STUDY PROTOCOL

Validation of the EULAR patient-derived rheumatoid arthritis impact of disease (RAID) preliminary score based on patients' perception of the impact of the disease on dimensions of health

Addition to the protocol:

Elaboration of the EULAR patient-derived rheumatoid arthritis impact of disease (RAID) score based on patients' perception of the impact of the disease on dimensions of health: weighting of domains

Date: 2 January 2008
Development Phase: 4
Study Design: Multi-center study in subjects with rheumatoid arthritis
Convenor Tore K Kvien, MD
Facilitator Laure GOSSEC, MD
Contact: Hôpital Cochin
Rhumatologie B
27 rue Fbg St Jacques
75014 PARIS, FRANCE
Phone: (33) 1 5841562
Fax: (33) 1 58412628
laure.gossec@cch.aphp.fr

Steering committee
Sandra CANADELO
Loreto CARMONA
Maarten DE WIT
Maxime DOUGADOS
Laure GOSSEC
Turid HEIBERG
Tore KVIEN

This study will be conducted in compliance with the protocol, Good Clinical Practice, appropriate patient consent, and all other applicable regulatory requirements.
# LG Validation of RAID

## Table of Contents

1.0 **Introduction** .................................................................4

1.1 Background ..............................................................................4

1.2 Summary of first steps of this project ...........................................5

1.3 Study objectives ........................................................................6

2.0 **Methods for validation of the RA Impact of Disease Score** .................................................................7

2.1 Translation of the score ..............................................................7

2.2 Feasibility of the RAID .............................................................9

2.3 Face, construct and external validity of the RAID .................................9

2.3.1 Study design ...........................................................................9

2.3.2 Inclusion/exclusion criteria .....................................................9

2.3.3 Data collection .........................................................................9

2.4 Final choice of domains ............................................................10

2.5 Assessment of reliability of the RAID ..........................................11

2.5.1 Study design ..........................................................................11

2.5.2 Inclusion criteria ......................................................................11

2.5.3 Data collection .........................................................................11

2.6 Sensitivity to change of the RAID ................................................11

2.6.1 Study design ..........................................................................12

2.6.2 Inclusion criteria ......................................................................12

2.6.3 Data collection .........................................................................12

2.7 Cutoff values of the RAID defining state and change ...........................12

2.7.1 Study design ..........................................................................13

2.7.2 Inclusion criteria ......................................................................13

2.7.3 Data collection .........................................................................13

2.8 Statistical Methods .....................................................................14

2.8.1 Face, construct and external validity ........................................14

2.8.2 Final choice of tools and domains ...........................................15

2.8.3 Reliability ................................................................................15

2.8.4 Sensitivity to change ................................................................15

2.8.5 Cutoffs defining low activity and relevant change ......................16

2.8.6 Sample size calculations .........................................................16

3.0 **Regulatory and ethical aspects** .............................................17

3.1 Study coordinators ......................................................................17

3.2 Regulations and Review Boards ..................................................17

3.3 Data Quality Assurance .............................................................17

4.0 **Practical aspects** ....................................................................18

LG Validation of RAID  modified draft 2 January 2008
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Participants</td>
<td>18</td>
</tr>
<tr>
<td>4.2</td>
<td>Financial aspects</td>
<td>18</td>
</tr>
<tr>
<td>4.3</td>
<td>Project milestones: planned timetable</td>
<td>19</td>
</tr>
<tr>
<td>5.0</td>
<td>References</td>
<td>20</td>
</tr>
<tr>
<td>6.0</td>
<td>Annex 1. Proposed wording of RAID to be evaluated</td>
<td>22</td>
</tr>
<tr>
<td>7.0</td>
<td>Annex 2. Affiliations of participants</td>
<td>29</td>
</tr>
</tbody>
</table>
1.0 Introduction

1.1 Background

Rheumatoid arthritis (RA) is currently recognized as a heterogeneous entity that is usually diagnosed with reference to the American College of Rheumatology (ACR) classification criteria [1]. The clinical course of RA is variable and its prognosis is difficult to predict [2, 3]. In many patients, the disease process is severe and results in pain and functional loss, then progressive joint destruction and severe disability.

The aim of treatment in RA is double: on the one hand, to treat current patients’ symptoms, and on the other hand, to prevent ultimate disability.

Disability is largely predicted by structural degradation [4], which in its turn is closely related to synovitis [5]. Thus, for this axis of treatment, efficacy of treatment should be assessed largely by structural evaluation (i.e., radiographs) and by synovitis counts, which may be helped by ultrasonography. Biological markers of inflammation may also play a role here.

To treat current patients’ symptoms, one must rely on patients’ assessments. Here, it is possible to simply rely on global assessment (i.e., “how are you today?”). However, data suggest that global assessment is insufficient to reflect patients’ wellbeing, and does not correctly reflect many domains of importance for the patient, such as fatigue [6].

A universally recognized “core set” has been elaborated in 1992 [6a, 6b, 6c]. This core set includes some patient-reported domains (pain, function and patient global) but also included “physicians” domains. Furthermore, it was not patient-derived. A widely used composite index, the Disease Activity Score (DAS) [7] takes into account both joint counts, acute phase reactants and patient perceived activity; the difficulty with this score is that when there is a discordance between these fundamentally different aspects of the disease, results will be identical, e.g. high patient perceived activity with low synovitis counts, versus low patient symptoms with high synovitis counts.

