ARTHRITIS & RHEUMATOLOGY
Vol. 66, No. 2, February 2014, pp 428–432
DOI 10.1002/art.38248

BRIEF REPORT

The Value of a Patient Global Assessment of Disease Activity in Granulomatosis With Polyangiitis (Wegener's)

Gunnar Tomasson,[1] John C. Davis,[2] Gary S. Hoffman,[3] W. Joseph McCune,[4] Ulrich Specks,[5] Robert Spiera,[6] E. William St.Clair,[7] John H. Stone,[8] and Peter A. Merkel[9]

Objective. **To 1) describe the distribution of patient global assessment (PtGA) scores of disease activity in patients with granulomatosis with polyangiitis (GPA; Wegener's), 2) explore the discordance between PtGA scores and physician global assessment (PhGA) scores of disease activity, and 3) explore whether PtGA scores during disease remission are associated with future disease relapse.**

Methods. **Data from the Wegener's Granulomatosis Etanercept Trial (WGET) were used. PtGA and PhGA scores were assessed on 100-mm visual analog scales (VAS). Presence of active disease was determined using the Birmingham Vasculitis Activity Score for WG (BVAS/WG), and remission was defined as a BVAS/WG score of 0. Disease relapse was defined as a BVAS/WG score of >0 after remission had been achieved. Discordance between PtGA and PhGA scores was defined as a difference of ≥20 points between the two measures. Mixed linear models were used in longitudinal analysis of PtGA scores.**

Results. **Data were obtained from 180 patients in the WGET cohort, seen at a total of 1,719 study visits. The mean ± SD PtGA and PhGA disease activity scores (on 100-mm VAS) at baseline were 64.2 ± 27.4 and 55.5 ± 23.4, respectively. PtGA–PhGA discordance occurred in 53% of patients at baseline, and this was inversely associated with newly diagnosed disease (as opposed to relapsing disease) at baseline (odds ratio 0.37, 95% confidence interval [95% CI] 0.20–0.68) but not with age, sex, or presence of renal or pulmonary disease. Patients were in disease remission during 62% of the study visits. The mean PtGA score during visits immediately prior to relapse was 4.52 points higher (95% CI 0.66–8.4) than that at other remission visits ($P = 0.03$).**

Conclusion. **PtGA–PhGA discordance is common in GPA. A rise in the PtGA disease activity score during times defined by physicians as periods of remission is associated with subsequent occurrence of disease relapse. These findings support the addition of PtGA as an outcome measure for GPA.**

Granulomatosis with polyangiitis (GPA; Wegener's) is a multisystem disease that often has a chronic course with frequent periods of relapse and remission. Disease remission is attained for the great majority of patients, and longitudinal studies demonstrate that patients with GPA are in apparent remission at most study visits (1–4). The Birmingham Vasculitis Activity Score (BVAS) (5) and its modifications are the most widely accepted tools for disease activity assessment in antineutrophil cytoplasmic antibody–associated vasculitis, with remission usually defined as a BVAS score of 0. However, the BVAS is a purely physician-based disease measure that relies on the physician's judgment regarding involvement of individual organ systems and attribution of manifestations to disease activity but not to

The Wegener's Granulomatosis Etanercept Trial was supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases [NIAMS] grant N01-AR-92240 and National Center for Research Resources, General Clinical Research Center grants M01-RRO-00533 to Boston University, M01-RRO-0042 to University of Michigan, M01-RR-30 to Duke University, and M01-RRO-2719 to Johns Hopkins University School of Medicine) and by the FDA, Office of Orphan Products (grant FD-R-001652). Drs. St.Clair, Stone, and Merkel's work was supported by the NIH (NIAMS grants K24-AR-02126-04, K24-AR-049185-01, and K24-AR-2224-01A1).

[1]Gunnar Tomasson, MD: University of Iceland, Reykjavík, Iceland; [2]John C. Davis, MD, MPH: Genentech Inc., South San Francisco, California; [3]Gary S. Hoffman, MD: Cleveland Clinic, Cleveland, Ohio; [4]W. Joseph McCune, MD: University of Michigan, Ann Arbor; [5]Ulrich Specks, MD: Mayo Clinic, Rochester, Minnesota; [6]Robert Spiera, MD: Hospital for Special Surgery, New York, New York; [7]E. William St.Clair, MD: Duke University and Duke University Medical Center, Durham, North Carolina; [8]John H. Stone, MD, MPH: Massachusetts General Hospital, Boston; [9]Peter A. Merkel, MD, MPH: University of Pennsylvania, Philadelphia, and Boston University, Boston, Massachusetts.

