

# A Composite Endpoint Measure to Consolidate Multidimensional Impact of Treatment on Gouty Arthritis

Andrew J. Sarkin<sup>1</sup>, Ari Gnanasakthy<sup>2</sup>, Rachel S. Lale<sup>1</sup>, Kyle J. Choi<sup>1</sup>, Jan D. Hirsch<sup>3</sup>

<sup>1</sup>Health Services Research Center, University of California, San Diego, USA; <sup>2</sup>Novartis Pharmaceuticals Corporation, East Hanover, USA; <sup>3</sup>Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, USA.  
Email: [asarkin@ucsd.edu](mailto:asarkin@ucsd.edu)

Received June 21<sup>st</sup>, 2013; revised July 21<sup>st</sup>, 2013; accepted July 28<sup>th</sup>, 2013

Copyright © 2013 Andrew J. Sarkin *et al.* This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## ABSTRACT

**Objective:** To create a multidimensional composite outcomes endpoint for gouty arthritis treatment in order to consolidate disparate measures of comparative effectiveness. **Methods:** One solution is to create a multidimensional composite endpoint that consolidates the complexity of outcomes into a single scale, as was done in this study. The psychometrics of the multidimensional scale and subgroup differences were investigated. **Results:** Cronbach's alpha for the multidimensional composite endpoint created in this study was 0.76, indicating good internal reliability. Similar results were found across age, race, and gender. Removing any single item did not increase Cronbach's alpha beyond 0.77, indicating that none of the items were interfering with the reliability of the scale. However, a reduction in serum urate levels was not significantly correlated with the overall multidimensional endpoint scale with that variable removed,  $r = 0.03$ ,  $p > 0.05$ . **Conclusion:** This study demonstrated the feasibility and usefulness of creating a composite multidimensional endpoint for assessing treatment outcomes among individuals with gouty arthritis.

**Keywords:** Gout; Response Criteria; Composite Endpoint

## 1. Introduction

Obtaining a comprehensive assessment of treatment impact usually requires the use of multiple outcome measures such as self-reported pain, physical limitations, flare frequency, and biochemical markers, but the results of multiple measures can be challenging to consolidate. Gouty arthritis (GA) is one such disorder where different indicators of improvement can be challenging to interpret when there are different impacts on the multiple outcomes, and GA domains are not all impacted equally across all treatments and patients who experience improvement in gout symptoms. Management of GA involves medications and lifestyle modification to prevent flares (attacks) from occurring and medications to treat acute symptoms such as pain, inflammation, and swelling when a flare does occur [1-3]. Effectiveness of treatments varies among patients who may have marked response, poor response, no response, and/or adverse reactions to the medications themselves. Different levels of response or disagreement among multiple outcome measures complicates treatment decisions, therefore causing action decisions to be based on clinical experience or

pathophysiological principles [4].

Because GA is a complex disorder that has multiple impacts on patients' quality of life [5,6] that differ by treatment type and individual patient characteristics, it would be useful to reflect this complexity by using multidimensional response criteria. Physicians and patients may not always agree on the relative importance of different outcomes [7]. In addition to the difficulties with interpretation that arise from disparate results from separate statistical tests of different response criteria, there is also an increase in the probability of making a Type I error, or a loss of statistical power if Type I error probability is kept constant by adjusting the significance criteria for the individual tests. Further, missing data can make these challenges even more complex if different patients are included in different analyses because of systematic biases in the patterns of missing data.

A possible solution is to create a composite scale measuring the multidimensional impact of treatment by combining the different outcome measures. Several studies have demonstrated the benefits of utilizing composite scale measures. To quantify and assess the multidimensional impact of asthma severity, this approach was em-