If it is considered that the aim of treatment in RA is BOTH to prevent later disability AND to treat patients’ symptoms, it seems important to use 2 different tools, one related to prediction of disability (i.e., synovitis assessment), and a different score to assess patient perceived symptoms.

Some patient questionnaires are available, such as the Health Assessment Questionnaire, HAQ [8] but they only take into account part of the domains which are important for the patient (e.g., functional assessment for the HAQ). Work has already been performed by an OMERACT, Outcome Measures in RA Clinical Trials group regarding the patient’s perspective on domains of importance in RA [9, 10], and by focus groups; this qualitative work showed that some domains are not being sufficiently considered in RA, such as fatigue, well-being, sleep pattern, return to normal life.

One scale has been developed to incorporate three different dimensions, physical, pain and patient global (the patient activity scale, PAS (10a)), but patients were not directly involved in the development of this scale.
Thus, to date, to our knowledge, a validated composite index reflecting patient-perceived impact of RA and taking into account all the domains of importance for the patient does not exist.

1.2 Summary of first steps of this project.

As a first step to the current project, a patient-derived impact of disease scale has been developed under the EULAR umbrella (see previous protocol).

Choice of domains or dimensions of health to be included
Ten patients from different European countries participated in the elaboration of the score by determining domains or areas of health important for the patient. They determined a list of 17 important domains in March 2007. Then 100 patients ranked these 17 domains in order of decreasing importance (10 per country in 10 countries). Thus an order of decreasing importance was obtained.

The steering committee decided that for feasibility reasons, the final score should include 6 to 8 domains. Thus, the 7 most important domains were selected. These domains are in order of decreasing importance: pain, function, fatigue, sleep, emotional well-being, physical well-being, and coping/helplessness.

Choice of tools or questionnaires for each domain
Tools or specific questionnaires were chosen for each domain, by expert opinion (steering committee and principal investigators) based on an extensive literature review.

Several questionnaires could be chosen for a given domain, and will be compared in terms of psychometric properties during the validation phase.

Relative importance of the different domains
It was decided to obtain a relative weight for the domains, based on the patient’s perspective. The relative weight for the patient of these domains was determined in 500 patients by a questionnaire where they were asked to “spend 100 points” across the 7 domains. Demographic and severity data were also collected (analysis is ongoing).

The relative weights of the domains are the following: pain 28%, function 16%, fatigue 14%, sleep 11%, coping 11%, psychological well-being 10%, physical well-being 10%.

Results
Thus a preliminary RA impact of disease (RAID) score is now available. The present format of the RAID is in Annex 1. Several questions or questionnaires are included for each domain, and will be compared in the present study.
1.3 Study objectives

The objectives of this study are to finalize and validate the preliminary RAID.

This score will be validated according to international recommended methodology, with reference to the OMERACT filter [11], i.e. the composite index must be feasible, must be truthful, i.e. must have face and construct validity, must be reproducible, and must be sensitive to change. This score is planned to be used for clinical trials in RA. Further research, beyond the limits of this research protocol, will include assessing the possible usefulness of the RAID at the individual level, i.e. could this tool designed for clinical research also be useful for clinical practice.

The current protocol concerns the following steps:
- assessment of psychometric properties of the RAID score, including its feasibility, its face, construct and external validity (correlation with other validated scores), its reproducibility, and sensitivity to change.
- Final choice of domains and of tools. During this phase, the number of domains may be brought down from 7 to 6 domains and there will be a choice between the different available tools (questions or questionnaires) for each domain, based on their psychometric properties.
- We also plan to identify cut-offs measuring state and change with RAID; that is the value of the RAID score defining a binary state (Patient Acceptable Symptom State) and if possible cut-offs defining a binary change (minimal clinically important improvement).

Discriminant capacity will not be evaluated as part of this protocol.
2.0 Methods for validation of the RA Impact of Disease Score

2.1 Translation of the score

A necessary step is the translation of the score into European languages. If available, existing validated translations of the tools integrated in the RAID will be used (see Annex 1). If no translations are available, a validated translation procedure will be followed.

Translation and cross cultural adaptation will be performed according to published recommendations [13, 14].
A. Two or three persons (at least 1 rheumatologist and 1 teacher of English, all as bilingual as possible, of whom at least one bilingual person (if possible also the patient representative participating in the first Zurich meeting – alternatively this patient participates under item D) native in the target language) translate independently the English version into the target language.

B. A single preliminary version is obtained during a simple consensus meeting with the 2 to 3 translators. Please keep in mind at this phase that the final wording needs to be understood by lay people including low-education people (“the aim is for a 9-yr old child to understand the wording”).

C. Backward translation is then performed by an independent bilingual native English speaker, blinded to the English original version.

D. A multidisciplinary consensus committee then meets, including the initial 2-3 translators, at least 2 rheumatologists (who may also be part of the translators), one person very familiar with cross-cultural adaptation, and at least one patient fluent in English. The group will compare the initial version and the back-translation and will discuss the phrasing of the target-language version, and by consensus will produce a final version. The committee has to ensure that the translation is fully comprehensive and to verify cross-cultural equivalence of the source and final versions. It may be interesting or even necessary during this meeting to contact by telephone the “author” of the initial English-version questionnaire (LG) to verify that the initial concepts are perfectly understood. Please keep in mind again at this phase that the final wording needs to be understood by lay people including low-education people.

