# Chapter 7

## Assessment of global disease activity in RA patients monitored in the METEOR database: The patient's versus the rheumatologist's opinion.

E. Gvozdenović
R. Koevoets
R. Wolterbeek
D. van der Heijde
T.W.J. Huizinga
C.F. Allaart
R. Landewé.

## ABSTRACT

**Objectives**

To compare the patient's (PtGDA) and physician's (PhGDA) assessment of global disease activity and to identify factors that might influence these differences, as well as factors that may influence the patients and the physicians score separately.

**Methods**

Anonymous data were used from 2118 Dutch patients included in the METEOR database. PtGDA and PhGDA were scored independently on a 100mm visual analogue scale (VAS) with 0 and 100 as extremes. The agreement, Intraclass correlation coefficients (ICC), was calculated and a Bland Altman plot was created to visualize the differences between PtGDA and PhGDA. Linear Mixed Model analysis was used to model PtGDA and PhGDA. Logistic repeated measurements were used to model the difference in PtGDA and PhGDA (PtGDA>PhGDA vs. PtGDA≤PhGDA). Gender, age, swollen joint count, tender joint count, VAS pain, disease duration and ESR were considered as possible determinants in both models.

**Results**

Mean (SD) age was 57 (15) years and 67% of the patients were female. Agreement between PtGDA and PhGDA was moderate (ICC: 0.57). Patients scored on average 11 units higher (worse) than rheumatologists (95% limits of agreement: -25.2 to 47.6). Patient's perception of pain (VAS) was positively associated with a PtGDA being higher than PhGDA. Similarly, ESR and swollen joint counts were positively associated with a PtGDA being lower or equal to the PhGDA.

**Conclusion**

Patients rate global disease activity consistently higher than their rheumatologists. Patients base their judgment primarily on the level of pain; physicians on the level of SJC and ESR.

## INTRODUCTION

The importance and use of patient reported outcomes (PROs) in health care increased during the past decades. PROs are considered valuable in measuring status and change in health care [1]. However, in addition to the PRO, similar information is also collected by the physician, e.g. assessment of level of disease activity. As patients and physicians may differ in their perception of health status, discordant observations may occur and may affect patient care. For example, patients are likely to report dissatisfaction with a treatment if their physician underestimates their perceived level of disease activity [2-4].

The 100mm visual analogue scale (VAS) is an instrument used to measure global disease activity (GDA) in rheumatoid arthritis (RA). It can be completed by the patient (PtGDA) (and is then considered then a PRO) as well as by the physician (PhGDA). Discordances between patients and rheumatologists rating their impression of GDA on a VAS have been reported; patients tend to score their GDA higher than their physician [5,6]. It is not clear which factors determine the occurrence and magnitude of these discrepancies between patient's and physicians' perceptions.

The METEOR (Measurement of efficacy of Treatment in the Era of Rheumatology) database provides data on several patient- and physician-reported outcome measures in RA. Here we have compared PtGDA and PhGDA reported in individual patients, and identified which factors determined the discordance in PtGDA and PhGDA.



## METHODS

### *Patients*

Data collected in the ongoing prospective international METEOR database were used. METEOR is an acronym for Measurement of efficacy of Treatment in the Era of Rheumatology hat has been started in 2008. METEOR is used by rheumatologists to monitor patients with rheumatic diseases. Data are collected in a central database in a completely anonymous way. Both newly diagnosed patients and patients with more advanced disease are included in de database. Measures of disease activity and Health Assessment Questionnaire data are registered every visit. Currently, the tool is used worldwide and data is available from 100 hospitals, which included more than 14.800 patients. More details on the METEOR database are described elsewhere [9] .

A sample of 2.118 patients was taken from the METEOR database covering the time span between 2008 and 2011. The number of visits (8.509 in total) varied with a range of 1 to 17 visits per patient as did time intervals between visits. PtGDA and PhGDA were measured on a 100mm visual analysis scale (VAS) with 0 (best possible) and 100 (worst possible) as extremes. PtGDA and PhGDA separately were operationalized as continuous variables. The

20mm difference between PtGDA and PhGDA was used as a binary outcome variable (patient scores higher versus rheumatologist scores equal or higher). A difference in rating of 20mm between PtGDA and PhGDA score was chosen as cutoff value, since it was considered to be a relevant discordance in previous literature[5].