ployed to develop a Composite Asthma Severity Index (CASI), which accounts for disease symptoms and impairment, lung function, controller medication usage and frequency of hospitalizations with oral corticosteroid bursts. When validated using an independent sample, the CASI demonstrated a 32% greater magnitude of improvement within the treatment group compared with measuring symptom days alone [8]. Likewise, this approach was taken with hand osteoarthritis, where researchers combined the responsiveness of patient-reported measures (e.g. the Australian/Canadian osteoarthritis hand index [AUSCAN] and visual analogue pain subscale [VAS]), and counts of distal interphalangeal, proximal interphalangeal, metacarpophalangeal, and carpometacarpal joints to calculate patient activity and clinical disease activity composite scores. Researchers found the composite scores to be superior to the AUSCAN score in detecting the difference between the mean change in baseline values of pain, disability, and joint stiffness between the two treatment groups. The composite scores showed similar responsiveness to treatment effects as VAS pain single item measure; however, the use of composite indices appears to improve the ability to capture and quantify multiple important aspects of disease impact and activity, and may be more sensitive to detect change over time [9]. Similarly, in a study of chronic obstructive pulmonary disease (COPD), researchers hypothesized that a multidimensional grading system that assessed the respiratory, perceptible, and systemic aspects of COPD would better predict risk of death due to COPD than the use of a single physiological variable, forced expiratory volume in one second (FEV1). This was done by assigning point values from 0 - 3 to four factors that predicted risk of death (body mass index, degree of airflow obstruction, dyspnea, and exercise capacity), and then adding up the points for each factor to create the composite index score. This study showed that the sensitivity and specificity of the composite index in classifying patients with COPD as either dying or surviving was greater than FEV1 alone [C statistic of 0.74, compared to 0.65 for FEV1 alone] [10].

Because different measures may use different metrics (e.g. serum urate levels versus categorical responses to self-reported pain measures), a common metric needs to be created. One way to do this is by categorizing outcomes into response criteria. For the current study, we used the dichotomous measure of whether there was a markedly important difference on each variable, using criteria previously defined in the literature. For example, a reduction of more than two points on a 10-point pain scale or a 25% reduction in urate levels could each be response criteria. A recent Outcome Measures in Rheumatology (OMERACT) meeting asked members to give input on multidimensional response criteria for gout [11].

Core-set domains such as serum uric acid, number of tophi, flare frequency, and health assessment questionnaire disability index (HAQ-DI) that were derived from prior work with patient profiles were examined using 1000Minds™ by two groups; the gout experts and the OMERACT registrants. In the present study, we created a multidimensional composite endpoint that is disease-specific for gout based on recommendations discussed in the OMERACT results [11].

## 2. Materials and Methods

### 2.1. Participants

This analysis used pooled data from the  $\beta$ -RELIEVED program, which included patients meeting the American College of Rheumatology (ACR) 1977 preliminary criteria for acute GA, and contraindicated, intolerant, or unresponsive to non-steroidal anti-inflammatory drugs (NSAIDs) and/or colchicine. Both core studies ( $\beta$ -RELIEVED [N = 228];  $\beta$ -RELIEVED II [N = 226]) were 12-week, multiregional, active controlled, double-blind, parallel-group, double-dummy, phase 3 studies [12]. Patients were enrolled to receive a single dose canakinumab 150 mg s.c. or TA 40 mg i.m. to treat an acute GA attack and were re-dosed “on demand” on each new attack. Demographic characteristics of the participants are shown in **Table 1**.

**Table 1. Demographic characteristics of the pooled sample.**

| Characteristic               | Pooled Sample       |
|------------------------------|---------------------|
| <b>Gender, n (%)</b>         |                     |
| Male                         | 414 (91.2)          |
| Female                       | 40 (8.8)            |
| <b>Age, years</b>            |                     |
| Mean (SD)                    | 53.0 (11.7)         |
| Median (range)               | 53 (20 - 85)        |
| <b>Race, n (%)</b>           |                     |
| White                        | 343 (75.6)          |
| Black                        | 50 (11.0)           |
| Asian                        | 25 (5.5)            |
| Other                        | 36 (7.9)            |
| <b>Ethnicity, n (%)</b>      |                     |
| Hispanic or Latino           | 30 (6.6)            |
| Mixed Ethnicity              | 12 (2.6)            |
| <b>Weight (kg)</b>           |                     |
| Mean (SD)                    | 98.0 (19.0)         |
| Median (range)               | 95.5 (57.9 - 170.5) |
| <b>Height (cm)</b>           |                     |
| Mean (SD)                    | 175.5 (8.6)         |
| Median (range)               | 176 (140.0 - 200.6) |
| <b>BMI, kg/m<sup>2</sup></b> |                     |
| Mean (SD)                    | 31.7 (5.2)          |
| Median (range)               | 30.9 (19.5 - 44.9)  |