E. The final version is pre-tested with 15-20 target-language-native patients. These patients fill in the questionnaire in the presence of a physician, and each question is discussed with the patient, to check whether it is fully understood for all items and whether the patients have problems with the formulation. Patients’ comments are collected and the initial translators may need to go back to the translation and modify it, if comments are frequent and consistent.

QUESTIONS FOR ZURICH

Skip step E?

How many instruments need translation according to these recommendations? (each PI should check before the meeting the availability of instruments in their language)
2.2 Feasibility of the RAID

This important aspect means: will the RAID be usable in trials and/or in clinical practice. Because it is a questionnaire, it is not painful, not costly (the final RAID will be available free of cost on the EULAR website) and not difficult to obtain. However length of the questionnaire and time necessary to complete it are issues that need consideration.

2.3 Face, construct and external validity of the RAID

Face validity is ascertained by feedback from the group of patients initially involved in the development of the RAID, and by feedback from the expert participants.

Construct and external validity are assessed by obtaining RAID results and comparing them cross-sectionally to other scores used in RA.

This step will also allow us to perform a final choice of tools for each domain.

2.3.1 Study design

Cross-sectional assessment; international study at 1 time point. The aim is to obtain data from 600 patients, i.e. 50 per country, by establishing a collaborative study in 1 to 3 sites of the participating countries of consecutive patients with RA. For each patient, data will be collected once.

2.3.2 Inclusion/exclusion criteria

The objective is to obtain data from a large number of patients, with a large range of severities and treatments.
All adult patients with RA consulting a rheumatologist for follow-up are eligible to participate except those who decline.
Early arthritis which is not definitely RA will be excluded. Patients with concomitant other inflammatory disease will be also excluded. Patients unable to fill in a questionnaire will be excluded.

2.3.3 Data collection

Collection of patient data – Data collection will take place over a 1-3 month period, to include 60 patients with RA at each country. Each site will be asked to distribute a questionnaire to every consecutive, unselected patient with RA if
inclusion criteria are satisfied. The patient will fill in the questionnaire in his/her own language. During this assessment, the patients will complete all questions in the present format of the RAID, and will also complete other relevant health status measures (i.e. SF-36, HAQ, global assessment on VAS) as well as demographic variables. External validity will be assessed cross-sectionally by comparison of RAID to these validated measures.

Concomitant collection of physician data
The investigator will complete a 2-page clinical evaluation of RA which may be translated or completed in English. This will include:
  a. Clinical features of RA, ACR classification criteria
  b. All medications used to treat arthritis.
  c. Standard 44 joint count.
  d. Investigator’s global assessment of disease activity
  e. Current results of laboratory tests for erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), and rheumatoid factor.
This will also allow to compare the RAID results with other validated activity scores such as the Disease Activity Score, DAS28 as well as other patient reported outcomes.
A modified version of the QUEST questionnaire will be used (copyright, Ted Pincus: agreement for use obtained).

2.4 Final choice of domains

In the literature, the domains chosen by patients in the previous steps of elaboration of the RAID all appear to be important for other patients in qualitative studies [9, 10, 17, 18], except coping. Furthermore, coping is less sensitive to change than the other domains. Coping was the seventh domain in order of importance and the steering committee arbitrarily decided to keep 7 domains.

It is possible that when analyzing the results of the validation, it is found that coping is not sensitive to change and/or does not bring more information into the RAID. If this was the case, it is anticipated that the final RAID questionnaire may not include coping, i.e. may include only the other 6 domains, pain, function, fatigue, sleep, physical and emotional well-being.

In this case however, the relative weights of the different domains, which were assessed previously based on patients’ perspective, will not be appropriate since we do not know where the patients “would have distributed” the points that they actually gave for coping.

Therefore, we plan in the present study to perform once again a weighting process, but only for the 6 domains (without coping). This is possible by adding
one question: “please spend 100 points between the following 6 domains” to the face validity questionnaire.

If the final score includes the 7 domains i.e. with coping, then the first weighting will be used (RAID 7 domains). If coping is finally excluded, the second weighting will be used (RAID 6 domains).

2.5 Assessment of reliability of the RAID

2.5.1 Study design

Longitudinal assessment international study at 2 time points within a 2-7 days interval.
The aim is to obtain data from 60 patients, i.e. 5 per country, a sub-group of the face validity study (above).

2.5.2 Inclusion criteria

Sub-group of the face validity study (above).
Patients must have a disease which is in a stable phase, with stable treatment (no change in DMARDs (or DMARD dose) over last 2 months, no change in NSAIDs over last week).
Changes in treatment between the 2 assessments will be an exclusion criterion.
A random selection of stable patients will be performed among patients participating in the face validity study. However patients unwilling or unable to come back one week later for a second assessment will be excluded.

2.5.3 Data collection

It is planned to obtain RAID results from 60 patients at a 1-week interval on stable patients. Patients participating above, will be asked to complete a second questionnaire at home 1 week later and send it to the investigator by mail. The questionnaire will include a question regarding treatment changes. Otherwise the assessment will be identical to baseline, except for the assessments performed by investigator.

DISCUSS IN ZURICH: patient comes in for new assessment and has also clinical examination including lab tests (for comparative reliability of DAS). If not, use RADAI (is this translated in every language?)

2.6 Sensitivity to change of the RAID
2.6.1 Study design

Longitudinal assessment international study at 2 time points. For each patient, data will be collected twice (8 weeks interval between assessments).

