LUNDEX, a new index for drug efficacy in clinical practice.
Results from a 5-year observational study in southern Sweden of treatment with infliximab and etanercept in rheumatoid arthritis patients.

Lars Erik Kristensen, Tore Saxne and Pierre Geborek
Department of Rheumatology, Lund University Hospital, SE-221 85 LUND, Sweden

**Corresponding author**
Lars Erik Kristensen
Department of Rheumatology
Lund University Hospital
SE-221 85 LUND, Sweden
E-mail: LarsErik.Kristensen@skane.se
Phone: +46 4617 2287

**Acknowledgements**
We are indebted to all colleagues and staff in the South Swedish Arthritis Treatment Group for cooperation and data supply. This study was supported by grants from Österlund and Kock Foundations, King Gustav V 80 year fund, and Reumatikerförbundet.

**Key Words**
Rheumatoid arthritis; infliximab; etanercept; treatment efficacy
Abstract

Objective. To present and employ LUNDEX, a new index for drug efficacy suitable for comparing long-term efficacy and tolerability of biologic therapies in rheumatoid arthritis (RA) patients treated in clinical practice.

Methods. Patients (n=949) with active RA, not responding to at least 2 DMARDs including methotrexate, initiating biologic therapy for the first time, were included in a structured clinical follow up protocol. The protocol included diagnosis, disease duration, previous and ongoing DMARDs, treatment start and termination. In addition efficacy measures used for calculating validated response criteria, i.e. EULAR and ACR response criteria, were collected at fixed time-points. Data were prospectively registered from March 1999 thru January 2004. To compare efficacy of the different therapies we designed LUNDEX, a new index combining the proportion of patients fulfilling a selected response criterion with the proportion of patients adhering to a particular therapy.

Results. Etanercept had higher overall LUNDEX values compared to infliximab, mostly because of lower adherence to therapy for infliximab. The relationship between the drugs was consistent irrespective of the response criteria used.

Conclusion. LUNDEX is a valuable tool for evaluating drug efficacy in observational studies. It has the advantage of integrating both clinical response as well as adherence to therapy in a composite value. Moreover, LUNDEX has a practical and a potential universal application independent of diagnosis and response criteria.
Introduction
Treatment of rheumatoid arthritis (RA) has undergone remarkable changes over the past few years following introduction of biologic therapies, such as tumor necrosis factor (TNF) blockers and interleukin-1 receptor antagonist. Several randomized controlled clinical trials (RCT’s) have provided documentation for the effectiveness of these drugs (1-14). However, these trials have limitations compared to observational studies. Due to the strict inclusion criteria frequently used, patients enrolled in RCT’s are often limited to RA patients with moderate severity and without any significant co-morbidity. Thus results obtained in RCT’s cannot be uncritically applied to clinical practice, since RA patients are heterogeneous regarding severity, duration, and co-morbidity (15-17) prior to therapy initiation. In other words, the external validity of RCT’s have limitations.

Moreover, RCT’s are often restricted in duration and the number of patients included, which in turn reduces the power for detecting long-term efficacy and tolerability. In addition, rare or co-morbidity-associated side effects are difficult to detect. Finally, comparisons of different biologic therapies are mainly indirect, as RCT’s often compare new drugs to conventional therapy.

Conversely, by using open observational studies following a clinical protocol it is possible to include patients continuously and without limits regarding number or co-morbidity. Furthermore, observational studies allow the inclusion of different treatments in heterogeneous patient groups independent of industry support.

Previously, the concept of `adherence to therapy´, i.e. the number of patients continuing on a drug, has been employed to compare different drugs in observational studies (18-20). However, the `adherence to therapy´ fraction only provide information about the proportion of patients receiving a drug regardless of clinical response. Thus, only a subgroup of patients adhering to a treatment actually experiences a considerable clinical effect. To meet these limitations we introduce LUNDEX, which is the fraction of patients, who not only remain on a particular therapy but also fulfill certain response criteria, such as the American College of Rheumatology response criteria of at least 20% (ACR20) (21).