## 2.2. Measures

In addition to clinical measures such as serum urate and recorded flares, patients also reported frequency of flares, number of flares during the past four weeks, and global treatment response. They also completed the Gout Impact Scales (GIS) and Short Form-36 v2 (SF-36; Acute Form). The SF-36 was completed only by patients who reported GA symptoms in their lower extremities. Due to the lack of available translations required for all the study centers, the GIS was completed by participants where their preferred language was available. In addition to the two questionnaires, patients also responded to the separate questions pertaining to their overall experience of gout shown in **Table 2**.

### 2.2.1. Gout Impact Scales (GIS)

The GIS contains five scales; three assessing the impact of GA overall (Gout Concern Overall, Gout Medication Side Effects, Unmet Gout Treatment Need) and two assessing the impact of GA during an attack [13]. Response options for GIS items are on a five-point Likert scale (*i.e.*, from strongly agree to strongly disagree or all of the time to none of the time). GIS scales are scored from 0 to 100, with higher scores on each scale indicating “worse condition” or “greater gout impact”.

### 2.2.2. Short Form-36 (SF-36)

The SF-36 v2 (Acute Form) is a widely used measure in clinical trials assessing health-related quality of life by assessing recent function and symptoms, including pain. Normative based scoring is used with a scale range between 0 and 100, where higher scores indicate higher levels of well-being [14].

## 2.3. Procedures

The composite response endpoint representing overall

change in GA related health outcomes from baseline to 12 weeks included clinical markers (serum urate and flare activity), patient-reported data from the Gout Impact Scale (GIS) of the Gout Assessment Questionnaire 2.0 (including 6 items related to pain and quality of life), and the SF-36 bodily pain scale. Variables were chosen based on expert opinion, including the published literature and the results of the OMERACT meeting [11]. The 12 items representing five domains are shown in **Table 3** along with the criteria used as the responder definition. For each variable, the markedly important difference was determined based on published research and/or expert opinion [15-17].

A total score was calculated in two ways for each patient. One composite score was calculated as the percentage of all response criteria met out of the total number of response criteria. However, patients who could not be evaluated on each response criteria because of missing values would have scores that might underestimate their true improvement using this first method. Thus, a second composite was calculated as the percentage of response criteria for which data were available for each patient. This second method of calculating the score is less likely to be influenced by missing data, assuming that the amount and nature of missing data is the same between treatment groups. Reliability of the whole scale including all response criteria was measured by Cronbach’s alpha, which was also calculated with each criterion removed. Corrected item-total correlations were also calculated as the relationship between each variable and the total number of other criteria met, in order to evaluate how the variables related to the overall construct of GA improvement.

## 3. Results

The correlation between the two ways of handling missing data when calculating the composite scale was 0.63

**Table 2. Self-report questions related to patient’s overall experience and impact of gout symptoms.**

| Item/Question  | Response Options    |      |      |      |           |           |   |   |   |   |    |                         |
|--|---------------------|------|------|------|-----------|-----------|---|---|---|---|----|-------------------------|
| Because of your gout, how would you rate your <u>physical health</u> in the past 4 weeks?                                      | Very Poor           | Poor | Fair | Good | Very Good | Excellent |   |   |   |   |    |                         |
|  | O                   | O    | O    | O    | O         | O         |   |   |   |   |    |                         |
| Because of your gout, how would you rate your <u>quality of life</u> in the past 4 weeks?                                      | Very Poor           | Poor | Fair | Good | Very Good | Excellent |   |   |   |   |    |                         |
|  | O                   | O    | O    | O    | O         | O         |   |   |   |   |    |                         |
| Because of your gout, how would you rate your <u>mental health</u> in the past 4 weeks?  | Very Poor           | Poor | Fair | Good | Very Good | Excellent |   |   |   |   |    |                         |
|  | O                   | O    | O    | O    | O         | O         |   |   |   |   |    |                         |
| Considering all the ways gout affects you, circle a number on the scale for how well you have been doing for the past 4 weeks. | No Disease Activity | 1    | 2    | 3    | 4         | 5         | 6 | 7 | 8 | 9 | 10 | Severe Disease Activity |
| Circle a number on the scale indicating the severity of pain you have experienced within the past 4 weeks.                     | No Pain             | 1    | 2    | 3    | 4         | 5         | 6 | 7 | 8 | 9 | 10 | Severe Pain             |