The aim is to obtain data from at least 20 patients, i.e. at least 10 per country; a sub-group of the face validity study (above) or another subset of patients fulfilling the specific inclusion criteria (see below). 120 is a minimal number, a higher number is preferred.

2.6.2 Inclusion criteria

Sub-group of the patients participating in the face validity study (above) or another subgroup who requires a therapeutic change. The change will be start of systemic corticosteroids, biological agents or start of conventional DMARDs in DMARD naïve patients.
These patients will be included only if the therapeutic change is known to be effective (since it is used as the gold standard for change) and if expected effect is rapid (less than 2 months).
Patients unwilling or unable to come back two months later for a second assessment will be excluded.

DISCUSS IN ZURICH: precise list of treatments that are considered as entry criteria – consider the time for second assessment

2.6.3 Data collection

Sensitivity to change will be assessed by comparing data from the 2 assessments.

Patients answering the criteria will come to the outpatient clinic again, 2 months after the first assessment. The baseline assessments will be repeated, including the clinical assessment.

2.7 Cutoff values of the RAID defining state and change

Results obtained from the RAID will be expressed as one continuous value. However, it is important to be able to express these results as dichotomous (i.e. as number of patients with a clinically significant improvement and also as number of patients with a low disease activity [12, 15, 16]).

To this end, we plan to add questions to the questionnaire concerning tolerable levels of symptoms, to determine the Patient Acceptable Symptom State, PASS and change from the patient's perspective, the Minimal Clinically Important Improvement, MCII.
Thus, a secondary objective of this study is to obtain cutoff values of the RAID, allowing the determination of binary elements:

- **state**: value of RAID defining low disease activity based on the patient’s perspective, the Patient Acceptable Symptom State.
- **Change**: value of change in RAID defining the Minimal Clinically Important Improvement.

For these 2 cutoffs, study design differs.

### 2.7.1 Study design

Cutoff defining state: cross-sectional study. Thus a question will be added to the cross-sectional face/construct validity study. (objective, 600 patients)

Cutoff defining change: longitudinal study. Thus a question will be added to the longitudinal sensitivity to change study (objective, 120 patients).

**DISCUSS IN ZURICH: all patients longitudinal assessment? (for MCII) ? or 120?**

### 2.7.2 Inclusion criteria

The same patients participating in the above studies will participate here:
- validity cross-sectional study for cutoff for state
- sensitivity to change longitudinal study for cutoff for change.

### 2.7.3 Data collection

The gold standard for state and change will be the patient’s assessment. Two questions will be added to the questionnaire [19]:

- for PASS: . If you were to remain for the next few months with the same level of impact of disease you had during the last 48 hours, would this be acceptable or unacceptable for you?
- For MCII: Compared to when you filled in the first questionnaire, how has the impact of disease been during the last 48 hours? (answers, improved – less impact versus no change versus worse – more impact) THEN If you answer “improved” at the previous question, how important is this improvement to you? (answers very important to not at all important).

**DISCUSS IN ZURICH: additional anchor for PASS, also using 5 response categories – very acceptable, acceptable, neither acceptable or unacceptable, unacceptable, very unacceptable (since very acceptable can provide an
opportunity to identify more ambitious cut-points for PASS than achieved with the standard question)

2.8 Statistical Methods

As explained above, validation of the RAID will include several steps:
- Assessment of face, construct and external validity.
- Final choice of tools and of domains
- Reliability
- Sensitivity to change
- Determination of cutoffs defining state and change.
The statistical methods will be adapted to these steps.

2.8.1 Face, construct and external validity

Face validity is ascertained by feedback from the group of patients initially involved in the development of the RAID, and by feedback from the expert participants.

Construct validity is evaluated by several tests.
To assess validity, results of the RAID will be compared to “gold standard” questions to assess impact of RA. These questions are global assessment of impact questions. Because they are imperfect gold standards, we will compare RAID to several questions. We will use the following questions:
- Considering all the ways in which illness and health conditions may affect you at this time, please make a mark below to show how you are doing.
- In general, how active has your rheumatic condition been over the PAST SIX MONTHS?
- In general, would you say your health is …very good to poor, Likert scale (question 1 of the SF-36)
Results of RAID will also be compared to SF-36, a validated quality of life questionnaire.
Non-parametric correlations will be performed between the different items of the RAID in order to check the % patients for whom a discrepancy exists between the levels of each of the domains in the RAID score.

Internal validity will be assessed by Cronbach’s alpha coefficient [20]. Discrimination item/full score will be analysed as well as correlations between domains and multiple factorial analyses will also be performed to determine the different dimensions in the full score.

External validity is measured through divergent correlation. Non-parametric correlation between RAID scores and a largely validated score measuring activity
of disease i.e. the disease activity score, DAS28 will be performed for the complete RAID score, and separately for each item of the RAID.

2.8.2 Final choice of tools and domains

For each domain of the RAID, several tools have been selected, e.g. numerical rating scale versus full questionnaire. The number of total combinations between tools is 36, using the 7 domains; and 12 using only 6 domains (if coping is excluded). Therefore we plan to test 48 combinations. These combinations will be compared in terms of construct validity and external validity as well as reliability and sensitivity to change (see below). We will also consider feasibility (a shorter questionnaire is more feasible). Therefore, we will determine if shorter questionnaires lead to loss of psychometric properties to determine the optimal tools. This will also allow us to compare combinations with 6 versus with 7 domains included.