*Statistical analyses*

Descriptive statistics were performed using the mean and standard deviation (SD) or median and interquartile ranges (IQR) as appropriate for continuous variables, and number and percentages for categorical variables.

A Bland and Altman plot was performed to visualize the differences between PtGDA and PhGDA. This is based on the standard deviation of the differences in PtGDA and PhGDA calculated from variance components in a linear mixed model (LMM), and used to construct the 95% limits of agreement [7]. The agreement between patient and physician was expressed as intraclass correlation coefficient (ICC) using variance components in a LMM with a random intercept for patients.

LMM was also used to model the PtGDA and PhGDA. Gender, age, swollen joint count, tender joint count, pain (VAS), disease duration (diagnosis until first visit) and erythrocyte sedimentation rate (ESR) were considered as possible determinants for the model.

Non-linear mixed modeling (repeated measures logistic regression) was used to model the difference in PtGDA and PhGDA as binary outcome (patient's score higher than physician's score as "event"). Gender, age, swollen joint count, tender joint count, pain (VAS), disease duration and ESR were considered as possible determinants for the model.

Software programs SAS version 9.2 and SPSS version 17.0 were used for the analyses and P-values smaller than 0.05 were considered statistically significant.

## RESULTS

Of the 2118 patients, 1338 (67%) were female. The mean (SD) age at entry was 57 [15] years (table 1).

Agreement between PtGDA and PhGDA was moderate (ICC: 0.57; $p<0.01$). Patients rated their GDA on average 11mm higher (worse) than rheumatologists at the first registered visit (95% limits of agreement: -25.2 to 47.6). A few scores showed large discrepancy between the PtGDA and PhGDA score, on average 75 mm (figure 1). Patients scored the GDA significantly higher when the number of tender joint count, and VAS pain increased ($p<0.01$). VASpain ($p<0.01$), number of swollen and tender joint count ($p=0.04$ and ESR ($p<0.01$) independently contributed to an increase in GDA score by the physician. Physician's score decreased by increasing disease duration ($p=0.03$) and patient's scores increased by decreasing swollen joint count, ($p=0.04$) (table 2).

Pain (VAS), ESR and the number of swollen joints all independently contributed to the differences between patient's GDA and physician's GDA. A higher patient GDA score compared to the physicians score is positively correlated with pain (VAS). A higher or equal GDA score of the physician compared to the patient is positively correlated with ESR and swollen joint count. (Table 3)

**Table 1: Baseline characteristics (Visit 1)**

| Variables | Patients | N total (N=2118) |
|---|---|---|
| Age, mean (SD) | 57 (15) | 1879 |
| Female, N (%) | 1338 (67) | 2007 |
| CRP, median (IQR) | 5 (3-13) | 167 |
| ESR, median (IQR) | 14 (6-29) | 1489 |
| DAS 28, mean (SD) | 3.2 (1.4) | 1408 |
| HAQ, mean (SD) | 0.9 (0.3) | 575 |
| Duration complaints until diagnosis (mo), median (IQR) | 4 (1-12) | 758 |
| Duration complaints until first registered visit (yrs), median (IQR) | 7 (2-15) | 775 |
| Duration diagnosis until first registered visit (yrs )median (IQR) | 6 (1-13) | 790 |
| CCP positive, N (%) | 212 (64) | 334 |
| RF positive, N (%) | 726 (74) | 987 |
| Erosions present, N (%) | 596 (65) | 923 |
| Swollen joint count 28, median (IQR) | 1 (0-3) | 1872 |
| Tender joint count 28, median (IQR) | 2 (0-4) | 1872 |
| VAS, median (IQR) | | |
| Global health physician | 21 (10-41) | 903 |
| Global health patient | 34 (14-55) | 1615 |
| Pain patient | 39 (15-60) | 1476 |

N: number; SD: standard deviation; CRP: C-reactive protein; ESR: erythrocyte sedimentation rate; IQR: inter quartile range; DAS: disease activity score; HAQ: health assessment questionnaire; Mo: months; yrs: years; CCP: anti cyclic citrullinated peptides antibodies; RF: rheumatoid factor; VAS: visual analogue scale

