 0.0006107 | 0.0004518 | 1.352 | 0.1770 |
| Gender | 0.001734 | 0.004397 | 0.394 | 0.6934 |
| DAS28 | -0.02793 | 0.002273 | -12.288 | <0.0001 |
| Sharp SHS | -0.0003112 | 0.00006532 | -4.764 | <0.0001 |
| F-ratio | | | | 34.6440 |
| Significance level | | | | $p$<0.001 |
| *Disease duration ≤3 years (n=198)* | | | | |
| (Constant) | 0.7433 | | | |
| Age, years | -0.0002735 | 0.0003676 | -0.744 | 0.4579 |
| Comorbidity | -0.006570 | 0.003314 | -1.982 | 0.0491 |
| Educational level | -0.001353 | 0.001298 | -1.042 | 0.2990 |
| Gender | 0.01645 | 0.01108 | 1.484 | 0.1396 |
| DAS28 | -0.03011 | 0.004093 | -7.357 | <0.0001 |
| Sharp SHS | -0.00008613 | 0.0002176 | -0.396 | 0.6928 |
| F-ratio | | | | 11.2545 |
| Significance level | | | | $p$<0.001 |
| *Disease duration >3 years (n=385)* | | | | |
| (Constant) | 0.7359 | | | |
| Age, years | -0.0003480 | 0.0002139 | -1.627 | 0.1046 |
| Comorbidity | -0.004046 | 0.002050 | -1.974 | 0.0492 |
| Educational level | 0.0009286 | 0.0004830 | 1.922 | 0.0553 |
| Gender | -0.0009529 | 0.004812 | -0.198 | 0.8431 |
| DAS28 | -0.02665 | 0.002835 | -9.400 | <0.0001 |
| Sharp SHS | -0.0003856 | 0.00008067 | -4.780 | <0.0001 |
| F-ratio | | | | 22.7684 |
| Significance level | | | | $p$<0.001 |

with the number of co-morbidities (58, 61-66).

The strength of our study relies on the use of a large sample of RA patients usually evaluated in rheumatology practices and not on the typical trial data which may be misleading, given the typical exclusion criteria (*e.g.* due to co-morbidities). However, our research presents several limitations as well. Firstly, because of the nature of the sample, the results are not generalisable beyond RA patients being treated in rheumatology practices. The second limitation is related to the cross-sectional study design which does not allow test-retest reliability evaluation and does not provide information on the sensitivity to change after treatment.

In conclusion, on the basis of our data the EQ-5D and SF-6D in the assessment of RA appeared not equivalent despite some similarities. These instruments both detected changes over different health status among RA patients, but for worse health status the median EQ-5D scores were significantly lower than the median SF-6D scores. Moreover, EQ-5D and SF-6D appeared both significantly influenced by disease activity, radiological damage and co-morbidity. These properties should be carefully considered in planning utility assessments in RA patients, particularly in those studies assessing patients and impact of treatment in severe disease. Although there is not yet a definitive utility measure and we advice caution in the employment of these instruments in RA assessment, at present, the EQ-5D remain the utility measure recommended by the OMERACT (22) and NICE (23) for he economic evaluation in RA. Considering the ongoing controversial debate (64), further deeper studies focused on the psychometric properties of the EQ-5D and SF-6D in clinical or resource allocation decision-making are needed.

