FACULDADE DE MEDICINA
UNIVERSIDADE D
COIMBRA

MESTRADO INTEGRADO EM MEDICINA – TRABALHO FINAL

ANA RAQUEL MACHADO TEIXEIRA

PHYSICIAN GLOBAL ASSESSMENT OF DISEASE ACTIVITY IN RHEUMATOID ARTHRITIS: PSYCHOMETRIC PROPERTIES AND DETERMINING FACTORS

ARTIGO CIENTÍFICO ORIGINAL

ÁREA CIENTÍFICA DE REUMATOLOGIA

Trabalho realizado sob a orientação de:

DRA. CÁTIA DUARTE

PROFESSOR DOUTOR JOSÉ ANTÓNIO PEREIRA DA SILVA

ABRIL/2022

Physician Global Assessment of disease activity in rheumatoid arthritis: psychometric properties and determining factors

ARTIGO CIENTÍFICO

ÁREA CIENTÍFICA DE REUMATOLOGIA

Ana Raquel Machado Teixeira[1,2]
Cátia Duarte, MD, PhD MD[3,4,5]
José António Pereira da Silva, MD, PhD[3,4,6]

[1] Faculty of Medicine, University of Coimbra, Portugal.

[3] Department of Rheumatology - Centro Hospitalar e Universitário de Coimbra. Coimbra, Portugal.

[4] Coimbra Institute for Clinical and Biomedical Research - Faculty of Medicine, University of Coimbra, Coimbra, Portugal.

[2] raquelteixeira98@gmail.com, Rua Ferreira e Neto, nº 180, 4585-814, Astromil.

[5] catiacmduarte@gmail.com, Praceta Mota Pinto, 3004-561, Coimbra.

[6] jdasilva@ci.uc.pt, Praceta Mota Pinto, 3004-561, Coimbra.

ABSTRACT

Background

Physician Global Assessment (PhGA) is widely used for assessing disease activity in Rheumatoid Arthritis, is important in the assessment of disease activity, being a component of different indexes of disease activity, integrating in crucial therapeutic decisions. Several factors may influence PhGA, affecting its validity, responsiveness and reliability.

Objectives

To assess the validity, responsiveness, and reliability of PhGA and identify its determinants.

Material and Methods

Data from RAID.PT, an observational, prospective and multicenter study, including adult patients fulfilling RA classification criteria, were used. Socio-demographic and clinical data collected in two consecutive visits, separated by 2 to 6 months, included disease duration, comorbidities, medication, tender and swollen joint counts, C-Reactive Protein (CRP), Erythrocyte Sedimentation Rate (ESR), Disease activity Score (DAS28-4v-CRP, DAS28-4v-ESR); PhGA, Patient Global Assessment, Pain, Health Assessment Questionnaire (HAQ), RAID.7 and total RAID score. Construct Validity was assessed through Spearman's correlation between PhGA and DAS. Responsiveness was assessed through the Standardized Response Mean (SRM) in EULAR responder's subgroup and reliability through the Intra-Class Correlation Coefficient (ICC) in the subgroup of patients with stable disease (ΔDAS28-4v-ESR<|0.6|). Variables with $p<0.01$ in univariate analysis were included in 4 different multivariable linear regression models, to define the best model to explain PGA.

Results

In total, 299 patients (81.3% women, mean age of 57.4±12.0 years and disease duration 9.4±9.5 years) with a mean DAS28-4v-ESR of 2.9±1.4 and moderate impact of the disease (RAID score: 3,9±2.2), were included. PhGA was strongly associated with DAS28-4v-CRP (r=0.74) and DAS28-4V-ESR (r=0.70), but moderately with individual items of DAS (r= 0.55 to 0.60). PhGA showed a moderate correlation with the RAID score (r=0.50), being weakly correlated with all domains of RAID, except for the pain domain (r=0.51).

In EULAR responder´s subgroup, PhGA was poorly responsive (SRM:0.43, 95%CI:0.37-0.45). Among patients with a stable condition, PhGA was moderately reliable with an ICC of 0.7 (95% CI:0.6-0.8).

In multivariate analysis, VAS pain (β=0.28, 95%CI: 0.17-0.36), ESR (β=0.18, 95%CI 0.09-0.27), SJC44 (β=2.55, 95%CI 1.79-3.31) and TJC44 (β=0.92, 95%CI 0.36-1.47) remained as independent positive correlates of PhGA (r^2 =0.61 $p<0.05$).

Conclusions

PhGA is a valid and reliable tool, thought poorly sensitive to change in this cohort. Although the responsiveness of PhGA is weak, it can discriminate patients who changed their disease activity from non-responders. Objective indicators of inflammation were the most relevant factors associated with PhGA, with smaller effect of subjective domains. Further research is needed to better understand PhGA due to its importance to clinical practice.

Keywords: rheumatoid arthritis; physician global assessment; responsiveness; reliability; determinants.