(n = 454). Cronbach's alpha for the multidimensional composite endpoint was 0.76 for the 93 participants who had no missing data on any of the response criteria, indicating good internal reliability despite the breadth of the measure.

As shown in **Table 3**, removing any single item did not increase Cronbach's alpha beyond 0.77, indicating that none of the items were interfering with the reliability of the scale by not belonging with the others. However, **Table 3** shows that a reduction in serum urate levels was not significantly correlated with the overall multidimensional endpoint scale with that variable removed ( $r = 0.03$ ,  $p > 0.05$ ), indicating that reduction in urate levels was not associated with changes in other outcome measures used in this study.

Cronbach's alpha was similar for subgroups, showing that reliability was consistent across age, race, and gender. People older than 53 (the mean and median for age) demonstrated a reliability of 0.77 while people up to 53 had a reliability of 0.76. The reliabilities of Caucasian, Black, and Asian people were 0.76, 0.80, and 0.79, respectively. Cronbach's alpha indicated that the reliability was 0.76 for males and 0.66 for females.

#### 4. Discussion

This analysis supports the creation of a multidimensional composite outcomes endpoint for GA treatment in order to consolidate disparate measures of comparative effectiveness. The multidimensional scale was reliable across

age, race, and gender groups. By defining dichotomous endpoints based on markedly important differences and combining them into a scale, a single metric is created for judging the difference between treatments, thus addressing the confusion that may arise when outcome measures show varying levels of evidence of treatment impact. Consolidation enhances the ability to summarize, interpret, and communicate the overall treatment impact.

Items making up the composite measure were chosen based on expert opinion, mostly drawing from recommendations of the OMERACT consensus project [11], providing face validity for the multidimensional scale. The scale also showed good internal reliability. Although less related to the overall scale in this study, serum urate level change is probably still an important part of any multidimensional assessment of GA, despite its lack of relationship to the overall multidimensional composite endpoint scale in this study. The treatment in this study was not designed to have its main mechanism through serum urate levels, but many important GA treatments do have their impact through serum urate levels and this was one of the key indicators that came from the OMERACT participants. It is debatable whether the scale should include serum urate, which is more of a medication "process outcome" that mediates the effect of treatment rather than a "health outcome" that is experienced by the patient. In this case, very few patients (7.7%) had a change in serum urate. Indeed, the best future treatment for GA may not involve lowering serum urate, but rather making sure excess urate does not impact health by causing GA.

**Table 3. Response criteria, item-total correlations, and Cronbach's alpha with criteria removed for the multidimensional composite endpoint.**