Dr. Davis owns stock or stock options in Genentech Inc.

Address correspondence to Gunnar Tomasson, MD, University of Iceland, Stapi Hringbraut, 101 Reykjavík, Iceland. E-mail: gunnar.tomasson@gmail.com.

Submitted for publication May 7, 2013; accepted in revised form October 22, 2013.

permanent damage, adverse effects from treatments, or other nonvasculitis disease processes. It is increasingly recognized that patients can provide insight to disease assessment beyond what can be obtained from physician-based disease assessments or laboratory values (6–8).

The objectives of this study were to 1) describe the distribution of scores for patient global assessment (PtGA) of disease activity in patients with GPA enrolled in a clinical trial, 2) explore the discordance between PtGA scores and physician global assessment (PhGA) scores of disease activity, and 3) explore whether PtGA scores during disease remission are associated with future occurrence of disease relapse.

PATIENTS AND METHODS

Data source. Data from the Wegener's Granulomatosis Etanercept Trial (WGET) (1) (see Appendix A for a list of participating centers and investigators) were used. The WGET was a randomized, double-blind, placebo-controlled trial conducted at 8 clinical centers in North America, testing the addition of etanercept or placebo to standard remission-induction therapy (glucocorticoids plus cyclophosphamide or methotrexate) for patients with GPA. Patients were enrolled at a time of active vasculitis and had study visits at baseline, 6 weeks, and every 3 months thereafter.

Study variables. PtGA scores of disease activity were assessed at every study visit, on a 100-mm visual analog scale (VAS) (scored 0–100), with the question, "Please mark the line below indicating how active you believe your Wegener's granulomatosis has been in the past 28 days. Consider only how much your Wegener's (the disease itself) is causing you problems. Do not count the effects of other medical problems or side effects of medications." PhGA scores of disease activity were assessed at every study visit, on a 100-mm VAS (scored 0–100), with the question, "Mark line to indicate the amount of WG disease activity (not including longstanding damage) within the previous 28 days."

Disease activity was assessed with the BVAS/WG (9). The BVAS/WG measures activity in 34 items, categorized into 9 groups. BVAS/WG values range from 0 to 63, with higher scores representing more manifestations of active disease. Active disease was defined as a BVAS/WG score of >0, and remission was defined as a BVAS score of 0. Disease relapse was defined as a BVAS/WG score of >0 when the BVAS/WG score had been 0 at the preceding study visit.

Renal disease and pulmonary disease were each defined on the basis of corresponding BVAS/WG items. Health-related quality of life (HRQOL) was assessed with the Short Form 36 (SF-36) health survey at the baseline visit. The SF-36 contains 36 items that assess HRQOL in 8 health dimensions: physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, and mental health (10,11). Scores for each dimension/subscale range from 0 to 100, with higher scores indicating better HRQOL. Two summary scores are derived from the 8 subscales: the physical component summary score and the mental component summary score, both of which are norm-based scores standardized to the values for the US general population and transformed to have a mean of 50 and SD of 10 in the referent population.

PtGA–PhGA discordance was defined as a difference between the PtGA and PhGA scores (on 100-mm VAS) of ≥20 points. We also explored whether our findings were sensitive to discordance thresholds of ≥10 points or ≥30 points.

Statistical analysis. The distribution of outcome measures at baseline was expressed as the mean ± SD. Pearson's correlation coefficients were calculated to determine the correlations between the global assessment scores and the SF-36 subscores. Logistic regression was used to explore the association of PtGA–PhGA discordance with age, sex, newly diagnosed disease or relapsing disease at baseline, and presence of renal or pulmonary disease. The association of the demographic and disease-associated factors with PtGA–PhGA discordance was expressed as the odds ratio (OR) with 95% confidence intervals (95% CIs).

To explore whether the PtGA score was increased prior to overt disease relapse, analysis was limited to data from visits during remission. Mixed linear models were used to explore differences in PtGA scores between study visits during remission but immediately prior to overt relapse, and other study visits during remission; random intercepts were included in the models to account for multiple observations from each subject. The difference in PtGA scores at visits prior to relapse compared to other remission visits was expressed as points on the VAS with 95% CIs. All statistical analyses were done using SAS version 9.2.

RESULTS

Patients. One-hundred eighty patients participated in the WGET and were seen at a total of 1,719 study visits. At baseline, when all patients had active disease (mean ± SD BVAS/WG score 7.0 ± 3.4), the mean ± SD PtGA score was 64.2 ± 27.4, and the mean ± SD PhGA score was 55.5 ± 23.4. At study entry, 100 patients had relapsing disease and 80 patients had a new diagnosis of GPA. Renal involvement was present at baseline in 97 patients (54%), and pulmonary involvement was present at baseline in 108 patients (60%).