RESUMO

Introdução

A avaliação global da doença na perspetiva do médico (PhGA) é amplamente utilizada na avaliação da Artrite Reumatoide (AR), integrando vários índices de atividade da doença, sendo crucial na decisão terapêutica. Vários fatores podem influenciar o PhGA, afetando sua validade, responsividade e reprodutibilidade.

Objetivos

Avaliar a validade, responsividade e reprodutibilidade do PhGA e identificar os preditores.

Materiais e Métodos

Foram utilizados dados do RAID.PT, um estudo observacional, prospetivo e multicêntrico, incluindo pacientes adultos que preenchem os critérios de classificação de AR. Dados sociodemográficos e clínicos coletados em duas visitas consecutivas, separadas por 2 a 6 meses, incluíram duração da doença, comorbilidades, medicação, contagem de articulações tumefactas e dolorosas, proteína C reativa (CRP), velocidade de sedimentação (ESR), escala de atividade da doença (DAS28-4v-CRP, DAS28-4v-ESR); PhGA, avaliação global da doença pelo doente, dor, questionário de avaliação de saúde (HAQ), RAID.7 e pontuação total do RAID. A validade de construto foi avaliada pelo coeficiente de correlação de Spearman entre PhGA e DAS. A responsividade foi avaliada pela *Standardized Response Mean* (SRM) no subgrupo de respondedores EULAR e a reprodutibilidade pelo Coeficiente de Correlação Intraclasse (ICC) no subgrupo com doença estável (ΔDAS28-4v-ESR $<|0,6|$). Variáveis com $p<0,01$ na análise univariada foram incluídas em 4 diferentes modelos de regressão linear multivariada, para definir o melhor modelo explicativo do PhGA.

Resultados

No total, 299 pacientes (81,3% mulheres, idade média de 57,4±12,0 anos e duração da doença 9,4±9,5 anos) com DAS28-4v-ESR médio de 2,9±1,4 e impacto moderado da doença (pontuação do RAID:3,9 ±2,2), foram incluídos. PhGA foi fortemente associado com DAS28-4v-CRP (r=0,74) e DAS28-4V-ESR (r=0,70), mas moderadamente com as componentes individuais do DAS (r=0,55 a 0,60). O PhGA apresentou correlação moderada com o resultado do RAID (r=0,50), sendo fracamente correlacionado com todos os domínios do RAID, exceto o domínio da dor (r=0,51).

No subgrupo de respondedores EULAR, PhGA foi pouco responsivo (SRM:0,43, IC95%: 0,37-0,45). Nos pacientes com doença estável, o PhGA foi moderadamente reprodutível com um ICC de 0,7 (IC95%:0,6-0,8).

Na análise multivariada, dor (β=0,28, IC95%:0,17-0,36), ESR (β=0,18, IC95%:0,09-0,27), SJC44 (β=2,55, IC95%:1,79-3,31) e TJC44 (β=0,92, IC95%:0,36-1,47), explicam 60.8% dos valores do PhGA.

Conclusões

O PhGA é uma ferramenta válida, embora pouco sensível a mudanças nesta coorte. Embora a responsividade tenha sido fraca, o PhGA consegue discriminar os doentes que alteraram a atividade da doença dos não-respondedores. Os indicadores objetivos de inflamação foram os fatores mais associados ao PhGA, com menor efeito dos domínios subjetivos. Mais investigação é necessária dado a sua relevância clínica do PhGA.

Palavras-chave: artrite reumatoide; avaliação global pelo médico; capacidade de resposta; confiabilidade; preditores.

INTRODUCTION

Rheumatoid Arthritis (RA) is a systemic, chronic, and inflammatory autoimmune disease of unknown etiology characterized by synovial inflammation, leading to pain, swelling, morning stiffness, destruction of synovial joints and relevant impact upon patients' lives [1].

The prognosis of RA patients changed drastically over the last years due to the available novel treatments and the early diagnosis, but also due to the treat-to-target [2] strategy, where the immunosuppressive treatment is regularly adjusted as needed to achieve the target of remission as soon and as consistently as possible.

Therefore, the continuous assessment of disease activity is mandatory for patients´ management and treatment decisions. Besides joint counts, inflammatory markers and patient perspective of disease activity, the physician global assessment of disease activity (PhGA) plays an important role in the current assessment of disease activity [3], being a component of different indexes of disease activity as CDAI or SDAI [4].

The relevance of PhGA is reinforced by its impact, when individually considered, in outcomes and physician's decisions. Recent research showed that PhGA is a strong predictor of outcome at long-term in RA, with lower rates at 3 months being associated with sustained remission at one year of follow-up [5]. In other hand, some studies demonstrated that PhGA is an important factor of physicians´ decision´s to change treatment, even more important than the Disease Activity Score (DAS). Therefore, ensuring good psychometric properties of PhGA and understanding the factors underlying its scoring is crucial to promote good quality of care for RA patients.