| Domain                       | Variable  | Response Criterion             | Overall % responder | Cronbach's alpha if item removed* | Corrected item-total correlation |
|------------------------------|---|--------------------------------|---------------------|-----------------------------------|----------------------------------|
| Urate                        | Serum urate   | >25% reduction                 | 7.7 (n = 403)       | 0.77                              | 0.03                             |
|                              | Flare past 4 weeks                                  | No                             | 79.1 (n = 225)      | 0.74                              | 0.49                             |
| Flare frequency              | New flare during trial                              | No                             | 61.2 (n = 454)      | 0.77                              | 0.24                             |
|                              | Use of rescue medications                           | No                             | 48.5 (n = 454)      | 0.77                              | 0.20                             |
| Pain                         | Gout pain severity past 4 weeks (GIS, 1 - 10 scale) | >2 point reduction             | 79.6 (n = 226)      | 0.73                              | 0.50                             |
|                              | Bodily pain (SF-36, 0 - 100 scale)                  | >10 point reduction            | 62.5 (n = 373)      | 0.72                              | 0.57                             |
|                              | How well doing past 4 weeks (GIS, 1 - 10 scale)     | >2 point reduction             | 63.7 (n = 226)      | 0.74                              | 0.47                             |
| Patient global response      | Global treatment response                           | Acceptable, good, or excellent | 89.9 (n = 424)      | 0.74                              | 0.42                             |
|                              | GIS Global Control Scale (GIS, 0 - 100)             | >8 points                      | 75.8 (n = 227)      | 0.76                              | 0.28                             |
| HRQoL (Disease specific)     | Gout related quality of life                        | >1 point improvement           | 28.7 (n = 150)      | 0.73                              | 0.56                             |
|                              | Gout related physical health                        | >1 point improvement           | 24.5 (n = 143)      | 0.74                              | 0.45                             |
| (GIS, very poor - excellent) | Gout related mental health                          | >1 point improvement           | 19.4 (n = 139)      | 0.73                              | 0.53                             |

\*Cronbach's alpha for the whole scale is 0.76

When studies do not have sufficient power to show significance for all endpoints, different conclusions may be drawn for different endpoints, and this can be exacerbated by the presence of missing data. Creating a multidimensional composite endpoint in this way improves handling of missing data as well as facilitates interpretation and communication of findings. For example, the presence of missing data could have been problematic in this study where different measures were given to different participants, but the composite endpoint allowed inclusion of all participants despite missing data points, and enabled the interpretation and communication of the treatment much more efficiently and effectively. When using this method to compare treatments, it is important to verify that the amount and nature of missing data do not vary by treatment group assignment. Even with random assignment, the treatment differences could be related to differential attrition [18]. If some response criteria are more difficult to meet because of missing data, then missing data can still bias scores.

The primary strength of this method is the ability to consolidate disparate results across multiple outcome measures. A limitation of this study was reliance on a single sample of patients receiving a specific type of treatment; results may differ in other patient groups. Like many instruments, this multidimensional index has multiple measures of some domains. In order to confirm that the results are not driven by any single domain, results can be checked by eliminating any item or domain. Decisions to give all items in the composite endpoint the same weight and to include only disease-specific quality of life measures are areas open to further investigation.

## 5. Conclusion

In summary, creation of composite multidimensional endpoints should be useful for existing data, ongoing studies, and future study designs. GA researchers usually obtain responses to these items and there is minimal burden for calculating the response criteria. The score could be calculated retrospectively in existing databases allowing for new analyses, especially when results were difficult to interpret. Future studies should explore better determination and weighting of the criteria and possible inclusion of other measures. Although there may be aspects of GA that were not represented, this study demonstrated the usefulness of calculating a composite endpoint as a way to examine the multidimensional impact of treatments in clinical trials, and as an individual clinical indicator of treatment success in healthcare settings.

## 6. Acknowledgements

The authors acknowledge Erik Groessl, Theodore Ganiats, Poorva Nemlekar, Kimberly Center, and Jennalee Wool-

bridge for contributions to editing the manuscript and Novartis Pharmaceuticals for funding the study.