Correlations of measures at baseline. PtGA scores were modestly correlated with both PhGA scores ($r = 0.30$, $P < 0.0001$) and BVAS/WG scores ($r = 0.28$, $P < 0.0001$). PhGA scores were highly correlated with BVAS/WG scores ($r = 0.62$, $P < 0.0001$). The PtGA score correlated weakly or moderately with all of the subscores of the SF-36; no correlation or much weaker correlations were observed between the PhGA score and SF-36 subscores (Table 1).

Discordance between patient and physician global assessments at baseline. At baseline, there was substantial discordance regarding the assessment of dis-

Table 1. Correlations of the patient and physician global assessment scores of disease activity with subscores of the SF-36*

	Patient global assessment		Physician global assessment	
SF-36 domain	Pearson's r	*P*	Pearson's r	*P*
Pain	−0.36	<0.0001	−0.09	0.250
Physical function	−0.31	<0.0001	−0.21	0.0039
Role physical	−0.33	<0.0001	−0.10	0.16
Role emotional	−0.23	0.0023	−0.13	0.08
Social function	−0.45	<0.0001	−0.20	0.0064
Vitality	−0.42	<0.0001	−0.13	0.0788
General health	−0.30	<0.0001	0.00	0.98
Mental health	−0.27	0.0003	−0.08	0.26
PCS	−0.38	<0.0001	−0.13	0.08
MCS	−0.30	<0.0001	−0.12	0.11

* Correlations of the 36-item Short Form (SF-36) subscores with patient and physician global assessment scores on 100-mm visual analog scales were determined using Pearson's correlation coefficients. PCS = physical component summary score; MCS = mental component summary score.

ease activity between the patients and the physicians. For 95 patients (53%), there was a difference of at least 20 points between the PtGA and PhGA scores. Having newly diagnosed disease at study entry was inversely associated with PtGA–PhGA discordance (OR 0.37, 95% CI 0.20–0.68). This means that patients with newly diagnosed disease were more likely than those with relapsing disease to agree with the physician's global assessment of disease activity. This inverse association was not dependent on the defined threshold for PtGA–PhGA discordance (Table 2). No demographic or disease-related factor was significantly associated with the PtGA–PhGA discordance.

Patient global assessment during remission and its association with future disease relapse. Of the 180 patients with GPA, 162 (90%) achieved disease remission (BVAS/WG score of 0) on at least one study visit. Among these 162 patients, 74 (46%) experienced at least one disease relapse. The 162 patients who achieved remission had a total of 1,058 visits during the times of remission, of which 103 visits were followed by a disease relapse. At visits immediately prior to relapse, the mean PtGA score was 4.52 points higher (95% CI 0.66–8.4) compared to that at other remission visits ($P = 0.03$) (Table 3). Among the 74 patients who experienced disease relapse after achievement of remission, the PtGA score was 3.0 points higher (95% CI −1.68–7.76) at the visit immediately preceding relapse compared to that at other remission visits among this group ($P = 0.21$).

Table 3. Patient global assessment scores during times of remission and disease relapse*

	Score	*P*†
Remission (n = 885)‡	15.7	Referent
Two visits prior to relapse (n = 70)	17.7	0.41
One visit prior to relapse (n = 103)	20.2	0.03
Relapse visit (n = 103)	28.2	<0.001

* Values are the mean patient global assessment scores (n = number of study visits) during remission, at visits leading up to disease relapse, and at visits after disease relapse.
† Versus the referent group.
‡ Remission that is not followed by a relapse during the next 2 study visits.

DISCUSSION

This study demonstrates that a simple patient-reported outcome, in this case a global score of disease activity measured by a 100-mm VAS, captures novel and important disease aspects in GPA. The PtGA only modestly correlates with other outcome measures, and patient and physician disease assessments were commonly discordant. Correlations between subscores of the SF-36 and the PtGA score were weak or modest, but statistically significant. These findings support several intriguing concepts: 1) the PtGA captures disease domains not assessed by physician-based measures; 2)

Table 2. Associations of demographic and disease factors with discordance between the patient and physician global assessments of disease activity*

Variable	≥10-point discordance (n = 131)	≥20-point discordance (n = 95)	≥30-point discordance (n = 65)
Age (per 1 year)	1.00 (0.98–1.03)	1.00 (0.99–1.02)	1.00 (0.99–1.03)
Male sex	2.37 (1.20–4.67)	1.44 (0.79–2.64)	0.99 (0.53–1.85)
Newly diagnosed disease (at baseline)	0.57 (0.30–1.11)	0.37 (0.20–0.68)	0.32 (0.16–0.61)
Renal disease	1.17 (0.60–2.28)	0.68 (0.38–1.23)	0.61 (0.33–1.13)
Pulmonary disease	0.91 (0.46–1.81)	0.90 (0.49–1.65)	0.81 (0.44–1.51)