This study is observational and uses a structured clinical protocol developed by the South Swedish Arthritis Treatment Group for monitoring new biologic therapies in RA (19). The objectives are to present LUNDEX and apply it in the evaluation of long-term efficacy and tolerability of etanercept and infliximab in RA patients treated in clinical practice.

Patients and Methods
The structured clinical protocol was developed from previous nationwide protocols for early RA monitoring, but was modified and extended to make it more suitable for drug monitoring. The inherent element of quality control characterizing the protocol meets the legislative documentation required in Sweden, and therefore no formal approval from the ethical committee was necessary.
Patients. The patients eligible for the study had a diagnosis of RA according to clinical judgement of the treating physician. In a systematic review of 150 patients with a clinical diagnosis, we found 98 % to fulfil the American College of Rheumatology 1987 classification criteria for RA (unpublished data). The patients were treated at 8 centers in southern Sweden serving a population of about 1.3 million individuals during the period March 1999 thru January 2004. Subjects eligible for biologic therapy were selected by physicians based on disease activity and/or unacceptable glucocorticoid use. There were no formal level of disease activity required; however, the patients should have received at least 2 DMARDs including methotrexate without satisfactory response. The selection of particular treatment depended primarily on drug availability. Patients having received biologic therapy prior to inclusion were excluded from this study. The dosage of the different drugs followed the recommendations by the manufacturers. Etanercept 25 mg subcutaneously was administered twice a week, while infliximab was infused at 3 mg/kg at 0, 2, 6, 12 weeks and then every 8th week. Depending on primary or secondary failure the dosage of infliximab could be increased in steps of 100 mg to a maximum of 500 mg administered at 4 to 8 week intervals.

At inclusion, the following data were recorded: Primary diagnosis, other rheumatic diagnoses, previous and concomitant DMARD treatment, and systemic prednisolone dosage.

Method. Clinical data were prospectively collected at 0, 3, 6, 12 months, and subsequently every 3-6 months. No patients were excluded due to lack of registrations at any particular follow-up time. Initially, and at each follow-up the following data were recorded: Health Assessment Questionnaire score, patient scored visual analogue scale for pain and general health, physician’s global assessment of disease activity on a five grade scale (Evalglobal), 28 joint tender and swollen joint count, erythrocyte sedimentation rate (ESR), and C-reactive protein (CRP).

Any withdrawal from treatment was registered prospectively and classified by the treating physician as withdrawal caused by adverse events, lack of response/treatment failure, or miscellaneous. No criteria for inefficacy were predefined, and the decision relied upon the judgment of the treating physician. The category “miscellaneous” mainly consisted of patients with poor compliance or subjects moving away from southern Sweden. In cases where cause of withdrawal was registered as both treatment failure and adverse event (AE), the reason of withdrawal was classified as an adverse event.

To compare efficacy of the different therapies we designed LUNDEX, a new index combining the proportion of patients fulfilling i.e. ACR20 response criteria with the proportion of patients adhering to a particular therapy. LUNDEX is calculated as the fraction of patients adhering to therapy multiplied by the fraction of patients fulfilling a selected response criterion at a given time, see figure 1. Adherence to therapy was calculated using life-table analysis. Improvement in the American College of Rheumatology response criteria of at least 20%, 50% and 70% (ACR20, ACR50 and ACR70 respectively) (21) and the EULAR
responses using 28 joint Disease Activity Score (DAS28) were calculated at
given times of follow up (21).

Statistical analysis. Demographic and baseline clinical characteristics were
analysed by Mann-Whitney U-test for comparison of groups for continuous
variables. Pearson’s Chi-square test was used for discrete or ordinal variables.
Values are reported as the mean ± SD except where stated otherwise.

Adherence to therapy was estimated using Kaplan Meier plots and analysed with
life-table technique using log-rank statistics for comparing different treatments.
ACR and DAS28 responses were analyzed using Pearson’s Chi2 test.
Differences are indicated by p-values where p<0.05 was considered significant.

Results
Baseline data. During the observational period 949 patients initiated etanercept
(n=309) or infliximab (n=640) treatments for the first time. Less than 4% in each
of the different treatment groups had incomplete clinical data at entry (table 1)
and were excluded from efficacy calculations but remained in the survival
analyses.