It is possible that 2 combinations of tools will be proposed: one simple, short questionnaire reflecting each domain of the RAID but with few questions, for use in the clinic; and one longer, perhaps more optimal but less feasible questionnaire, for use in clinical trials.

2.8.3 Reliability

This will be assessed through comparisons of results obtained at a 1-week interval, for patients considering themselves as in a stable state. The results will be compared by intra-class coefficient correlations type III and kappa statistics, as appropriate. A graphic analysis according to Bland and Altman will also be performed.

2.8.4 Sensitivity to change

Results obtained before and after treatment will be compared and sensitivity to change will be measured using the standardised response mean, SRM, defined as change divided by standard deviation of change, Friedman tests and if possible an ANOVA for repeated measures.
2.8.5 Cutoffs defining low activity and relevant change

The gold standard will be defined by the patient's opinion regarding the level of impact and/or of change (see above). A “non-acceptable symptom state threshold” will be evaluated by using both a ROC analysis and also the 75th percentile technique. This can be performed for the full score and for each item of the RAID score.

The same techniques will be applied to obtain cutoffs for change (from the patient's point of view).

2.8.6 Sample size calculations

Cross-sectional construct validity and correlation validity

There are no published data allowing a precise sample size calculation in this case. We consider that to be able to ascertain that the RAID is valid, we need 500 questionnaires. We plan to have 10-20% missing or incomplete data and thus propose to assess validity in 600 patients (50 per country).

Reliability

It is usually considered that 30 test-retests are enough to obtain a calculation. We plan to have missing data and thus propose to assess reliability in 60 patients (5 per country).

Longitudinal sensitivity to change

It is usually considered that 60 patients are enough to obtain an estimation of sensitivity to change. However, because this is an international study and because we expect variability, we plan to include 120 patients in this study, i.e. 10 per country. We expect to have 10-20% missing or incomplete data therefore evaluable data will come from 100 patients.

Measure of cutoffs defining state and change

For state, we will use the data from the 600 cross-sectional assessments and for change, data from the 120 patients included in the sensitivity to change study. However for change, because only a subgroup a patients (defining themselves as improved) can be analysed, it is possible that the present study size will not be sufficient to adequately define the MCII value of RAID. We consider that the objective here is to obtain a preliminary MCII cutoff.
3.0 Regulatory and ethical aspects

3.1 Study coordinators

This project is a collaborative effort of several European centres, who contribute medical data collected on their patients for central data analysis.

In order to analyse the data and in the interest of all participants, the study coordinators are responsible for correct data procurement, delivery of analysis results to the participants and development of agreed publication strategies.

The study coordinators for this phase are:
- the members of the steering committee
- the advisory committee, comprising the Principal Investigator from each participating country.

3.2 Regulations and Review Boards

The study will be conducted in accordance with the protocol, ICH Good Clinical Practice, ethical principles that have their origin in the Declaration of Helsinki and all applicable local regulations.

Independent Ethics Committee or Institutional Review Board approval of the protocol will be obtained prior to commencing the program at each site, through the principal investigator and designated investigators. The details will be determined in each clinical setting that participates in this program.

This protocol may be submitted to Ethics Committees as an addendum to the previously approved “elaboration of RAID” protocol.

An Institutional Review Board approved, study-specific informed consent will be reviewed, signed and dated by the subject (and the investigator) prior to the performance of any study-related procedures.

3.3 Data Quality Assurance

Data treatment. All results will be forwarded anonymously to the data center in Paris, France to be entered into a pooled database. Patients will be identified by a local number at the site, which will have a 3-letter code for a site and a number for each patient. Each site will keep a confidential subject identification code list, so that if there are missing data this information will be available locally to clarify the information. Data will be stored and analysed anonymously.

Quality control

There are several levels for quality control.

A. Locally for each CRF.
Last question of the CRF for the investigator:
“Have you checked the patient has filled in all the questions of his/her CRF, and have you filled in all the questions of your questionnaire?”

B. National PI.
Each CRF is sent or faxed individually (one by one) as soon as data are collected, to the national PI. The national PI will have to check for potential missing data and, if yes, contact the center for correction.

C. Central quality control.
CRFs will be sent by the national PI to the statistical center, Paris, every 10 CRFs, to be checked centrally for quality control.

4.0 Practical aspects

4.1 Participants

It is anticipated that for this phase, 12 European countries will participate.

Inclusion and exclusion criteria for patients are as detailed above.

The following European countries will participate (see annex 2): Estonia, Finland, France, Germany, Greece, Italy, Netherlands, Norway, Romania, Spain, Turkey, United Kingdom.

Thus 12 countries will participate. In each country, the principal investigator may decide to include the patients in 1 to 3 centers. He or she is then responsible for his or her country data.

4.2 Financial aspects

Estimation of costs: total 55,000 euros.
- 1 meeting in Zurich in January 2008, 20,000 euros.
- CRF printing and mailing 2,800 euros
- Translation of CRF: 500 euros per country PI (total, 5500 euros because no translation for UK)
- For each patient included: 30 euros (ie for 600 patients, 18,000 euros) and 15 euros more for each second evaluation of a patient, ie for 60 reliability and sensitivity to change patients, 2700 euros
- Data entry and statistical analysis 6,000 euros.
EULAR has accepted to give a grant for this project.
4.3 Project milestones: planned timetable

01/2008 finalisation of study protocol

01/2008 nomination of a local investigator responsible for the conduct of the study in each country. Signing of a document by the investigator indicating that he/she adheres to the study protocol

01 to 02/2008 Translation of the RAID

02 to 03/2008 Ethical and administrative issues

03/2008-06/2008 Validation of the RAID: data collection

06/2008 presentation at EULAR of the preliminary RAID

07-10/2008 data analysis, preparation of manuscript.