Previous research found that PhGA is influenced by several factors associated with the patient, both related and unrelated with RA, but also by aspects of physician's profile. PhGA has been found to be more associated with objective measures of inflammation as swollen joint counts and inflammatory markers, and to be only weakly related to the impact of the disease upon patients' lives and comorbidities [6-11]. These influences upon PhGA may vary and affect the reliability and responsiveness of PhGA both in research and clinical practice. Despite its wide use and relevance in RA assessment, the information about reliability and responsiveness of PhGA is scarce. One study evaluated its reliability, during and 1 to 2 days after physician visit, reporting that in such conditions, PhGA is highly reliable [11]. However, the stability PhGA over follow-up in similar clinical circumstances has not been evaluated, to our knowledge.

Thus, the assessment of validity, reliability, and sensitivity to change of PhGA in clinical practice and research is very relevant. Additionally, identifying the factors that explain the variance of PhGA, and their respective weight is relevant to harmonize and standardize the physician global assessment and optimize its use in clinical practice.

With this study we intend to evaluate the psychometric properties of the PhGA in a cohort of Portuguese RA patients and its determinants.

MATERIAL AND METHODS

Study Design and Patients

This is a supplementary analysis of RAID.PT, an observational, prospective and multicentric study of patients with established RA followed in 9 Portuguese centers of rheumatology. Adult patients (≥ 18 years old), with Rheumatoid Arthritis, according to the ACR 1987 [12] or ACR/EULAR 2010 [13] classification criteria, with ability to understand and fill the questionnaires unaided were included. In order to reflect daily clinical practice, patients were evaluated in two consecutive visits, 2-6 months apart.

Ethics

RAID.PT Study was conducted according to the Declaration of Helsinquia [14], and was approved by the Ethics Committee of the Centro Hospitalar e Universitário de Coimbra (CHUC-160-17). All patients provided written informed consent.

Outcome of Interest

PhGA of the current disease activity was assessed on a 0–100 mm visual analogue scale (VAS), with anchors of 0 (not active at all) and 100 (extremely active).

Data Collection

The following socio-demographic and clinical variables were collected: gender, age, disease duration, comorbidities, and current medication for RA (prednisolone (PDN), classic Disease Modifying Anti-Rheumatic Drugs (cDMARDs), biologic DMARDs (bDMARDs)).

Swollen and tender joints in 28 and 44 joint counts performed by the rheumatologist, Erythrocyte Sedimentation Rate (ESR, mm/h) and C-Reactive Protein (CRP, mg/dl) were recorded.

PGA ("Considering all the way that your illness and health conditions affect you at this time, how do you feel?"), and Pain were scored by patients through a VAS 0-100 mm (where 0 corresponds to the best state and 100 to the worst). PGA and PhGA were converted to a 0 to 10 scale to calculate composite disease activity scores.

Disease activity was assessed through the Disease Activity Score 28-ESR (DAS28-ESR), with 4 and 3 variables and classified as *Remission (<2.6), low (2.6 ≤ LDA ≤3.2), moderate (3.2<MDA ≤5.1) and high (HDA >5.1) disease activity* [15]. Patients were classified as

responders according to DAS-based EULAR response criteria [16]. Patients with change of DAS 28-ESR < |0.6| between the two consecutive visits were considered stable.

Patient-perceived impact of RA was evaluated through the RAID score [17,18] at every visit: patients were asked to score the intensity of the impact of their disease in each of seven domains (pain, fatigue, physical function, sleep disturbance, emotional and physical well-being, and coping) on a 0 to 10 numerical rating scales and a combined score was derived according to the proposed algorithm. Each domain was also addressed individually (RAID.7) [19].

The Health Assessment Questionnaire (HAQ) is a self-questionnaire which assesses physical function of patients regarding activities of daily living [20]. The score varies between 0 to 3: 0< HAQ <1 means low disability, 1< HAQ < 2 moderate disability, and 2< HAQ <3high disability

Statistics Analysis

Statistical analysis was performed through Software Package for Social Sciences (SPSS) IBM for Windows® v24 and Medcalc® software version 18.11.3. Descriptive statistics are presented as mean ± standard deviation (SD) for continuous variables and as proportions (%) for categorical variables. Only patients with complete data measures were included in our analyses.

The conceptual framework for outcome measurement (OMERACT 2.0) [21] was used in this methodological study, to assess the psychometric properties of the PhGA.

Validity

Construct validity of PhGA was assessed through its Spearman´s correlation coefficient with DAS28, its individual components (PGA, SJC28, TJC28 and ESR/CRP), HAQ and RAID score. Correlations from 0.3 to 0.5 were considered weak, moderate from 0.5 to 0.7, and strong if >0.7 [22].

Discriminant validity, meaning the capacity to distinguish patients with different conditions, was evaluated comparing the mean changes in PhGA across different categories of disease activity through one-way analysis of variance (ANOVA), with post-hoc Bonferroni correction for pairwise comparisons.

Responsiveness

Longitudinal construct validity, i.e., responsiveness, was assessed by the Pearson's coefficient of correlation between change of PhGA (∆PhGA) and the change of disease activity (∆DAS28-ESR). A comparison between responders and non-responders was performed through independent samples t-test, with the same purpose.