* Discordance was defined as a ≥10-, ≥20-, or ≥30-point minimum difference between patient global assessment and physician global assessment scores. Values are the odds ratio (95% confidence interval).

change in a patient-reported outcome (PtGA) during remission precedes periods of active disease as detected by physicians; and 3) in GPA, there may be periods of occult disease activity that are not captured by the current standard of physician-based activity measures (BVAS, PhGA).

The finding that, when compared to patients in the WGET trial with relapsing disease at baseline, patients with newly diagnosed GPA at the baseline visit were more likely to agree with their physician with respect to assessment of disease activity was unexpected. It is possible that disease assessment in cases of relapsing disease presents more complexities than that in newly presented cases, with a need by both physicians and patients to differentiate active disease from the effects of disease- or treatment-related damage.

This study has important strengths. The data source was a well-defined patient cohort with GPA, characterized by high disease activity at baseline, followed by disease remission in a majority of the patients, and then by 1 or more relapses in a substantial number of patients. The assessments were made prospectively per a clearly defined protocol at centers expert in the care of patients with GPA. The dynamics of disease trajectories seen in this cohort and comprehensive, repeated standardized disease assessments make this a rich data source for assessing the utility of outcome measures in GPA.

The study also has some limitations. Direct measures of fatigue, pain, dyspnea, and other important manifestations of disease from the patient perspective (12) were not collected. Although the observed measures of correlation between subscores of the SF-36 and the PtGA score suggest that fatigue (vitality) and social function might be the domains of HRQOL that contribute the most to the PtGA, individual subscores of the SF-36 did not have optimal distribution for such calculations, and therefore these results should be interpreted with caution. Further exploration of the dimensions of disease captured by the PtGA is much needed. Based on findings from the WGET, etanercept is generally considered inefficacious for the treatment of GPA. Therefore, data originating from the WGET would not allow assessment of how outcome measures discriminate between treatment arms. Consistent with the trial's primary conclusion, PtGA scores did not discriminate between treatment arms in this study (results not shown).

The utility of the PtGA should be explored further with respect to its construct validity with other disease manifestations and its ability to discriminate between treatment arms. It is likely that the PtGA could contribute to a composite measure of disease activity in GPA, as it has for other diseases (13,14), and could provide unique information complementary to standard physician assessments. In the absence of tools better than those currently available for assessment of disease activity in GPA, data obtained through the use of this highly feasible patient-reported outcome should be collected in all longitudinal studies of GPA, especially in randomized controlled trials of therapies in GPA.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Tomasson had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Tomasson, Merkel.

Acquisition of data. Davis, Hoffman, McCune, Specks, Spiera, St.Clair, Stone, Merkel.

Analysis and interpretation of data. Tomasson, Merkel.

ADDITIONAL DISCLOSURES

Dr. Davis is an employee of Genentech Inc.

REFERENCES

1. Wegener's Granulomatosis Etanercept Trial Research Group. Etanercept plus standard therapy for Wegener's granulomatosis. N Engl J Med 2005;352:351–61.
2. Jayne D, Rasmussen N, Andrassy K, Bacon P, Tervaert JW, Dadoniene J, et al. A randomized trial of maintenance therapy for vasculitis associated with antineutrophil cytoplasmic autoantibodies. N Engl J Med 2003;349:36–44.
3. Jones RB, Tervaert JW, Hauser T, Luqmani R, Morgan MD, Peh CA, et al. Rituximab versus cyclophosphamide in ANCA-associated renal vasculitis. N Engl J Med 2010;363:211–20.
4. Stone JH, Merkel PA, Spiera R, Seo P, Langford CA, Hoffman GS, et al. Rituximab versus cyclophosphamide for ANCA-associated vasculitis. N Engl J Med 2010;363:221–32.
5. Luqmani RA, Bacon PA, Moots RJ, Janssen BA, Pall A, Emery P, et al. Birmingham Vasculitis Activity Score (BVAS) in systemic necrotizing vasculitis. QJM 1994;87:671–8.
6. Hewlett S, Sanderson T, May J, Alten R, Bingham CO III, Cross M, et al. 'I'm hurting, I want to kill myself': rheumatoid arthritis flare is more than a high joint count—an international patient perspective on flare where medical help is sought. Rheumatology (Oxford) 2012;51:69–76.
7. Kirwan JR, Boonen A, Harrison MJ, Hewlett SE, Wells GA, Singh JA, et al. OMERACT 10 Patient Perspective Virtual Campus: valuing health; measuring outcomes in rheumatoid arthritis fatigue, RA sleep, arthroplasty, and systemic sclerosis; and clinical significance of changes in health. J Rheumatol 2011;38:1728–34.
8. Neogi T, Xie H, Felson DT. Relative responsiveness of physician/assessor-derived and patient-derived core set measures in rheumatoid arthritis trials. J Rheumatol 2008;35:757–62.
9. Stone JH, Hoffman GS, Merkel PA, Min YI, Uhlfelder ML, Hellmann DB, et al, for the International Network for the Study of the Systemic Vasculitides. A disease-specific activity index for