## References


1. SALAFFI F, DE ANGELIS R, GRASSI W: MArche Pain Prevalence; INvestigation Group (MAPPING) study. Prevalence of musculoskeletal conditions in an Italian population sample: results of a regional community-based study. I. The MAPPING study. *Clin Exp Rheumatol* 2005; 23: 819-28.
2. SALAFFI F, CAROTTI M, GASPARINI S, INTORCIA M, GRASSI W: The health-related quality of life in rheumatoid arthritis, ankylosing spondylitis, and psoriatic arthritis: a comparison with a selected sample of healthy people. *Health Qual Life Outcomes* 2009; 7: 25.
3. STRAND V, KHANNA D: The impact of rheumatoid arthritis and treatment on patients' lives. *Clin Exp Rheumatol* 2010; 28: S32-40.
4. KILTZ U, VAN DER HEIJDE D: Health-related quality of life in patients with rheumatoid arthritis and in patients with ankylosing spondylitis. *Clin Exp Rheumatol* 2009; 27: S108-11.
5. TURK DC, DWORKIN RH, BURKE LB *et al*.: Developing patient-reported outcome measures for pain clinical trials: IMMPACT recommendations. *Pain* 2006; 125: 208-15.
6. SALAFFI F, MIGLIORE A, SCARPELLINI M *et al*.: Psychometric properties of an index of three patient reported outcome (PRO) measures, termed the CLinical ARthritis Activity (PRO-CLARA) in patients with rheumatoid arthritis. The NEW INDICES study. *Clin Exp Rheumatol* 2010; 28: 186-200.
7. KALYONCU U, DOUGADOS M, DAURÈS JP, GOSSEC L: Reporting of patient-reported outcomes in recent trials in rheumatoid arthritis: a systematic literature review. *Ann Rheum Dis* 2009; 68: 183-90.
8. SCHOELS M, WONG J, SCOTT DL *et al*.: Economic aspects of treatment options in rheumatoid arthritis: a systematic literature review informing the EULAR recommendations for the management of rheumatoid arthritis. *Ann Rheum Dis* 2010; 69: 995-1003.
9. RÄSÄNEN P, ROINE E, SINTONEN H, SEMBERG-KONTTINEN V, RYYNANEN OP, ROINE R: Use of quality-adjusted life years for the estimation of effectiveness of health care: A systematic literature review. *Int J Technol Assess Health Care* 2006; 22: 235-41.
10. BRAZIER J, ROBERTS J, DEVERILL M: The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002; 21: 271-92.
11. BRAZIER J, USHERWOOD T, HARPER R *et al*.: Deriving a preference-based single index from the UK SF-36 health survey. *J Clin Epidemiol* 1998; 51: 1115-28.
12. ARA R, BRAZIER J: Predicting the short form-6D preference based index using the eight mean short form-36 health dimension scores: estimating preference-based health-related utilities when patient level data are not available. *Value Health* 2009; 12: 346-53.
13. OLIVIERI I, DE PORTU S, SALVARANI C *et al*.: The psoriatic arthritis cost evaluation study: a cost-of-illness study on tumour necrosis factor inhibitors in psoriatic arthritis patients with inadequate response to conventional therapy. *Rheumatology* 2008; 47: 1664-70.
14. LEARDINI G, SALAFFI F, MONTANELLI R,