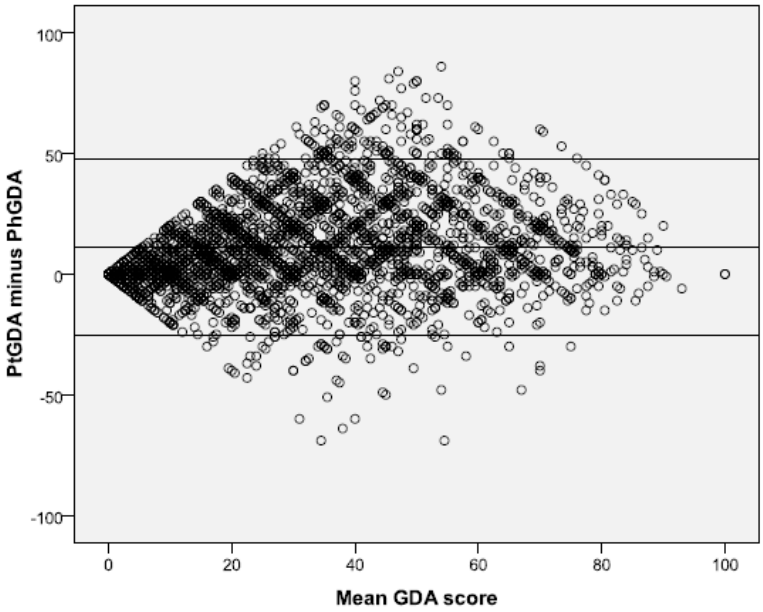


**Figure 1:** Bland Altman's plot: GDA patient versus GDA physician
GDA patient: global disease activity according to the patient; GDA physician: global disease activity according to the physician; PtGDA: global disease activity according to the patient; PhGDA: global disease activity according to the physician

**Table 2:** Linear mixed model predictors of GDA disease activity by patients and physicians

| | PtGDA | | PhGDA | |
|---|---|---|---|---|
| **Variable** | **β Estimate, 95% CI** | **p-value** | **β Estimate, 95% CI** | **p-value** |
| Male | -0.73<br>-2.58, 1.12 | 0.44 | -0.25<br>-2.55, 2.05 | 0.07 |
| Age | -0.03<br>-0.10, 0.03 | 0.33 | -0.07<br>-0.15, 0.01 | 0.08 |
| Disease duration (yrs) | -0.02<br>-0.11, 0.08 | 0.71 | -0.10<br>-0.17, -0.03 | 0.03 |
| ESR | 0.04<br>-0.00, 0.09 | 0.07 | 0.09<br>0.05, 0.12 | <0.01 |
| SJC28 | -0.51<br>-0.94, -0.09 | 0.02 | 0.53<br>0.02, 1.04 | 0.04 |
| TJC28 | 0.74<br>0.43, 1.05 | <0.01 | 0.41<br>0.03, 0.80 | 0.04 |
| VAS pain patient | 0.45<br>0.41, 0.49 | <0.01 | 0.20<br>0.15, 0.25 | <0.01 |

PtGDA: global disease activity according to the patient; PhGDA: global disease activity according to the physician; β: beta; CI: confidence interval; yrs: years; ESR: erythrocyte sedimentation rate; SJC28: swollen joint count in 28 joints; TJC28: tender joint count in 28 joints; VAS: visual analogue scale

**Table 3:** Non-linear mixed model predictors of GDA difference between patients and physicians

| | PtGDA versus PhGDA[a] | |
|---|---|---|
| **Variable** | **β Estimate, 95% CI** | **p-value** |
| Male | 0.18<br>-0.12, 0.48 | 0.24 |
| Age | -0.00<br>-0.01, 0.00 | 0.38 |
| Disease duration (yrs) | 0.01<br>-0.01, 0.02 | 0.26 |
| ESR | -0.01<br>-0.00, -0.02 | <0.01 |
| SJC28 | -0.29<br>-0.37, -0.20 | <0.01 |
| TJC28 | -0.04<br>-0.10, 0.02 | 0.17 |
| VAS pain patient | 0.05<br>-0.06, 0.06 | <0.01 |

PtGDA: global disease activity according to the patient; PhGDA: global disease activity according to the physician; β: beta; CI: confidence interval; yrs: years; ESR: erythrocyte sedimentation rate; SJC28: swollen joint count in 28 joints; TJC28: tender joint count in 28 joints; VAS: visual analogue scale; [a] 1=patient scores higher; 0=physician scores higher or equal; reference category=0



## DISCUSSION

On average, patients tend to score GDA systematically higher than rheumatologist. The agreement between both is only moderate. Determinants of the differences in scores are pain (more pain means higher GDA by the patient), and swollen joint count as well as ESR (higher SJC/ESR means higher GDA by physician). PtGDA and PhGDA separately are partly associated with the same determinants: tender joint count and pain are both taken into consideration by the patients and physicians assessment of GDA. In addition, the objective measures, swollen joint count and ESR, are taken into consideration by the physician. Expectedly, physicians put more weight on the value of ESR and SJC, whilst patients put more weight on pain.

