Chapter 4

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.

Clinical Rheumatology 2014 Apr; 33(4):461-6

ABSTRACT

Objective

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 2.117 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 patient, gender physician, 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 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. Determinants reported to be of influence on the discrepancies between patients' and physicians' perceptions are pain, swollen joint count, tender joint count and erythrocyte sedimentation rate.[5-7] However, the magnitude and direction of the influence of these factors is unclear. A determinant, to our knowledge not studied yet, which might be of influence on the difference between physicians' and patients' perception is the gender of the physician. This might be a plausible factor of difference in score since male and female physician perceptions differ in clinical practice regarding communication of information, compliance and satisfaction of the patient.[8]

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, including gender of the rheumatologist. 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 2.117 Dutch 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 19 visits per patient as did time intervals between visits. PtGDA and PhGDA were measured on a 100mm visual analogue 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 cut-off value, since it is considered to be a frequent chosen value for minimum clinically important improvement in PtGDA.[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.[9] 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 patient, gender rheumatologist, 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. Furthermore LMM was used to estimate means of DAS28, ESR, tender and swollen joint count between male and female. Non-linear mixed modelling (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 patient, gender rheumatologist, 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 2.117 patients, 1.338 (67%) were female. The mean (SD) age at entry was 57 (15) years (table 1).

Table 1. Baseline characteristics (Visit 1)

Variables	Patient	n total (n=2.117)
Age (years), mean (SD)	57 (15)	1879
Patient female, n (%)	1339 (67)	2007
Physician female, n (%)	1072 (67)	1598
CRP, median (IQR)	5 (3 to 13)	167
ESR, median (IQR)	14 (6 to 29)	1491
DAS 28, mean (SD)	3.2 (1.4)	1408
HAQ, median (IQR)	0.7 (0.1 to 1.4)	573
Duration complaints until diagnosis (months), median (IQR)	4 (1 to 12)	855
Duration complaints until first visit in METEOR (years), median (IQR)	6 (1 to 14)	862
Duration diagnosis until first visit in METEOR (years), median (IQR)	3 (0 to 11)	996
CCP positive, n (%)	215 (64)	334
RF positive, n (%)	734 (77)	959
Erosions present, n (%)	605 (66)	923
Swollen joint count 28, median (IQR)	1 (0 to 3)	1799
Tender joint count 28, median (IQR)	2 (0 to 4)	1799
VAS (visual analogue scale), median (IQR)		
Global disease activity physician	21 (10 to 41)	903
Global disease activity patient	34 (14 to 55)	1615
Pain patient	39 (15 to 60)	1474

n=number, SD=standard deviation, IQR= interquartile range, CRP=C-reactive protein, ESR= erythrocyte sedimentation rate, DAS28=Disease Activity Score 28 joints, HAQ= Health Assessment Questionnaire, CCP=cyclic cictrullinated ceptide antibody, RF=rheumatoid factor.

978 of the observed patient scores were higher (20mm) compared to the physicians score, 2.747 of patients and physicians score where concordant (-20 mm until 20 mm) and 102 patients score were lower (20 mm) than the physician's score.

Agreement between PtGDA and PhGDA was moderate (ICC: 0.57; p<0.01; 95% limits of agreement: -25.2 to 47.6). Patients rated their GDA on average 11mm higher (worse) than rheumatologists at the first registered visit. A few scores (n=19) showed a discrepancy between the PtGDA and PhGDA of 70 or more (patient scored higher) (figure 1).

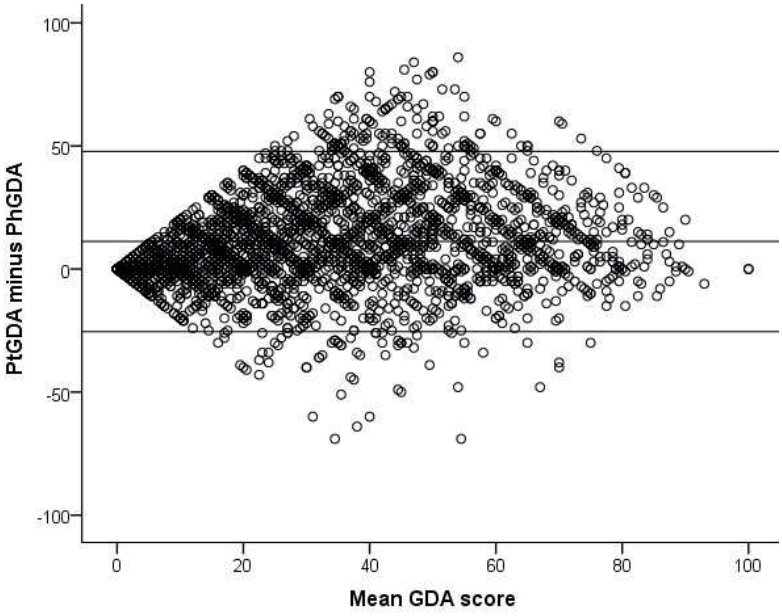

Figure 1. Bland and Altman's plot; global disease activity patient (PtGDA) versus global disease activity physician (PhGDA)