The project progress will be monitored by two-month-reports prepared by the investigators and forwarded to the study coordinators.

It is planned to disseminate the results by presentations at national and international meetings and by publications in the rheumatology literature. It is estimated that 2 to 3 overall publications summarizing results of this protocol will emerge.

All national PIs and the steering committee will be coauthors.
5.0 References


## Annex 1. Proposed wording of RAID to be evaluated

<table>
<thead>
<tr>
<th>Domain / weight</th>
<th>Tools</th>
<th>Translations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pain (28%)</strong></td>
<td><strong>TOOL ONE</strong></td>
<td>Available in 6 languages from REFLECT study</td>
</tr>
<tr>
<td></td>
<td>CATEGORY: single question</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pain Numerical rating scale, NRS (see below)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PHRASING</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Circle the number that best describes the pain you felt due to your rheumatoid arthritis during the last 48 hours</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ANCHORS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0= none TO 10 = extreme</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RANGE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0 to 10. Higher scores indicate worse status</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>TOOL TWO</strong></td>
<td>Available in many countries</td>
</tr>
<tr>
<td></td>
<td>CATEGORY: questionnaire</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pain questions from the SF36 questionnaire but modified to a 48-hour timeframe</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PHRASING</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1. How much bodily pain have you had over the past 48 hours?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. During the past 48 hours, how much did pain interfere with your normal work (including both work outside the home and housework)?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ANCHORS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Likert, question 1 : 6 point Likert: none , very mild, mild, moderate, severe, very severe</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Question 2 : 5 point Likert: not at all, a little bit, moderately, quite a bit, extremely</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RANGE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Higher scores indicate worse status. We propose to normalise to a 0-10 range.</td>
<td></td>
</tr>
<tr>
<td><strong>Function (16%)</strong></td>
<td><strong>TOOL ONE</strong></td>
<td>Available in 6 languages from REFLECT study</td>
</tr>
<tr>
<td></td>
<td>CATEGORY: single question</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Function Numerical rating scale, NRS (see below)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PHRASING</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Circle the number that best describes the difficulty you had in doing daily physical activities due to your rheumatoid arthritis during the last 48 hours.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ANCHORS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0= physical activity is no problem TO 10 = physical activity is a major problem</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RANGE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0 to 10. Higher scores indicate worse status</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>TOOL TWO</strong></td>
<td>Validated in many languages</td>
</tr>
</tbody>
</table>
**TOOL ONE**
*CATEGORY: single question*
NRS (see below) modified from the phrasing used by F Wolfe (J Rheumatol. 2006 Oct;33(10):1942-51).

**PHRASING:**
Circle the number that best describes how much of a problem sleep (i.e., resting at night) has been for you due to your rheumatoid arthritis during the last 48 hours.

**ANCHORS**
0 = sleep is no problem TO 10 = sleep is a major problem.

**TOOL TWO**
*CATEGORY: questionnaire*
MOS sleep disturbance subscale questionnaire (see below). 4 questions modified to a 48 hour timeframe.

**PHRASING**
1. How long did it usually take for you to fall asleep during the past 48 HOURS?
2. How often in the during the past 4 weeks did you ...
3. feel that your sleep was not quiet (moving restlessly, feeling tense, speaking, etc., while sleeping)?
4. have trouble falling asleep?
5. awaken during your sleep time and have trouble falling asleep again?

**ANCHORS**
Question 1: 0-15 minutes (score 1), 16-30 minutes (score 2), 31-45 minutes (score 3), 46-60 minutes (score 4), More than 60 minutes (score 5)

Other questions: scored 1 to 6: All of the time, Most of the time, A good bit of the time, Some of the time, A little of the time, None of the time

**SCORING**
To score this domain you need to:
reverse the scoring on items 3, 7, 8 as follows: 1, 2, 3, 4, 5, 6
<table>
<thead>
<tr>
<th>Fatigue (14%)</th>
<th>CATEGORY: single question</th>
<th>Not available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatigue NRS (see below) with phrasing derived from F Wolfe (Wolfe F. J Rheumatol. 2004 Oct;31(10):1896-902.)</td>
<td>PHRASING: Circle the number that best describes how much of a problem fatigue has been for you due to your rheumatoid arthritis during the last 48 hours.</td>
<td></td>
</tr>
<tr>
<td>ANCHORS 2 versions of anchors 1. According to J Kirwan’s work: no fatigue TO totally exhausted 2. If translation difficulties: alternative anchors: fatigue is no problem to fatigue is a major problem</td>
<td>RANGE 0 to 10. Higher scores indicate worse status</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coping (11%)</th>
<th>TOOL ONE</th>
<th>Available in several languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>NRS created for this study: Considering your arthritis overall, how much of a problem has coping been for you? Circle the number that best describes how much of a problem coping has been for you during the last 48 hours.</td>
<td>ANCHORS 0= coping is no problem TO 10 = coping is a major problem</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TOOL TWO</th>
<th>CATEGORY: questionnaire</th>
<th>Only English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arthritis Helplessness Index (AHI) Helplessness subscale, 5-question questionnaire.</td>
<td>PHRASING Arthritis is controlling my life. I would feel helpless if I couldn’t rely on other people for help with my arthritis. No matter what I do, or how hard I try, I just can’t seem to get relief from my pain. I am coping effectively with my arthritis. It seems as though fate and other factors beyond my control affect my arthritis.</td>
<td></td>
</tr>
<tr>
<td>ANCHORS 6 point Likert : (1 = strongly disagree, 2 = moderately disagree, 3 = disagree, 4 = agree, 5 = moderately agree, 6 = strongly agree).</td>
<td>RANGE 5 to 30. Higher scores indicate worse status. Scoring To score: reverse scoring on item 4; sum all items for total score We propose to normalise AHI to a 0-10 range</td>
<td></td>
</tr>
</tbody>
</table>
## TOOL THREE
CATEGOR Y: questionnaire
Newth modified coping questionnaire (20 questions).
See below for full wording of the Newth modified questionnaire.
ANCHORS
Likert, see below
RANGE
Unknown. Higher scores indicate worse status. We propose to normalise to a 0-10 range