Patients and physicians take partly the same determinants in consideration when they assess global disease activity; they both consider tender joints and pain of the patient. However the patient does not take into consideration objective measures. In fact the swollen joint count even has a negative association with PtGDA which we cannot explain. The physician takes both objective (swollen joint count, acute phase reactants and disease duration) and subjective (patients pain and tender joint count) measures in consideration in assessing the GDA. The discrepancy in factors both patients and physicians take into consideration for their

GDA assessment might lead to the systematic difference in patients and physicians scores of almost 11 units (on a scale from zero to 100) and to the only moderate agreement between patients and physicians. Other studies also reported discordances between patients and physicians in rating the GDA. Barton et al. showed that patients' GDA score was on average 15 points higher than the physicians' mean GDA score [8]. Also, the QUEST-RA study showed a higher mean GDA of patients (approximately 11 points) than GDA of physicians [5]. In concordance with the latter study, we also found a difference of approximately 11 points. However, it is questionable if 11 points is a clinical relevant discrepancy between patient' and physician' GDA score since we defined 20 points to be a difference. On the other hand, the moderate agreement between patients and physicians might support that patients and physicians rate RA disease activity differently. This confirms the statement of an earlier study that patient and physicians differ in perception of disease activity [6].

A previous study, carried out in several European countries, also showed only a moderate agreement between GDA patient and GDA physician [5]. Other studies, performed in the United States and in Europe showed low correlations and low agreement between physician and global health assessments [9, 10]. The discrepancies between the results of previous studies might suggest differences between countries in GDA of patient and physician due to cultural factors.

Our study shows that the difference in scoring might be explained by differences in the interpretation of ESR, swollen joints and pain. Pain is more likely to be associated with an equal or higher score of the patient. This statement was confirmed by the large QUEST-RA study, which studied factors on discordance between GDA of the patient and that of the physician. Pain was one of the most important factors that caused discordances. Pain increased significantly when patient scored GDA higher compared to the physician. Furthermore, the QUEST-RA also used 20mm difference in GDA score as the cut off value of a true difference between patient and physician [5].

In our study, patients with a high ESR and swollen joint count are more likely to be scored higher by the physician. A previous study confirms this result. [8]. Another study showed that, besides swollen joints, physician put more weight on ESR than patients [6].

As we can see from the results of our study, patients and physicians focus on different factors when assessing disease activity. Patients are more influenced by subjective feelings, such as pain, while physicians base their score more on objective measures, such as number of swollen joints and 'blood levels'. This is supported by previous literature [11]. Patients base their assessments on needs, priorities, experiences, expectations and attitude, which are all subjective domains. Physicians, on the other hand, rely on the patient's physical health status, which is considered more objective in nature [12, 13] .

This study has some limitations. The first is missing values, as these might not be randomly missing. Patients that perform worse in their opinion may stay at home and miss an appointment with the physician. This can result in selection of patients with unknown consequences.

Another limitation is that the included patients were not always newly diagnosed RA patients. Some patients are already treated for years and patients expectations and perceptions can change as a result of improvement or worsening of their health [14]. Therefore, long treatment duration might influence patient's assessment of GDA.

In conclusion, patients and physicians both assess GDA using partly similar determinants. Differences in GDA scores may be explained by pain, ESR and swollen joint count. Patients put more weight on pain and physicians on ESR and swollen joint count. Also cultural differences may have contributed to the moderate level of agreement between patients and physicians. We already see a difference in agreement between patient's and physician's score by comparing studies performed in several countries.

In clinical practice, it should be recommended to spend more time educating patients on how to rate the global disease activity. Patients need to be clearly informed on the difference between the disease activity and pain, as patients let pain influence their GDA score. A good understanding of the GDA score by the patient is important since a previous study showed that patients with a high PtGDA score, while having a normal ESR and low SJC and TJC, are not in remission [15].