Responsiveness of PhGA was also assessed by the standardized response mean (SRM) (mean/ SD of the change). SRMs were categorized as large (>0.80), moderate (0.5–0.80) and small (0.2 to 0.5) [23]. Furthermore, we compared the change of PhGA in responders and non-responders according to the EULAR response criteria. SRM were calculated and compared for both groups, with an *a priori* hypothesis that it would be higher in the responder group.

Reproducibility

Reliability of PHGA over time was evaluated between the 2 visits in the subgroup of patients with stable disease activity, defined as -0.6<ΔDAS28ESR<0.6. Reliability was assessed through the Intra-Class Correlation Coefficient (ICC, two-way mixed model absolute agreement) with 95% CI. An ICC >0.8 was considered indicative of excellent reliability [24].

Determinants of PhGA

To identify determinants of PhGA, we first used Spearman's correlational analyses to identify variables significantly associated with PhGA. In the next step, these variables were tested in multivariate analyses, by linear stepwise multivariate regression modeling to determine their independent contribution to the PhGA, having excluded collinearity. Different models were tested to define the best model to explain the variance of PhGA.

RESULTS

Patients Characteristics

In total, 299 patients were included (81.3% women with a mean age of 57.4±12.0 years and a mean disease duration of 9.4±9.5 years), with 43.5% (n=130) of the patients being in remission at baseline. Global impact of disease, as reflected by RAID score, was moderate 3,9±2.2, with most of its individual items (RAID7i) scoring higher than 4, except for Sleep Disturbance 3.3±3.0, Emotional Wellbeing 3.7±2.7 and Coping 3.4±2.7. Socio-demographic and clinical characteristics are detailed in Table1.

Table 1: Socio-demographic and clinical characteristics of RA patients.

Characteristics	1st	2nd
Female, n (%)	243 (81.3)	--
Age, years	57.4±12.0	--
Disease duration, years	9.4±9.5	--
Comorbidities sum	0.5±0.7	
0 n (%)	186(62.2)	--
1 n (%)	88(29.4)	--
>2 n (%)	25(8.3)	--
PDN, current intake n (%)	195 (65.2)	--
cDMARDs, current intake n (%)	274 (91.6)	--

bDMARDs, current intake n (%)	109 (36.5)	--
tsDMARDs, current intake n (%)	3 (1.0)	--
PGA	39.9±27.3	38.3±25.0
VAS Pain	40.0±28.2	39.0±26.2
PhGA	21.0±20.5	16.9±17.5
ESR (mm/h)	20.3±19.1	18.1±17.1
CRP (mg/dL)	6.8±9.7	6.6±22.4
DAS 28-3v-ESR	2.8±1.3	2.5±1.3
DAS 28-4v-ESR	2.9±1.4	2.7±1.4
DAS 28-3v-CRP	2.4±1.1	2.2±1.1
DAS 28 4v-CRP	2.6±1.2	2.4±1.2
TJC28	1,8±3.3	1.4±2.7
SJC28	1.4±2.5	1.2±2.5
TJC44	2.1±3.8	1.7±3.2
SJC44	1.5±2.7	1.2±3.2
Categories of Disease Activity (DAS28-4v-ESR)		
Remission n (%)	130 (43.5)	144 (48.2)
Low disease activity n (%)	63 (21.1)	65 (21.7)
Moderate disease activity n (%)	78 (26.1)	74 (24.7)
High disease activity n (%)	28 (9.4)	16 (5.4)
RAID score	3.9±2.3	3.9±2.2
Pain (0-10)	4.4±2.7	4.3±2.6
Function (0-10)	4.2±2.8	4.2±2.7
Fatigue (0-10)	4.2±2.8	4.3±2.7
Sleep Disturbance (0-10)	3.3±3.0	3.4±2.9
Physical Wellbeing (0-10)	4.1±2.6	4.1±2.5
Emotional Wellbeing (0-10)	3.7±2.7	3.5±2.5
Coping (0-10)	3.4±2.7	3.3±2.5
HAQ	0.8±0.7	0.8±0.7

All values are presented as mean ± (standard deviation), unless marked as otherwise. PDN, prednisolone; cDMARDs, classic Disease Modifying anti-Rheumatic Drugs; bDMARDs, biologic Disease Modifying anti-Rheumatic Drugs; tsDMARDs, target Disease Modifying anti-Rheumatic Drugs; RAID, Rheumatoid Arthritis Impact of Disease; PGA, Patient Global Assessment of disease activity; VAS Pain, visual analogic scale for pain; PhGA, Physician Global Assessment of disease activity; ESR, erythrocyte Sedimentation Rate; CRP, C- Reactive Protein; DAS28-ESR-3v, Disease Activity Score-Erythrocyte Sedimentation Rate-3 variables; DAS28-ESR-4v, Disease Activity Score-Erythrocyte Sedimentation Rate-4 variables; DAS28-CRP-3v, Disease Activity Score-C-Reactive Protein-3 variables; DAS28-CRP-4v, Disease Activity Score-C-Reactive Protein-4 variables; TJC 28, tender joints in 28 joint count; SJC28, swollen joints in 28 joint count; TJC44, tender joints in 44 joint count; SJC44, swollen joints in 44 joint count; HAQ, Health Assessment Questionnaire.