Demographic data and characteristics of patients enrolled are summarized in
table 1. At baseline several significant differences were found between treatment
groups. Patients receiving infliximab showed significantly lower HAQ score, CRP
level and DAS28 score when compared to etanercept. Also the patients receiving
infliximab had significantly shorter disease duration and fewer previous DMARD’s
when compared to the etanercept group. Furthermore, patients in the infliximab
group were treated with concomitant methotrexate more frequently than patients
in the etanercept group. Both treatment groups consist of patients with
longstanding, therapy resistant, severe RA.

LUNDEX data. LUNDEX was calculated for the treatment groups at 3, 6, 12, 24,
and 36 months for the ACR20, ACR50, EULAR good and EULAR moderate
responders. Data are presented in figure 2A-D. For comparison the figures also
includes the proportion responders at each follow up using the per protocol
technique, i.e. the proportion responders of those actually evaluated. Etanercept
had the highest overall LUNDEX values with nearly 55% of patients started on
etanercept treatment fulfilling ACR20 response criteria during the first year. This
fraction declined to around 40% after 3 years of follow up. On the other hand,
around 45% of patients started on infliximab fulfilled ACR20 response criteria at
12 months, dropping to about 30% after 3 years of follow up.

Adherence to therapy data. Kaplan Meier estimated adherence to therapy for the
treatments is shown in figure 3A. Infliximab had a significantly lower level of
adherence to therapy when compared to etanercept (p<0.001).

Figures 3B-C display the proportion of patients withdrawing from a treatment due
to failure (3B) or adverse events (3C). The main reason for withdrawing from
treatment with infliximab was adverse events (p<0.001), but there was also
significantly larger withdrawal due to treatment failure when compared to
etanercept (p=0.018). The reason for withdrawal from etanercept was equally
distributed between treatment failure and adverse events.
There were no significant differences between the treatments owing to withdrawal for the reason “miscellaneous” (data not shown). The percentages were 1.3 and 4.3 for etanercept and infliximab, respectively. **Response criteria data.** The proportion of patients fulfilling ACR20, ACR50, ACR70 and EULAR DAS28 responses is shown in Table 2. Etanercept showed significantly higher response rates versus infliximab at 3, 6 and 12 months for ACR20 response (p<0.001, p=0.002 and p=0.001, respectively). Also, a statistically significant difference is found for etanercept versus infliximab at 3 months of follow up for DAS28 moderate responders (p=0.033).

The changes in prednisolone dosage show no significant differences between the treatment groups during the observational period (data not shown). Patients with missing efficacy data at certain times of follow up did not show differences in response rates when compared to patients with complete follow up records.

**Discussion**

This study launches LUNDEX as a suitable index for comparing biologic therapies in observational studies. As illustrated by figures 2A-D LUNDEX gives considerably lower values compared to the per protocol technique. Previously, evaluating response rates at fixed times of follow up using intention to treat analyses with last observation carried forward (LOCF) has been used to compare biologic treatment (1-10). Both LOCF and completer (per protocol) analyses inflate the apparent proportions of responses in clinical studies. In observational studies, isolated use of patients fulfilling particular response criteria (Table 2) does not yield information about the true fraction of patients actually responding to a particular therapy, since not all patients are adhering to the different therapies. Therefore, the response rates observed in this study reflect drug performances in selected groups of patients not accounting for differences in drop out among the treatment groups.

In many RCT’s this problem is solved by using intention to treat analysis with carry forward techniques. In observational studies this type of analysis is inappropriate. Patients are continuously entering and exiting the study, and some patients switch to different treatment groups during the observational period. In addition, clinical information necessary for calculating response criteria for dropouts is sometimes missing. To meet this problem we developed LUNDEX, for measuring drug efficacy in RA patients. LUNDEX provides a unifying concept of the fraction of patients adhering to therapy, who truly achieve a specific response criterion after a defined follow up time. It is easy to utilize, and is calculated by multiplying the adherence to therapy proportion with the fraction of patients fulfilling a particular response criterion, as shown in figure 1. In this way, LUNDEX can be applied without having to use intention to treat analysis, and thus facilitates the process of evaluating therapies in clinical observational studies where patients are continuously initiating and stopping therapies. Furthermore, the concept of LUNDEX is not limited to RA and ACR or EULAR response criteria. It is a universal efficacy index, which can be applied to
evaluate drug efficacy in other well-defined diseases with validated response criteria.