## REFERENCES

- [1] K. M. Jordan, J. S. Cameron, M. Snaith, W. Zhang, M. Doherty, J. Seckl, A. Hingorani, R. Jaques and G. Nuki, "British Society for Rheumatology and British Health Professionals in Rheumatology Guideline for the Management of Gout," *Rheumatology*, Vol. 46, No. 8, 2007, pp. 1372-1374. [doi:10.1093/rheumatology/kem056a](https://doi.org/10.1093/rheumatology/kem056a)
- [2] W. Zhang, M. Doherty, T. Bardin, E. Pascual, V. Basrkova, P. Conghan, P. J. Gerster, J. Jacobs, B. Leeb, F. Lioté, G. McCarthy, P. Netter, G. Nuki, F. Perez-Ruiz, A. Pignone, J. Pimentão, L. Punzi, E. Roddy, T. Uhliq and I. Zimmermann-Görska, "EULAR Evidence Based Recommendations for Gout. Part II: Management. Report of a Task Force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ES-CISIT)," *Annals of Rheumatic Diseases*, Vol. 65, No. 10, 2006, pp. 1312-1324. [doi:10.1136/ard.2006.055269](https://doi.org/10.1136/ard.2006.055269)
- [3] T. Neogi, "Clinical Practice. Gout," *New England Journal of Medicine*, Vol. 364, No. 5, 2011, pp. 443-452. [doi:10.1056/NEJMc1001124](https://doi.org/10.1056/NEJMc1001124)
- [4] H. El-zawawy and B. F. Mandell, "Managing Gout: How Is It Different in Patients with Chronic Kidney Disease?" *Cleveland Clinic Journal of Medicine*, Vol. 77, No. 12, 2010, pp. 919-928. [doi:10.3949/ccjm.77a.09080](https://doi.org/10.3949/ccjm.77a.09080)
- [5] S. J. Lee, R. Terkeltaub, D. Khannam, J. A. Singhm A. J. Sarkin and A. Kavanaugh, "Perceptions of Disease and Health-Related Quality of Life among Patients with Gout," *Rheumatology*, Vol. 48, No. 5, 2009, pp. 582-586. [doi:10.1093/rheumatology/kep047](https://doi.org/10.1093/rheumatology/kep047)
- [6] J. D. Hirsch, R. Terkeltaub, D. Khanna, J. Singh, A. Sarkin, M. Shieh, A. Kavanaugh and S. J. Lee, "Impact of Gout on Health Related Quality of Life and the Relationship to Gout Characteristics," *Patient Related Outcome Measures*, Vol. 1, 2010, pp. 1-8. [doi:10.2147/PROM.S8310](https://doi.org/10.2147/PROM.S8310)
- [7] J. Sarkin, A. E. Levack, M. M. Shieh, A. F. Kavanaugh, D. Khanna, J. A. Singh, R. A. Terkeltaub, S. J. Lee and J. D. Hirsch, "Predictors of Doctor-Rated and Patient-Rated Gout Severity: Gout Impact Scales Improve Assessment," *Journal of Evaluation in Clinical Practice*, Vol. 16, No. 6, 2010, pp. 1244-1247. [doi:10.1111/j.1365-2753.2009.01303.x](https://doi.org/10.1111/j.1365-2753.2009.01303.x)
- [8] J. J. Wildfire, P. J. Gergen, C. A. Sorkness, H. E. Mitchell, A. Calatroni, M. Kattan, S. J. Szeffler, S. J. Teach, G. R. Bloomberg, R. A. Wood, A. H. Liu, J. A. Liu, J. A. Pongracic, J. F. Chmiel, K. Conroy, Y. Rivera-Sanchez, W. W. Busse and W. J. Morgan, "Development and Validation of the Composite Asthma Severity Index—An Outcome Measure for Use in Children and Adolescents," *Journal of Allergy and Clinical Immunology*, Vol. 129, No. 3, 2012, pp. 694-701. [doi:10.1016/j.jaci.2011.12.962](https://doi.org/10.1016/j.jaci.2011.12.962)
- [9] I. K. Haugen, B. Slatkowsky-Christensen, J. Lessem and T. K. Kvien, "The Responsiveness of Joint Counts, Patient-Reported Measures and Proposed Composite Scores in Hand Osteoarthritis: Analyses from a Placebo-Con-