Validity of PhGA

Correlations between PhGA and other variables is shown in Table 2. PhGA was strongly associated with DAS28-4V-ESR (r=0.70) and DAS28-4v-CRP (r=0.74). However, showed moderate correlation of individual items of DAS (r from 0.55 to 0.60). PhGA was weakly correlated with all domains of impact of disease activity (RAID).

Table 2: Correlation between PhGA with disease activity variables.

			95% CI	
	r	p	Lower	Upper
PGA	0.58	<0.01	0.48	0.66
VAS pain	0.61	<0.01	0.50	0.67
ESR	0.37	<0.01	0.25	0.48
CRP	0.31	<0.01	0.21	0.41
TJC28	0.55	<0.01	0.45	0.65
SJC28	0.60	<0.01	0.51	0.67
DAS28-3v-ESR	0.62	<0.01	0.52	0.71
DAS28-3v-CRP	0.67	<0.01	0.58	0.74
DAS28-4v-ESR	0.70	<0.01	0.62	0.77
DAS28-4v-CRP	0.74	<0.01	0.66	0.80
RAID score	0.49	<0.01	0.40	0.57
RAID pain	0.51	<0.01	0.42	0.59
RAID function	0.45	<0.01	0.36	0.54
RAID fatigue	0.37	<0.01	0.27	0.47
RAID sleep disturbance	0.27	<0.01	0.15	0.38
RAID physical wellbeing	0.44	<0.01	0.33	0.54
RAID emotional wellbeing	0.40	<0.01	0.30	0.50
RAID coping	0.39	<0.01	0.27	0.49
HAQ	0.29	<0.01	0.18	0.39

Data represents Pearson's correlation coefficient for Physician Global Assessment of disease activity (PhGA) in 299 patients at the first visit (baseline). PGA, Patient Global Assessment of disease activity; VAS Pain, visual analogic scale for pain; ESR, erythrocyte Sedimentation Rate; CRP, C- Reactive Protein; TJC28, tender joint in 28 joint count; SJC28, swollen joint in 28 joint count; RAID, Rheumatoid Arthritis Impact of Disease; DAS28-ESR-3v, Disease Activity Score-Erythrocyte Sedimentation Rate-3 variables; DAS28-ESR-4v, Disease Activity Score-Erythrocyte Sedimentation Rate-4 variables; DAS28-CRP-3v, Disease Activity Score-C-Reactive Protein-3 variables; DAS28-CRP-4v, Disease Activity Score-C-Reactive Protein-4 variables; HAQ, Health Assessment Questionnaire. 95% CI were obtained through bootstrap.

PhGA differed significantly across the disease activities categories, increasing progressively from patients in remission [8.8±1.0 (95% CI 6.8 to 10.8)] to patients with high disease activity [55.8±3.2 (95% CI 49.2 to 62.3)].

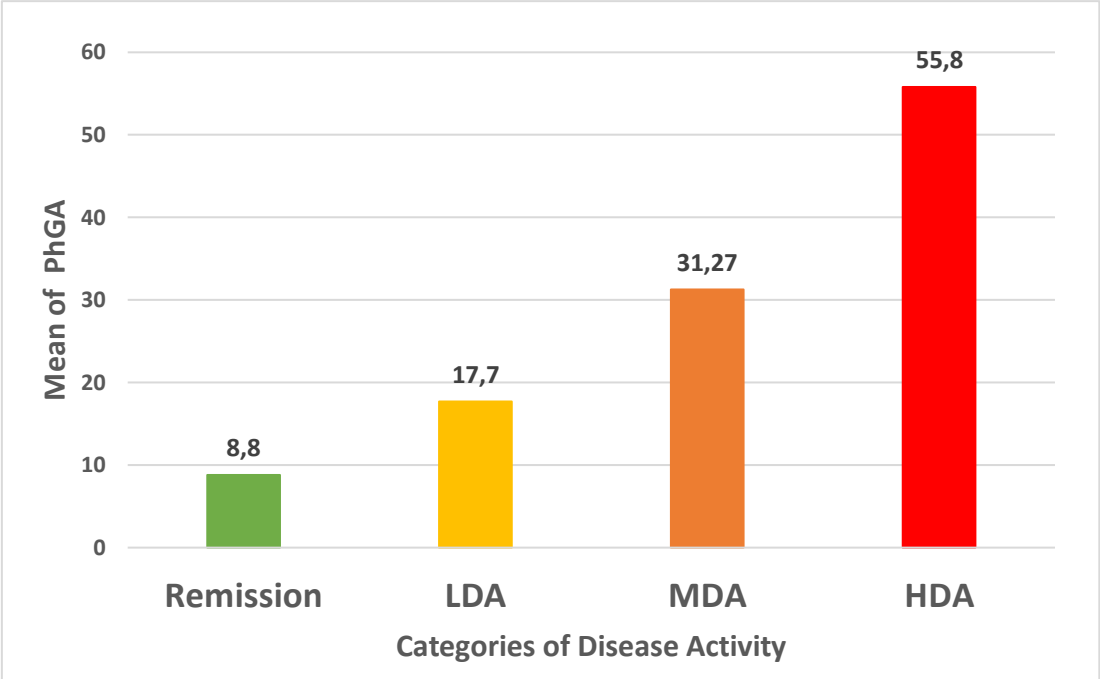

Figure 1: Mean changes in PhGA in patients according to their disease activity at the index visit.
LDA, low disease activity; MDA, moderate disease activity; HAD, high disease activity.