<table>
<thead>
<tr>
<th>Physical wellbeing (10%)</th>
<th>CATEGORY: single question</th>
<th>Not available</th>
</tr>
</thead>
<tbody>
<tr>
<td>NRS (see below) created for this study.</td>
<td>PHRASING: Considering your arthritis overall, how would you rate your level of physical well being during the past 48 hours? Circle the number that best describes your level of physical well-being on a 0-10 scale.</td>
<td></td>
</tr>
<tr>
<td>ANCHORS</td>
<td>0 = very good TO 10 = very bad.</td>
<td></td>
</tr>
<tr>
<td>RANGE</td>
<td>0 to 10. Higher scores indicate worse status</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Emotional well-being (10%)</th>
<th>CATEGORY: single question</th>
<th>Not available</th>
</tr>
</thead>
<tbody>
<tr>
<td>NRS (see below) created for this study.</td>
<td>PHRASING: Considering your arthritis overall, how would you rate your level of emotional well being during the past 48 hours? Circle the number that best describes your level of emotional well-being on a 0-10 scale</td>
<td></td>
</tr>
<tr>
<td>ANCHORS</td>
<td>0 = very good TO 10 = very bad.</td>
<td></td>
</tr>
<tr>
<td>RANGE</td>
<td>0 to 10. Higher scores indicate worse status</td>
<td></td>
</tr>
</tbody>
</table>

## ALL NRS
For homogeneity reasons, it is decided that all NRS will be presented in the same way, i.e.
- presented horizontally on a 0-10 scale,
- with as anchors 0: none or best state and 10: maximal symptom or worst state,
- with as time frame 48 hours
- with as phrasing one which is already validated if possible
- with an added notion of related to arthritis, such as “considering your arthritis overall,...”
- with no colours or other indicator of severity, only numbers to circle.

It is proposed that the final phrasing be translated to all countries and available on the EULAR website.
Health Assessment Questionnaire modified over 48 hours

We are interested in learning how your arthritis affects your ability to function in daily life.
Please check (√) the one best answer which best describes your usual abilities OVER THE PAST 48 HOURS:

<table>
<thead>
<tr>
<th>Activity</th>
<th>Without ANY Difficulty(0)</th>
<th>With SOME Difficulty(1)</th>
<th>With MUCH Difficulty(2)</th>
<th>UNABLE To Do(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRESSING &amp; GROOMING</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Dress yourself, including tying shoelaces and doing buttons?</td>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>- Shampoo your hair?</td>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>ARISING</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Stand up from a straight chair?</td>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>- Get in and out of bed?</td>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>EATING</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Cut your meat?</td>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>- Lift a full cup or glass to your mouth?</td>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>- Open a new milk carton?</td>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>WALKING</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Walk outdoors on flat ground?</td>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>- Climb up five steps?</td>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
</tbody>
</table>

Please check any AIDS OR DEVICES that you usually use for any of these activities:

- Cane
- Devices used for dressing (button hook, zipper pull, long-handed shoe horn, etc.)
- Walker
- Built up or special utensils
- Crutches
- Special or built up chair
- Wheelchair
- Other (Specify: ____________________________)

Please check any categories for which you usually need HELP FROM ANOTHER PERSON:

- Dressing and Grooming
- Eating
- Arising
- Walking
Please check the response which best describes your usual abilities OVER THE PAST 48 HOURS:

<table>
<thead>
<tr>
<th>HYGIENE</th>
<th>Without ANY Difficulty(0)</th>
<th>With SOME Difficulty(1)</th>
<th>With MUCH Difficulty(2)</th>
<th>UNABLE To Do(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are you able to:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Wash and dry your body?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Take a tub bath?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Get on and off the toilet?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>REACH</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Are you able to:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Reach and get down a 5 pound object (such as a bag of sugar) from just above your head?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Bend down to pick up clothing from the floor?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GRIP</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Are you able to:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Open car doors?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Open jars which have previously been opened?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Turn faucets on and off?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ACTIVITIES</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Are you able to:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Run errands and shop?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Get in and out of a car?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Do chores such as vacuuming or yard work?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. Please check any AIDS OR DEVICES that you usually use for any of these activities:

- Raised toilet seat
- Bathtub bar
- Bathtub seat
- Long-handled appliances for reach
- Jar opener (for jars previously opened)
- Long-handled appliances in bathroom
- Other (Specify: __________________)

6. Please check any categories for which you usually need HELP FROM ANOTHER PERSON:

- Hygiene
- Gripping and opening things
- Reach
- Errands and chores
Modified Coping Questionnaire
(Original version: Newth et al., 2004, Psychology & Health)

The following questionnaire is about possible mental strategies to bear up with your rheumatoid arthritis. Please indicate the extent to which you have used these strategies during the past 48 hours. Likert: not at all, rarely, some, a lot, not applicable.