Both patients and physicians scored the GDA significantly higher when the number of tender joints, number of swollen joints, and VAS pain was higher (p<0.01). Furthermore, a higher ESR (p<0.01) and male gender (p=0.02) were independently associated with a higher GDA score by the physician. Physician's scores decreased by increasing disease duration (p=0.01). The gender of the physician was not associated with the GDA score by physician or patient (table 2). Pain (VAS), ESR and the number of swollen joints all independently were

Table 2. Linear mixed model predictors of global disease activity by patients (PtGDA) and physicians (PhGDA)

	PtGDA		PhGDA	
Variable	**Estimate β, (95% CI)**	**p-value**	**Estimate β, (95% CI)**	**p-value**
Patient male	0.82 (-0.47 to 2.11)	0.21	1.86 (0.32 to 3.39)	0.02
Physician male	0.58 (-0.55 to 1.71)	0.31	1.21 (-0.18 to 2.60)	0.09
Age (years)	0.01 (-0.04 to 0.05)	0.82	-0.05 (-1.00 to 0.01)	0.08
Disease duration (years)	-0.05 (-0.12 to 0.01)	0.12	-0.10 (-0.17 to -0.02)	0.01
ESR	0.02 (-0.01 to 0.05)	0.82	0.11 (0.07 to 0,14)	<0.01
Swollen joint count 28	0.87 (0.60 to 1.15)	<0.01	3.24 (2.91 to 3.57)	<0.01
Tender joint count 28	0.41 (0.22 to 0.61)	<0.01	0.75 (0.49 to 1.01)	<0.01
VAS pain patient	0.72 (0.69 to 0.74)	<0.01	0.29 (0.26 to 0.32)	<0.01

CI=Confidence interval, β = beta, VAS=visual analogue scale, ESR= erythrocyte sedimentation rate.

associated with the difference between patient's GDA and physician's GDA score. Patient scored GDA higher than their physician by increasing VAS pain (p<0.01); and physician scored GDA higher than the patient by increasing swollen joint count and ESR (p<0.01). Gender of the patient or gender of the physician did not have an effect on the difference between patient's GDA and physician's GDA score (table 3).

Table 3. Non-linear mixed model predictors of global disease activity (GDA) difference between patients and physicians.

	PtGDA (n=978) versus PhGDA (2747)*	
Variable	**Estimate β, (95% CI)**	**p-value**
Patient male	-0.06 (-0.49 to 0.38)	0.79
Physician male	0.17 (-0.23 to 0.59)	0.40
Age (years)	0.00 (-0.01 to 0.02)	0.42
Disease duration (years)	0.00 (-0.02 to 0.02)	0.65
ESR	-0.02 (-0.03 to -0.00)	<0.01
SJC28	-0.38 (-0.53 to -0.22)	<0.01
TJC28	-0.06 (-0.14 to 0.12)	0.14
VAS pain patient	0.08 (0.05 to 0.11)	<0.01

*1= patient scores higher, 0= physician scores equal or higher; Reference category=0, VAS=visual analogue scale, ESR= erythrocyte sedimentation rate, CI=Confidence interval, β=beta.

Mean ESR and TJC28was lower in male patients compared to female patients (p=0.02). Also DAS28 was lower in male patients (p<0.01) (table 4).

Table 4. Linear mixed models for means of DAS and DAS components between man and women

	Male	**Female**	
	Mean (95% CI)	**Mean (95% CI)**	**p-value**
DAS28	2.8 (2.7 to 2.9)	3.1 (3.07 to 3.2)	<0.01
ESR	17.9 (16.5 to 19.3)	20.0 (19.3 to 21.0)	0.02
Swollen joint count 28	1.6 (1.5 to 1.8)	1.5 (1.42 to 1.63)	0.38
Tender joint count 28	2.3 (2.1 to 2.5)	2.6 (2.5 to 2.8)	0.02

DAS28=Disease Activity Score for 28 joints, ESR= erythrocyte sedimentation rate, CI=confidence interval.