GERZELI S, CANESI B: A multicenter cost-of-illness study on rheumatoid arthritis in Italy. *Clin Exp Rheumatol* 2002; 20: 505-15.
15. SAVOIA E, FANTINI MP, PANDOLFI PP, DALLOLIO L, COLLINA N: Assessing the construct validity of the Italian version of the EQ-5D: preliminary results from a cross-sectional study in North Italy. *Health Qual Life Outcomes* 2006; 4: 47.
16. APOLONE G, MOSCONI P: The Italian SF-36 Health Survey: translation, validation and norming. *J Clin Epidemiol* 1998; 51: 1025-36.
17. SALAFFI F, DE ANGELIS R, STANCATI A, GRASSI W: MArche Pain; Prevalence INvestigation Group (MAPPING) study. Health-related quality of life in multiple musculoskeletal conditions: a cross-sectional population based epidemiological study. II. The MAPPING study. *Clin Exp Rheumatol* 2005; 23: 829-39.
18. SALAFFI F, SARZI-PUTTINI P, GIROLIMETTI R, ATZENI F, GASPARINI S, GRASSI W: Health-related quality of life in fibromyalgia patients: a comparison with rheumatoid arthritis patients and the general population using the SF-36 health survey. *Clin Exp Rheumatol* 2009; 27: S67-74.
19. SALAFFI F, CIMMINO MA, MALAVOLTA N *et al*.: Italian Multicentre Osteoporotic Fracture Study Group. The burden of prevalent fractures on health-related quality of life in postmenopausal women with osteoporosis: the IMOF study. *J Rheumatol* 2007; 34: 1551-60.
20. PRIETO L, SACRISTÁN JA: Problems and solutions in calculating quality-adjusted life years (QALYs). *Health Qual Life Outcomes* 2003; 19: 80.
21. TORRANCE GW: Utility approach to measuring health-related quality of life. *J Chronic Dis* 1987; 40: 593-603.
22. BOONEN A, MAETZEL A, DRUMMOND M *et al*.: The OMERACT Initiative. Towards a reference approach to derive QALY for economic evaluations in rheumatology. *J Rheumatol* 2009; 36: 2045-9.
23. NATIONAL INSTITUTE FOR CLINICAL EXCELLENCE: Guide to the Methods of Technology Appraisal. *NICE*, London, 2004.
24. BRYAN S, LONGWORTH L: Measuring health-related utility: why disparity between EQ-5D and SF-6D? Eur. *J Health Econ* 2005; 6: 253-60.
25. ARNETT FC, EDWORTHY SM, BLOCH DA *et al*.: The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988; 31: 315-24.
26. PREVOO ML, VAN'T HOF MA, KUPER HH, VAN LEEUWEN MA, VAN DE PUTTE LB, VAN RIEL PL: Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995; 38: 44-8.
27. VAN GESTEL AM, PREVOO ML, VAN'T HOF MA, VAN RIJSWIJK MH, VAN DE PUTTE LB, VAN RIEL PL: Development and validation of the European league against rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. *Arthritis Rheum* 1996; 39: 34-40.
28. VAN GESTEL AM, HAAGSMA CJ, VAN RIEL PLCM: Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum* 1998; 41: 1845-50.
29. WELLS G, BECKER JC, TENG J *et al*.: Validation of the 28-joint Disease Activity Score (DAS28) and European League Against Rheumatism response criteria based on C-reactive protein against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on erythrocyte sedimentation rate. *Ann Rheum Dis* 2009; 68: 954-60.
30. PREVOO ML, VAN GESTEL AM, VAN HT, VAN RIJSWIJK MH, VAN DE PUTTE LB, VAN RIEL PL: Remission in a prospective study of patients with rheumatoid arthritis. American preliminary remission criteria in relation to the disease activity score. *Br J Rheumatol* 1996: 35: 1101-5.
31. VAN DER HEIJDE D: How to read radiographs according to the Sharp/van der Heijde method. *J Rheumatol* 2000; 27: 261-3.
32. KIND P: The EuroQol instrument: an index of health-related quality of life. In: Spiler B, ed. Quality of life and pharmaeconomics in clinical trials. Philadelphia: Lippincott-Raven, 1996.
33. WARE J, KOSINSKI M, BAYLISS M, ROGERS WH, RAZEC A: Comparison of methods for the scoring and statistical analysis of SF-36 health profile and summary measures: summary of results from the Medical Outcomes Study. *Med Care* 1995; 4: 264-79.
34. BLAND JM, ALTMAN DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307-10.
35. BARTON GR, SACH TH, DOHERTY M, AVERY AJ, JENKINSON C, MUIR KR: An assessment of the discriminative ability of the EQ-5D, SF-6D, and EQ-VAS, using socio-demographic factors and clinical conditions. *Eur J Health Econ* 2008; 9: 237-49.
36. KONTODIMOPOULOS N, PAPPA E, PAPADOPOULOS AA, TOUNTAS Y, NIAKAS D: Comparing SF-6D and EQ-5D utilities across groups differing in health status. *Qual Life Res* 2009; 18: 87-97.
37. SWETZ JA: Measuring accuracy of diagnostic systems. *Science* 1988; 240: 1285-93.
38. HANLEY JA, MCNEIL BJ: A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983; 148: 839-43.
39. WOLFE F, MICHAUD K: The loss of health status in rheumatoid arthritis and the effect of biologic therapy: a longitudinal observational study. *Arthritis Res Ther* 2010; 12:R35.
40. LILLEGRAVEN S, KRISTIANSEN IS, KVIEN TK: Comparison of utility measures and their relationship with other health status measures in 1041 patients with rheumatoid arthritis. *Ann Rheum Dis* 2010; 69: 1762-7.
41. MARRA CA, RASHIDI AA, GUH D *et al*.: Are indirect utility measures reliable and responsive in rheumatoid arthritis patients? *Qual Life Res* 2005; 14: 1333-44.
42. FLEURENCE R, SPACKMAN E: Cost-effectiveness of biologic agents for treatment of autoimmune disorders: structured review of the literature. *J Rheumatol* 2006; 33: 2124-31.
43. BRAZIER J, ROBERTS J, TSUCHIYA A, BUSSCHBACH J: A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Economics* 2004; 13: 873-84.
44. HARRISON MJ, LUNT M, VERSTAPPEN SM, WATSON KD, BANSBACK NJ, SYMMONS DP: Exploring the validity of estimating EQ-5D and SF-6D utility values from the health assessment questionnaire in patients with inflammatory arthritis. *Health Qual Life Outcomes* 2010; 8: 21.
45. KOPEC JA, WILLISON KD: A comparative review of four preference-weighted measures of health-related quality of life. *J Clin Epidemiol* 2003; 56: 317-25.
46. JORSTAD IC, KRISTIANSEN IS, UHLIG T, KVIEN TK: Performance of four utility measures in 1041 patients wth RA: Well correlated but differing widely in valuing health states. *Arthritis Rheum* 2005; 52: 660-1.
47. LONGWORTH L, BRYAN S: An empirical comparison of EQ-5D and SF-6D in liver transplant patients. *Health Econ* 2003; 12: 1061-7.
48. MARRA CA, WOOLCOTT JC, KOPEC JA *et al*.: A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. *Soc Sci Med* 2005; 60: 1571-82.
49. CONNER-SPADY B, SUAREZ-ALMAZOR ME: Variation in the estimation of quality-adjusted life-years by different preference-based instruments. *Med Care* 2003; 41: 791-801.
50. DROSSAERS-BAKKER KW, DE BUCK M, VAN ZEBEN D, ZWINDERMAN AH, BREEDVELD FC, HAZES JMW: Long-term course and outcome of functional capacity in rheumatoid arthritis. The effect of disease activity and radiological damage over time. *Arthritis Rheum* 1999; 42: 1854-60.
51. GUILLEMIN F, BRIANÇON S, POUREL J: Functional disability in rheumatoid arthritis: two different models in early and established disease. *J Rheumatol* 1992; 19: 366-9.
52. WELSING PM, VAN GESTEL AM, SWINKELS HL, KIEMENEY LA, VAN RIEL PL: The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. *Arthritis Rheum* 2001; 44: 2009-17.
53. WELSING PM, BORM GF, VAN RIEL P: Minimal clinically important difference in radiological progression of joint damage. A definition based on patient perspective. *J Rheumatol* 2006; 33: 501-7.
54. KNIJFF-DUTMER E, DROSSAERS-BAKKER W, VERHOEVEN A *et al*.: Rheumatoid factor measured by fluoroimmunoassay: a responsive measure of rheumatoid arthritis disease activity that is associated with joint damage. *Ann Rheum Dis* 2002; 61: 603-7.
55. MOLENAAR ET, VOSKUYL AE, DINANT HJ, BEZEMER PD, BOERS M, DIJKMANS BA: Progression of radiologic damage in patients with rheumatoid arthritis in clinical remission. *Arthritis Rheum* 2004; 50: 36-42.
56. WELSING PM, LANDEWÉ RB, VAN RIEL PL *et*