In this study, the treatment groups were not exactly matched because of the observational design. Therefore several significant differences were noted at baseline. From a clinical perspective, however, the groups were quite similar. Both patient groups were dominated by patients with long disease duration, failure on several previous DMARDs, and marked disability as well as high disease activity. Therefore, we believe that limited comparisons of the different treatments are justifiable. The present comparison is not hampered by including patients previously treated with biologics, and the indications for anti-TNF therapy as well as concomitant DMARD therapy remained stable during the study period. In this study, it is therefore reasonably clear that patients treated with etanercept have higher LUNDEX values compared to infliximab. The reason for this finding is mainly because of the lower level of drug adherence in patients treated with infliximab. In turn, this lower drug adherence in the infliximab group was mostly explained by withdrawals because of adverse events, but there also seemed to be more failures with this treatment. However, the lower degree of disease activity at baseline in the infliximab group, i.e. lower CRP, DAS28 score, and HAQ scores, may to some extent account for the lower adherence to therapy and LUNDEX values observed in this group, as the potential for improvement is lower in this group. Conversely, the infliximab group also showed a significantly higher proportion of patients receiving concomitant methotrexate at baseline reported to be a more efficacious regimen compared to monotherapy (14).

In accordance to previous reports (22, 23), we did not observe any consistent differences in per protocol response rates between the therapies. Etanercept showed significantly better responses at some points of follow up for ACR20, ACR50 and EULAR (moderate) when compared to infliximab. However, the lack of consistency combined with the heterogeneity of the baseline population makes it less likely that treatment with etanercept truly gives a better clinical response than infliximab.

The open non-randomized nature of this study may induce bias, both in the process of collecting data and during the selection of patients for particular treatments (17, 24). In order to minimize observational bias, all data entries were centralized, thereby ensuring uniform interpretation of registration forms. Confounding by indication cannot be excluded from this study. However, there is only sparse data directly comparing the different biologic drugs (24), thus giving no obvious reason for favoring prescription of one drug over another. A placebo effect improving the response to the drugs may be expected. However, there is no reason to believe that this effect is distributed unevenly between treatment groups. Actually, the treatments did in fact show lower response rates in this study when compared to previous controlled clinical trials (1, 2, 6-10, 13, 14). This can be explained by the natural variety of glucocorticoid usage and patients encountered in the clinical setting who are included in this
study. Also the long observational period of this study better reflects the chronic course of RA, and thus dilutes the bias occurring in many RCT's when observing treatment of flares (17).

Finally, we expect LUNDEX to become a valuable tool when evaluating results of observational studies in the future, due to the practical and potential universal application of this measure independent of diagnoses and response criteria.
References


Table 1. Demographic and clinical characteristics at baseline. Values are the mean +/- SD except where stated otherwise. MTX = methotrexate; HAQ = health assessment questionnaire; DMARD’s = disease-modifying antirheumatic drugs.

*Number of patients with complete data at entry.