- trolled Trial,” *Annals of Rheumatic Diseases*, Vol. 69, No. 8, 2010, pp. 1436-1440. [doi:10.1136/ard.2008.100156](https://doi.org/10.1136/ard.2008.100156)
- [10] B. R. Celli, C. G. Cote, J. M. Marin, C. Casanova, M. Montes de Oca, R. A. Mendez, V. P. Plata and H. J. Cabral, “The Body-Mass Index, Airflow Obstruction, Dyspnea, and Exercise Capacity Index in Chronic Obstructive Pulmonary Disease,” *New England Journal of Medicine*, Vol. 350, No. 10, 2004, pp. 1005-1012. [doi:10.1056/NEJMoa021322](https://doi.org/10.1056/NEJMoa021322)
- [11] W. J. Taylor, J. A. Singh, K. G. Saag, N. Dalbeth, P. A. MacDonald, N. L. Edwards, L. S. Simon, L. K. Stamp, T. Neogi, A. L. Gaffo, P. P. Khanna, M. A. Becker and H. R. Schumacher, “Bringing It All Together: A Novel Approach to the Development of Response Criteria for Chronic Gout Clinical Trials,” *Journal of Rheumatology*, Vol. 38, No. 7, 2011, pp. 1467-1470. [doi:10.3899/jrheum.110274](https://doi.org/10.3899/jrheum.110274)
- [12] N. Schlesinger, R. E. Alten, T. Bardin, H. R. Schumacher, M. Bloch, A. Gimona, G. Krammer, V. Murphy, D. Richard and A. K. So, “Canakinumab for Acute Gouty Arthritis in Patients with Limited Treatment Options: Results from Two Randomised, Multicentre, Active-Controlled, Double-Blind Trials and Their Initial Extensions,” *Annals of Rheumatic Diseases*, Vol. 71, No. 11, 2012, pp. 1839-1848. [doi:10.1136/annrheumdis-2011-200908](https://doi.org/10.1136/annrheumdis-2011-200908)
- [13] J. D. Hirsch, S. J. Lee, R. Terkeltaub, D. Khanna, J. Singh, A. Sarkin, J. Harvey and A. Kavanaugh, “Evaluation of an Instrument Assessing Influence of Gout on Health-Related Quality of Life,” *Journal of Rheumatology*, Vol. 35, No. 12, 2008, pp. 2406-2414. [doi:10.3899/jrheum.080506](https://doi.org/10.3899/jrheum.080506)
- [14] J. E. Ware, “SF-36 Health Survey Update,” *Spine*, Vol. 25, No. 24, 2000, pp. 3130-3139. [doi:10.1097/00007632-200012150-00008](https://doi.org/10.1097/00007632-200012150-00008)
- [15] M. Kosinski, S. Z. Zhao, S. Dedhiya, J. T. Osterhaus and J. E. Ware, “Determining Minimally Important Changes in Generic and Disease-Specific Health-Related Quality of Life Questionnaires in Clinical Trials of Rheumatoid Arthritis,” *Arthritis & Rheumatism*, Vol. 43, No. 7, 2000, pp. 1478-1487. [doi:10.1002/1529-0131\(200007\)43:7<1478::AID-ANR10>3.0.CO;2-M](https://doi.org/10.1002/1529-0131(200007)43:7<1478::AID-ANR10>3.0.CO;2-M)
- [16] D. Khanna, A. J. Sarkin, P. P. Khanna, M. M. Shieh, A. F. Kavanaugh, R. A. Terkeltaub, *et al.*, “Minimally Important Differences of the Gout Impact Scale in a Randomized Controlled Trial,” *Rheumatology (Oxford)*, Vol. 50, No. 7, 2011, pp. 1331-1336. [doi:10.1093/rheumatology/ker023](https://doi.org/10.1093/rheumatology/ker023)
- [17] B. Bruce and J. F. Fries, “The Stanford Health Assessment Questionnaire: A Review of Its History, Issues, Progress, and Documentation,” *Journal of Rheumatology*, Vol. 30, No. 1, 2003, pp. 167-178.
- [18] A. J. Sarkin, S. R. Tally, T. A. Cronan, G. E. Matt and H. W. Lyons, “Analyzing Attrition in a Community-Based Literacy Program,” *Evaluation and Program Planning*, Vol. 20, No. 4, 1997, pp. 421-431. [doi:10.1016/S0149-7189\(97\)00023-2](https://doi.org/10.1016/S0149-7189(97)00023-2)