- Reminded myself how much worse things could be.
- I thought about someone I know who is in a worse situation.
- Realized how, in some ways, I’m more fortunate than others.
- Rediscovered what is important in life.
- Changed or grew as a person in a good way.
- Kept others from knowing how bad it was.
- Tried to keep my pain to myself.
- Didn’t let it get to me; refused to think about it too much.
- Went on as if nothing had happened.
- Made light of the situation; refused to be upset.
- Tried to keep my pain from interfering with other things too much.
- Talked to someone about how I was feeling.
-Expressed anger.
- I let my feelings out somehow.
- Accepted sympathy and understanding from someone.
- Made a plan of action and followed it.
- Concentrated on what I had to do – the next step.
- I knew what I had to do so I increased my efforts to make things work.

If you had to rate the overall effectiveness of the above strategies to bear up with your rheumatoid arthritis which value would be most suitable to you? Please indicate your opinion on the rating scale below.
Anchors: (from 0 = very ineffective to 10 = very effective)

How helpless did you feel because of your rheumatoid arthritis although you used these strategies during the past 4 weeks? Please indicate the value which is most suitable to you on the rating scale below.
Anchors (0 = not at all helpless; 10 = very helpless)
7.0 Annex 2. Affiliations of participants.

**Steering committee**
- **Sandra Canadelo**, Vice President Eular Social Leagues
- **Loreto Carmona**, Research Unit, Spanish Foundation of Rheumatology, Calle Marques del Duero, 5, 1 28001 MADRID, Spain
- **Maarten de Wit** (MS), Vice President Eular Social Leagues, External representative of the Dutch Arthritis Patient League, Postbus 1370, 3800 BJ Amersfoort, The Netherlands
- **Maxime Dougados**, MD, Paris Descartes University, Cochin hospital, APHP, Rheumatology B Department, Paris, France
- **Laure Gossec**, MD, Paris Descartes University, Cochin hospital, APHP, Rheumatology B Department, Paris, France
- **Turid Heiberg** , senior advicer/research fellow, Dept for Research and Education, Ulleval University Hospital, 0407 Oslo, Norway
- **Tore K Kvien**, MD, Head of Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Professor University of Oslo, Norway

**External experts**
- **G. A. Wells**, MSc, PhD, Full Professor with the Department of Epidemiology and Community Medicine, University of Ottawa, Principal Investigator in the Institute of Population Health and Director, Cardiovascular Research Methods Centre, University of Ottawa Heart Institute. H1-1, 40 Ruskin Street, Ottawa, Ontario K1Y 4W7, Canada
- **Jean-Pierre Daurets**, PhD, Professor of Epidemiology and Statistics, head of department, Montpellier Medicine Faculty, Montpellier, France

**Principal investigators**
- **ESTONIA**: Kati Otsa, MD, Head of Rheumatology Department, Tallinn Central Hospital, Pärnu mnt 104, 11312 Tallinn, Estonia
- **FINLAND**: Tuulikki Sokka, Jyväskylä Central Hospital, 40620 Jyväskylä, Finland
- **FRANCE**: Maxime Dougados, MD, Paris Descartes University, Cochin hospital, APHP, Rheumatology B Department, Paris, France
- **GERMANY**: Georg Schett, Erlangen, Germany to be completed.
- **GREECE**: Dimitrios T Bournas, MD, FACP, Rheumatology, Clinical Immunology and Allergy, University of Crete, Faculty of Medicine, 1 Voutes str., 710 03 Heraklion, Greece
- **ITALY**: Marco Matucci Cerinic, professor of rheumatology, Dept of Medicine, Division of Rheumatology, villa Monna Tessa, viale Pieraccini 18, 50139 Firenze, University of Florence, Italy
- **NETHERLANDS**: Ben Dijkmans, MD, VU University medical center and Jan van Breemen Institute, the Netherlands
- **NORWAY**: Tore K Kvien, MD, Head of Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Professor University of Oslo, Norway
- **ROMANIA**: Andra Balanescu M.D. Ph.D., Consultant in Rheumatology and Internal Medicine, Associated Professor, Department of Internal Medicine and Rheumatology, Research Center of Rheumatic Diseases, “Sf. Maria” Hospital, University of Medicine and Pharmacy “Carol Davila”, 37-39 Ion Mihalache Bld., Sector 1, code 011172 Bucharest ROMANIA
- **SPAIN**: Emilio Martin Mola, MD, Head of the Rheumatology Department, Rheumatology Department, Hospital Universitario La Paz, Madrid, Spain
- **TURKEY**: Feride Gogus, MD, and Zafer Gunendi, MD, University of Gazi, Department of Physical Medicine and Rehabilitation, Division of Rheumatology, Ankara, Turkey
- **UNITED KINGDOM**: Professor John Richard Kirwan, MD, Consultant Rheumatologist and Professor of Rheumatic Diseases, University of Bristol, Academic Rheumatology Unit, Bristol Royal Infirmary, Bristol BS2 8HW, UK

LG Validation of RAID modified draft 2 January 2008