<table>
<thead>
<tr>
<th></th>
<th>Etanercept I (n = 309)</th>
<th>Infliximab II (n= 640)</th>
<th>Level of significant differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female (no. (%))</td>
<td>253 (82%)</td>
<td>481 (75%)</td>
<td>p=0.021</td>
</tr>
<tr>
<td>Age (ys)</td>
<td>55.1 (±13)</td>
<td>56.2 (±14)</td>
<td></td>
</tr>
<tr>
<td>Disease duration (ys)</td>
<td>14.7 (±10.1)</td>
<td>12.7 (±10.0)</td>
<td>P&lt; 0.001</td>
</tr>
<tr>
<td>Mean number of previous DMARD’s including MTX</td>
<td>4.2 (± 2.05)</td>
<td>3.6 (± 1.98)</td>
<td>p&lt; 0.001</td>
</tr>
<tr>
<td>MTX at inclusion, (no. (%)</td>
<td>96 (31%)</td>
<td>467 (73%)</td>
<td>p&lt; 0.001</td>
</tr>
<tr>
<td>Weekly MTX dosage of patients receiving MTX (mg)</td>
<td>15.7 (± 5.1)</td>
<td>14.0 (± 5.92)</td>
<td></td>
</tr>
<tr>
<td>DAS28 (0-10)</td>
<td>5.9 (± 1.06)</td>
<td>5.6 (± 1.20)</td>
<td>p&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>(n=297)*</td>
<td>(n=615)*</td>
<td></td>
</tr>
<tr>
<td>HAQ score (0-3)</td>
<td>1.6 (± 0.64)</td>
<td>1.4 (± 0.62)</td>
<td>p= 0.002</td>
</tr>
<tr>
<td></td>
<td>(n=301)*</td>
<td>(n=622)*</td>
<td></td>
</tr>
<tr>
<td>C-reactive protein, mg/dl (&lt;0.8)</td>
<td>4.0 (± 3.66)</td>
<td>3.6 (± 3.73)</td>
<td>p= 0.044</td>
</tr>
<tr>
<td></td>
<td>(n=298)*</td>
<td>(n=628)*</td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Response criteria at follow up times grouped according to biologic treatment. Values are shown as percentages of patients fulfilling the particular response criteria at follow up times 3, 6, 12, 24 and 36 month respectively.

<table>
<thead>
<tr>
<th></th>
<th>Etanercept</th>
<th>Infliximab</th>
<th>Levels of significant differences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACR20</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 month</td>
<td>63</td>
<td>45</td>
<td>P&lt;0.001</td>
</tr>
<tr>
<td>6 month</td>
<td>61</td>
<td>47</td>
<td>P=0.002</td>
</tr>
<tr>
<td>12 month</td>
<td>69</td>
<td>53</td>
<td>P=0.001</td>
</tr>
<tr>
<td>24 month</td>
<td>65</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>36 month</td>
<td>63</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td><strong>ACR50</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 month</td>
<td>38</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>6 month</td>
<td>34</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>12 month</td>
<td>44</td>
<td>32</td>
<td>P=0.011</td>
</tr>
<tr>
<td>24 month</td>
<td>39</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>36 month</td>
<td>39</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td><strong>ACR70</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 month</td>
<td>9</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>6 month</td>
<td>11</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>12 month</td>
<td>14</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>24 month</td>
<td>19</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>36 month</td>
<td>16</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td><strong>EULAR (moderate)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 month</td>
<td>51</td>
<td>37</td>
<td>P=0.033</td>
</tr>
<tr>
<td>6 month</td>
<td>48</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>12 month</td>
<td>51</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>24 month</td>
<td>44</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>36 month</td>
<td>46</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td><strong>EULAR (good)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 month</td>
<td>29</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>6 month</td>
<td>27</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>12 month</td>
<td>33</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>24 month</td>
<td>38</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>36 month</td>
<td>36</td>
<td>45</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1 displays the equation for LUNDEX.

\[
\text{LUNDEX} = (\text{Fraction of starters still in the study at time T}) \times (\text{Fraction responding at time T})
\]

Figure 2A-D illustrates LUNDEX using ACR20 (2A), ACR50 (2B), EULAR moderate plus good (2C), or EULAR good (2D) responders, for the treatment groups at different times of follow up. Also the respective responder proportions are included as blank columns.

- LUNDEX
- Responders actual follow up (per protocol analysis)

Figure 3. Adherence to therapy according to treatment group.
(A) withdrawal from therapy due to any reason.
(B) withdrawal due to adverse events.
(C) withdrawal due to failure of treatment.
ACR20

Months of follow up

Percent

- etanercept
- infliximab
- adalimumab
- anakinra

responders actual follow up
LUNDEX - responders of all initiated treatments