Responsiveness of PhGA

PhGA change from baseline was significantly different between the responder and non-responders (-20.39±2.93 vs 6.58±1.42, p<0,05). The correlation between ΔPhGA and ΔDAS28-4v-ESR was moderate (r=0.50).

Overall responsiveness of PhGA according to SRM was small (SRM: 0.20, 95%CI: 0.09 to 0.27). As we hypothesized previously, PhGA was more responsive among EULAR responders (SRM: 0.39, 95%CI:0.30 to 0.43). Similar results were found for PGA (SRM: 0.41, 95%CI:0.36 to 0.44) in this cohort.

Reliability of PhGA

Among patients with a stable condition (-0.6<ΔDAS283vESR<0.6 (n=145)), change of PhGA was minimal between the two visits (2.78±1.40). PhGA was moderately reliable in this sample, with an ICC of 0.7 (95% CI: 0.6-0.8). Similar results were found for PGA (ICC: 0.7, 95% CI: 0.6-0.8).

Determinants of PhGA

Among the variables tested in the univariate analysis, PhGA was significantly correlated with: ESR, CRP, SJC44, TJC44, SJC28, TJC28, DAS28-3v-ESR, DAS28-4v-ESR, DAS28-3v-CRP, DAS28-4v-CR, RAID score and each of its individual domains (pain, function, fatigue, sleep

disturbance, physical wellbeing, emotional wellbeing, and coping), PGA, VAS pain, and HAQ, with p<0.01 for all of them (Table 2).

In the multivariate analysis, four significant models were obtained (table 3). The most significant model includes VAS pain (β=0.278, 95% CI: 0.196-0.355), ESR (β=0.181, 95% CI 0.091-0.270), SJC44 (β=2.550, 95% CI 1.788-3.313) and TJC44 (β=0.917, 95% CI 0.360-1.473), explaining 61% of the variance of PhGA.

Table 3: Multivariate linear regression analysis regarding the four models.

Model 1	**Model 2**	**Model 3**	**Model 4**
VAS pain (β=0.28, 95% CI: 0.20-0.36) ESR (β=0.18, 95% CI 0.09-0.27) SJC44 (β=2.55, 95% CI 1.79-3.31) TJC44 (β=0.92, 95% CI 0.36-1.47)	PGA (β=0.27, 95% CI: 0.21-0.33) ESR (β=0.19, 95% CI: 0.10-0.28) SJC44 (β=2.59, 95% CI: 1.18-3.36) TJC44 (β=0.95, 95% CI: 0.38-1.51)	RAID score (β=2.45, 95% CI: 1.73-3.18) ESR (β=0.18, 95% CI: 0.09-0.28) SJC44 (β=2.60, 95% CI: 1.79-3.41) TJC44 (β=1.09, 95% CI: 0.50-1.68)	HAQ (β=2.73, 95% CI: 0.19-5.26) ESR (β=0.20, 95% CI: 0.10-0.31) SJC44 (β=2.58, 95% CI: 1.71-3.44) TJC44 (β=1.50, 95% CI: 0.89-2.12)
r^2: 0.61 p <0.01	r^2: 0.60 p <0.01	r^2: 0.56 p <0.01	r^2: 0.50 p <0.01

VAS Pain, visual analogic scale for pain; ESR, Erythrocyte Sedimentation Rate; TJC44, tender in 44 joint count; SJC44, swollen in 44 joint count; PGA, Patient Global Assessment of disease activity; RAID, Rheumatoid Arthritis Impact of Disease; HAQ, Health Assessment Questionnaire.

DISCUSSION AND CONCLUSION

This study brings to light the psychometric properties and drivers of PhGA, which is quite relevant given its determinant role in clinical decisions.

PhGA showed good construct validity with a strong correlation with composite indices of disease activity routinely used in clinical practice, as DAS28, and moderate correlation with its individual components. At the present time there are no studies that focus strictly on the validation of the PhGA, but our results are consistent with some studies which aimed to define drivers correlating with PhGA [6-11], with exception of one which showed a low correlation between PhGA and DAS28 [25]. PhGA showed also good capacity to discriminate groups of interest, with significant differences across the different disease activity categories. PhGA showed low to moderate correlation with measures of impact, reflecting the current belief that physicians are more focused on the biological process than on the impact of the disease in the different domains of patients' lives. Such findings can justify the low agreement between

patients and physicians' assessment of disease activity reported in the literature [6-8], leading to problems in patient-physician communication and jeopardize the recommended patient centered care and the holistic approach of the patient with RA.