*al*.: The relationship between disease activity and radiologic progression in patients with rheumatoid arthritis. *Arthritis Rheum* 2004; 50: 2082-93.
57. RUPP I, BOSHUIZEN HC, JACOBI CE, DINANT HJ, VAN DEN BOS G: Comorbidity in patients with rheumatoid arthritis: effect on health-related quality of life. *J Rheumatol* 2004; 31: 58-65.
58. BERKANOVIC E, HURWICZ ML: Rheumatoid arthritis and comorbidity. *J Rheumatol* 1990; 17: 888-92.
59. GABRIEL SE, CROWSON CS, O'FALLON WM: Comorbidity in arthritis. *J Rheumatol* 1999; 26: 2475-9.
60. MIKULS TR: Co-morbidity in rheumatoid arthritis. *Best Pract Res Clin Rheumatol* 2003; 17: 729-52.
61. ADAMS R, WALSH C, VEALE D, BRESNIHAN B, FITZGERALD O, BARRY M: Understanding the relationship between the EQ-5D, SF-6D, HAQ and disease activity in inflammatory arthritis. *Pharmacoeconomics* 2010; 28: 477-87.
62. WOLFE F, MICHAUD K, WALLENSTEIN G: Scale characteristics and mapping accuracy of the US EQ-5D, UK EQ-5D, and SF-6D in patients with rheumatoid arthritis. *J Rheumatol* 2010; 37: 1615-25.
63. KROOT EJ, VAN GESTEL AM, SWINKELS HL, ALBERS MM, VAN DE PUTTE LB, VAN RIEL PL: Chronic comorbidity in patients with early rheumatoid arthritis: a descriptive study. *J Rheumatol* 2001; 28: 1511-7.
64. BERESNIAK A, RUSSELL AS, HARAOUI B, BESSETTE L, BOMBARDIER C, DURU G: Advantages and limitations of utility assessment methods in rheumatoid arthritis. *J Rheumatol* 2007; 34: 2193-200.