Further research should be conducted to find out what the clinical impact is of these discrepancies between patients and physicians since previous research might suggest that treatment strategy is only based on the rheumatologist's opinion and not on the patient's opinion or the DAS28 [16 ]. Also differences in PtGDA and PhGDA score per country should be studied and whether GDA assessment is influenced by cultural factors.



## ACKNOWLEDGEMENTS

METEOR is a free-for-use online software program developed by the Merit Foundation and aims to improve treatment of RA by the measurement of patient outcomes and benchmarking data in a multi-national database. Data is securely sent and stored anonymously.

## REFERENCE LIST


1. Otter, S.J., Lucas, K., Springett, K., et al., *Identifying Patient-Reported Outcomes in Rheumatoid Arthritis: The Impact of Foot Symptoms on Self-Perceived Quality of Life.* Musculoskeletal Care, 2012.
2. Hewlett, S.A., *Patients and clinicians have different perspectives on outcomes in arthritis.* The Journal of Rheumatology, 2003. 30(4): p. 877-879.
3. Suarez-Almazor, M.E., et al., *Lack of congruence in the ratings of patients' health status by patients and their physicians.* Medical Decision Making, 2001. 21(2): p.113- 121.
4. Wartman, S.A., Morlock, L.L., Malitz, F.E., Palm, E., *Impact of divergent evaluations by physicians and patients of patients' complaints.* Public Health Reports, 1983. 98(2): p. 141.
5. Khan, N.A., Spencer, H.J., Abda, E., et al., *Determinants of discordance in patients' and physicians' rating of rheumatoid arthritis disease activity.* Arthritis Care & Research, 2012. 64(2): p. 206-214.
6. Nicolau, G., Yogui, M.M., Vallochi, T.L., Gianini, R.J., Laurindo, I.M.M., Novaes, G.S., *Sources of discrepancy in patient and physician global assessments of rheumatoid arthritis disease activity.* The Journal of Rheumatology, 2004. 31(7): p. 1293-1296.
7. Foti, C., Cisari, C., Carda, S., et al., *A prospective observational study of the clinical efficacy and safety of intra-articular sodium hyaluronate in synovial joints with osteoarthritis.* European journal of physical and rehabilitation medicine, 2011. 47(3): p. 407.
8. Barton, J.L., Imboden, J., Graf, J., Glidden, D., Yelin, E.H., Schillinger, D., *Patient-physician discordance in assessments of global disease severity in rheumatoid arthritis.* Arthritis Care & Research, 2010. 62(6): p. 857-864.
9. Kwoh, C.K., O'Connor, G.T., Regan-Smith, M.G., et al., *Concordance between clinician and patient assessment of physical and mental health status.* The Journal of Rheumatology, 1992. 19(7): p. 1031.
10. Ward, M., *Clinical measures in rheumatoid arthritis: which are most useful in assessing patients?* The Journal of Rheumatology, 1994. 21(1): p. 17.
11. Neville, C., Clarke, A.E., Joseph, L., Belisle, P. Ferland, D., Fortin, P.R., *Learning from discordance in patient and physician global assessments of systemic lupus erythematosus disease activity.* The Journal of Rheumatology, 2000. 27(3): p.675.
12. Liang, M., K. Cullen, and M. Larson, *In search of a more perfect mousetrap (health status or quality of life instrument).* The Journal of Rheumatology, 1982. 9(5): p. 775.
13. Pincus, T., Summey, J.A., Soraci, S.A., Wallston, K.A., Hummon, N.P., *Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire.* Arthritis & Rheumatism, 1983. 26(11): p. 1346-1353.
14. Ward, M.M., *Rheumatology care, patient expectations, and the limits of time.* Arthritis Care & Research, 2004. 51(3): p. 307-308.
15. Vermeer, M., Kuper, H.H., Bijl, van der, A.E., et al., *The provisional ACR/EULAR definition of remission in RA: a comment on the patient global assessment criterion.* Rheumatology, 2012.
16. Pyne, L., Bykerk., Boire., et al., *Increasing Treatment in Early Rheumatoid Arthritis Is Not Determined by the Disease Activity Score But by Physician Global Assessment: Results from the CATCH Study.* The Journal of Rheumatology, 2012.