Wegener's granulomatosis: modification of the Birmingham Vasculitis Activity Score. Arthritis Rheum 2001;44:912–20.
10. McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993;31:247–63.
11. Ware JE Jr, Sherbourne CD. The MOS 36-item Short-Form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473–83.
12. Herlyn K, Hellmich B, Seo P, Merkel PA for the Vasculitis Clinical Research Consortium. Patient-reported outcome assessment in vasculitis may provide important data and a unique perspective. Arthritis Care Res (Hoboken) 2010;62:1639–45.
13. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight–joint counts: development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum 1995;38:44–8.
14. Van der Heijde D, Lie E, Kvien TK, Sieper J, Van den Bosch F, Listing J, et al, for the Assessment of SpondyloArthritis international Society (ASAS). ASDAS, a highly discriminatory ASAS-endorsed disease activity score in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68:1811–8.

APPENDIX A: THE WGET RESEARCH GROUP

The WGET Research Group comprises the following investigators and centers: John H. Stone, MD, MPH (Chairman) (Johns Hopkins Vasculitis Center) and Gary S. Hoffman, MD (Co-Chairman) (Cleveland Clinic Foundation Center for Vasculitis Research and Care); for the coordinating center, Janet T. Holbrook, PhD, MPH (Director), Curtis L. Meinert, PhD (Associate Director), John Dodge (Systems Analyst), Jessica Donithan (Research Coordinator), Nancy Min, PhD (Biostatistician), Laurel Murrow, MSc (former Trial Coordinator), Jacki Smith (Research Data Assistant), Andrea K. Tibbs, BS (Trial Coordinator), and Mark Van Natta, MHS (Biostatistician) (Johns Hopkins University Center for Clinical Trials); for the clinical centers, Robert Spiera, MD, Rosanne Berman, MPH, and Sandy Enuha, MPH (Beth Israel Medical Center, New York), Peter A. Merkel, MD, MPH, Rondi Gelbard, BS, Melynn Nuite, RN, and Aileen Schiller, MS (Boston University), David Blumenthal, MD, Debora Bork, MFA, Tiffany Clark, CNP, Sonya L. Crook, RN, Leonard H. Calabrese, DO, Sharon Farkas, Sudhakar Sridharan, MD, Kimberly Strom, CNP, and William Wilke, MD (Cleveland Clinic Foundation), E. William St.Clair, MD, Nancy B. Allen, MD, Karen Rodin, RN, and Edna Scarlett (Duke University), David B. Hellmann, MD, Amanda M. Moore, BS, Lourdes Pinachos, RN, BSN, Michael J. Regan, MD, MRCP, and Misty L. Uhlfelder, MPH (Johns Hopkins University), Ulrich Specks, MD, Kristin Bradt, Kimberly Carlson, Susan Fisher, RN, Boleyn Hammel, Kathy Mieras, and Steven Ytterberg, MD (Mayo Clinic), John C. Davis, MD, MPH, Maureen Fitzpatrick, MPH, Ken Fye, MD, and Steve Lund, MSN, NP (University of California, San Francisco), and Joseph McCune, MD, Billie Jo Coomer, BS, Barbara Gilson, RN, Hilary Haftel, MD, Ana Morrel-Samuels, BA, and Sandra Neckel, RN (University of Michigan); for the resource center, Noel R. Rose, MD, PhD, C. Lynne Burek, PhD, Jobert Barin, BS, and Monica Talor, MS (Johns Hopkins University Immune Diseases Laboratory); and for the Data and Safety Monitoring Board, Paul L. Canner, PhD (Maryland Medical Research Institute), Doyt L. Conn, MD (Safety Officer) (Emory University), Jack H. Klippel, MD (Chair) (Arthritis Foundation), and J. Richard Landis, PhD (University of Pennsylvania).