DISCUSSION

On average, patients tend to score GDA systematically higher than rheumatologists. The agreement between patients and rheumatologists is only moderate. Physicians and patients both take into consideration tender joint count, swollen joint count and pain in their assessment of GDA. In addition, when rating GDA, the physician is influenced by the gender of the patient, disease duration and ESR. The difference in GDA score between patient and physician can best be explained by differences in pain, swollen joint count and ESR. 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. The physician takes both ‘objective’ (swollen joint count, acute phase reactants and disease duration) and ‘subjective’ (patients’ pain and tender joint count) variables into account when assessing the GDA. Furthermore, physicians tend to rate GDA in male patients higher than in female patients. The latter finding might be related to the difference in perception of disease activity between male and female patients, since male patients tend to underestimate their disease activity compared to female patients.[7] In our study male patients have indeed lower DAS and fewer tender joints when compared to female patients, while the number of swollen joints does not differ. This finding provides

input to the suggestion that the physician may implicitly compensate for this difference by rating disease activity in higher than in female.

The different factors patients and physician taken into consideration for their GDA assessment might be the explanation for 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 have 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.[10]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.[11,12] 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 perception 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.[10] 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.[13] 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.[14,15]

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.[16] 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.[17] 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.[18] Also differences in PtGDA and PhGDA score per country should be studied and whether GDA assessment is influenced by cultural factors.

REFERENCE LIST

1 Otter SJ, Lucas D, Springett K, et al. Identifying Patient-Reported Outcomes in Rheumatoid Arthritis: The Impact of Foot Symptoms on Self-Perceived Quality of Life. *Musculoskelet. Care* 10 (2012); 65–75.

2 Hewlett SA. Patients and clinicians have different perspectives on outcomes in arthritis. *J Rheumatol.* 2003; 30 (4) :877-9.

3 Suarez-Almazor ME, Conner-Spady B, Kendall CJ, et al. Lack of congruence in the ratings of patients' health status by patients and their physicians. *Med Decis Making*. 2001; 21(2): 113-21.

4 Wartman SA, Morlock LL, Malitz FE et al. Impact of divergent evaluations by physicians and patients of patients' complaints. *Public Health Rep*. 1983; 98(2):141-5

5 Khan NA, Spencer HJ, Abda E, et al. Determinants of discordance in patients' and physicians' rating of rheumatoid arthritis disease activity. *Arthritis Care Res (Hoboken)*. 2012; 64(2): 206-14.

6. Nicolau G, Yogui MM, Vallochi TL, et al. Sources of discrepancy in patient and physician global assessments of rheumatoid arthritis disease activity. *J Rheumatol.* 2004; 31(7): 1293-6.

7. Sokka T, Toloza S, Cutolo M, et al. Women, men, and rheumatoid arthritis: analyses of disease activity, disease characteristics, and treatments in the QUEST-RA study. *Arthritis Res Ther.* 009; 11(1): R7

8 Weisman CS,Teitelbaum MA. Physician gender and the physician-patient relationship: recent evidence and relevant questions. *Soc Sci Med.* 1985; 20(11):1119-27.

9 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. *Eur J Phys Rehabil Med.* 2011; 47(3):407-15

10 Barton JL, Imboden J, Graf J, et al. Patient-physician discordance in assessments of global disease severity in rheumatoid arthritis. *Arthritis Care Res (Hoboken)*. 2010; 62(6):857-64

11 Kwoh CK, O'Connor GT, Regan-Smith MG, et al. Concordance between clinician and patient assessment of physical and mental health status. *J Rheumatol.* 1992;19(7):1031-7

12 Ward M. Clinical measures in rheumatoid arthritis: which are most useful in assessing patients? J *Rheumatol.* 1994; 21(1):17-27

13 Neville C, Clarke AE, Joseph L, et al. Learning from discordance in patient and physician global assessments of systemic lupus erythematosus disease activity. *J Rheumatol.* 2000; 27(3) :675-9.

14 Liang M, Cullen K and Larson M. In search of a more perfect mousetrap (health status or quality of life instrument). *J Rheumatol*. 1982;9(5): 775-9.

15 Pincus T, Summey JA, Soraci SA Jr, et al. Assessment of patient satisfaction in

activities of daily living using a modified Stanford Health Assessment Questionnaire. *Arthritis Rheum.* 1983; 26(11): 1346-53

16 Ward MM. Rheumatology care, patient expectations, and the limits of time. *Arthritis Rheum.* 2004; 51(3):307-8

17 Vermeer M, Kuper HH, van der Bijl AE, et al. The provisional ACR/EULAR definition of remission in RA: a comment on the patient global assessment criterion. *Rheumatology (Oxford).* 2012; 51(6): 1076-80

18 Pyne L, Bykerk VP, Boire G, 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. *J Rheumatol.* 2012; 39(11) :2081-7