In our cohort, PhGA showed ability to discriminate patients who changed their disease activity from non-responders. However, responsiveness of PhGA in our cohort was weak, as reflected by a low SRM, and by the weak correlation between change in PhGA and change in disease activity score. The low disease activity at baseline, with 43% of patients being in remission, could influence such results, with low margin for clinical response over time. Despite the high remission rate, only 23.4% were scored as zero on the PhGA, with most of them scoring around 10 mm, leading to a narrow range for change overtime.

PhGA was moderately reliable in this cohort, as reflected by an ICC of 0.7, in contrast with a previous report that found PhGA strongly reliable (ICC: 0.96, 95% CI:0.95-0.97) [11]. Such discrepancy could be justified by the different approaches in both studies. In Rohekar´s study, reliability was assessed with only some hours of interval between the two assessments. Although this short interim period guaranteed that the clinical condition remained stable, it could be too short to avoid memory bias. In our study, the evaluations were performed 2 to 6 months apart, reflecting the typical scenario in clinical practice. However, we cannot exclude the occurrence of some events between the two visits that might have influenced physician assessment, as treatment changes in the first visit, flares reported by patients, adverse events or other clinical conditions, which were not assessed in our study [11].

Objective markers of inflammation were the main drivers of PhGA, in agreement with other studies [6-10]. Number of Swollen joints, including others besides the 28 joint count, remains the most determinant factor in all models with a lesser effect of tender joints. The patients' perspective has some, albeit small, impact on the evaluation of PhGA. Pain is the patient reported outcome (PROs) most related with PhGA, which may be explained by the expected straighter relationship between pain with disease activity than other PROs. PGA is more subjective and influenced by other factors usually not considered so relevant by the physicians, as fatigue, psychological aspects, life events or socio-cultural aspects [7-8, 26]. In our study, no association between PhGA and comorbidities, depression, and corticosteroid use were established, in disagreement with a recent study, that found as predictors of PhGA the positive rheumatoid factor, HAQ-DI and use of corticosteroids [10]. However, a significant proportion of variance of PhGA remains unexplained. Some factors not collected in our study, particularly related with the physician himself could be considered. A previous study showed that age, gender, clinical setting, and years of experience of the physician explained some of the variance of PhGA [3].

These findings are extremely relevant in clinical practice. To our knowledge there is no standardization of how to score PhGA in RA and no training is given to physician, which occurs

in other diseases such as Systemic Lupus Erythematosus (SLE). Given the relevance of PhGA in treatment decisions, some standardization with its use should be considered.
Some considerations must be taken into account when interpreting our results. Despite being a multicentric study, we only included Portuguese patients, which limits the generalizability of the results. The evaluations were conducted 2 to 6 months apart to replicate the most likely scenario in clinical practice, but the results could have been significantly different if the time between consecutive evaluations was shorter. The study was performed in a population with an average DAS in the low-disease-activity range. Results might have been different in a population with higher disease activity and with higher change of disease activity between the two visits.

In conclusion, PhGA is a valid and reliable construct, when clinical conditions remain stable. Although the responsiveness of PhGA is weak, PhGA is able to discriminate patients who changed their disease activity from non-responders. Drivers of PhGA are objective markers of inflammation, with low influence of impact of the disease upon patients´ lives. However, a large proportion of variance remains unexplained and further research is needed to better understand PhGA variation.

ACKNOWLEDGMENTS

My sincere gratitude to Dr. Cátia Duarte, my rheumatology teacher and mentor, for all the learning opportunities, guidance, availability, understanding, support, and patience throughout these months of work. To Professor José António Pereira da Silva, I would like to thank for his availability and passionate influence regarding rheumatology teaching. To Dr. Liliana Saraiva, for letting me be part of her work, for the patience and support during these months of work. My most sincere thanks for having transmitted to me the enthusiasm for Rheumatology.

To all the patients who agreed to participate in the RAID.PT study, and to all the rheumatologists who collected the data and generously allowed me to use the database.

To Professor Carla Henriques of the Instituto Politécnico de Viseu, for her time and guidance in the statistical analysis.

To my parents, sister, brother, and grandparents for being an example of resilience, for their patience and support in pursuing my dreams. Specially to my mother, for helping me get through not only the statistical dramas of this work, but all the others, over the last 6 years.

To Samuel, my everything, for showing me, countless times, what I am capable of and for all the love, support, and affection over the years.

To my godmother Maria Parreira, for being my companion on this journey and never letting me give up.

To my dear friends, for all the support throughout these last 6 years and for putting up with my daydreams about this work. Particularly to my friend Beatriz Vieira, for all her patience and trust in me, assuring me that I would be able to do it.

REFERENCES

1. Scott DL, Pugner K, Kaarela K, Doyle DV, Woolf A, Holmes J, et al. The links between joint damage and disability in rheumatoid arthritis. Rheumatology (Oxford). 2000;39(2):122-32.
2. Smolen JS, Breedveld FC, Burmester GR, Bykerk V, Dougados M, Emery P, et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. Ann Rheum Dis. 2016;75(1):3-15.
3. Turk M, Pope JE. Physician global assessments for disease activity in rheumatoid arthritis are all over the map! RMD Open. 2018;4(1):e000578.
4. Aletaha D, Smolen J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. Clin Exp Rheumatol. 2005;23(5 Suppl 39):S100-8.
5. Choy T, Bykerk VP, Boire G, Haraoui BP, Hitchon C, Thorne C, et al. Physician global assessment at 3 months is strongly predictive of remission at 12 months in early rheumatoid arthritis: results from the CATCH cohort. Rheumatology (Oxford). 2014;53(3):482-90.
6. Desthieux C, Hermet A, Granger B, Fautrel B, Gossec L. Patient-Physician Discordance in Global Assessment in Rheumatoid Arthritis: A Systematic Literature Review With Meta-Analysis. Arthritis Care Res (Hoboken). 2016;68(12):1767-73.
7. Brites L, Dinis de Freitas J, Costa F, Luís M, Prata AR, Assunção H, et al. Patient-Physician discordance in assessment of disease activity in Rheumatoid Arthritis patients. Acta Reumatol Port. 2021;46(2):103-9.
8. Studenic P, Radner H, Smolen JS, Aletaha D. Discrepancies between patients and physicians in their perceptions of rheumatoid arthritis disease activity. Arthritis Rheum. 2012;64(9):2814-23.
9. Kaneko Y, Takeuchi T, Cai Z, Sato M, Awakura K, Gaich C, et al. Determinants of Patient's Global Assessment of Disease Activity and Physician's Global Assessment of Disease Activity in patients with rheumatoid arthritis: A post hoc analysis of overall and Japanese results from phase 3 clinical trials. Mod Rheumatol. 2018;28(6):960-7.
10. Guimarães M, Pinto M, Resende GG, Machado CJ, Vargas-Santos AB, Amorim RBC, et al. Discordance between the patient's and physician's global assessment in rheumatoid arthritis: Data from the REAL study-Brazil. PLoS One. 2020;15(3):e0230317.
11. Rohekar G, Pope J. Test-retest reliability of patient global assessment and physician global assessment in rheumatoid arthritis. J Rheumatol. 2009;36(10):2178-82.
12. Arnett, F.C., et al., The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum, 1988. 31(3): p. 315-24.
13. Aletaha, D., et al., 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis, 2010. 69(9): p. 1580-8.
14. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. Jama. 2013;310(20):2191-4.

15. DAS-Score Website 2018 [Available from: https://www.das-score.nl/das28/en/difference-between-the-das-and-das28/how-to-measure-the-das28/how-to-calculate-the-das28/alternative-validated-formulae.html].
16. van Riel PL. The development of the disease activity score (DAS) and the disease activity score using 28 joint counts (DAS28). Clin Exp Rheumatol. 2014;32(5 Suppl 85):S-65-74.
17. Gossec L, Dougados M, Rincheval N, Balanescu A, Boumpas DT, Canadelo S, et al. Elaboration of the preliminary Rheumatoid Arthritis Impact of Disease (RAID) score: a EULAR initiative. Ann Rheum Dis. 2009;68(11):1680-5.
18. Gossec L, Paternotte S, Aanerud GJ, Balanescu A, Boumpas DT, Carmona L, et al. Finalisation and validation of the rheumatoid arthritis impact of disease score, a patient-derived composite measure of impact of rheumatoid arthritis: a EULAR initiative. Ann Rheum Dis. 2011;70(6):935-42.
19. Duarte C, Santos EJF, Ferreira RJO, Kvien TK, Dougados M, de Wit M, et al. Validity and reliability of the EULAR instrument RAID.7 as a tool to assess individual domains of impact of disease in rheumatoid arthritis: a cross-sectional study of 671 patients. RMD Open. 2021;7(1).
20. Fries JF, Spitz P, Kraines RG, Holman HR: Measure- ment of patient outcome in arthritis. Arthritis Rheum 23:137-145, 198.
21. Boers M, Kirwan JR, Gossec L, et al. How to choose core outcome measurement sets for clinical trials: OMERACT 11 approves filter 2.0. The Journal of rheumatology. 2014 May; 41(5):1025-1030.
22. Cheung PP, Lahiri M, March L, Gossec L. Patient-reported outcomes in Asia: evaluation of the properties of the Rheumatoid Arthritis Impact of Disease (RAID) score in multiethnic Asian patients with rheumatoid arthritis. Clin Rheumatol. 2017;36(5):1149-54.
23. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol. 2000;53(5):459-68.
24. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb). 2012;22(3):276-82.
25. Harrington JT. The uses of disease activity scoring and the physician global assessment of disease activity for managing rheumatoid arthritis in rheumatology practice. J Rheumatol. 2009;36(5):925-9.
26. Ferreira RJO, Dougados M, Kirwan JR, Duarte C, de Wit M, Soubrier M, et al. Drivers of patient global assessment in patients with rheumatoid arthritis who are close to remission: an analysis of 1588 patients. Rheumatology (Oxford). 2017;56(9):1573-8.