Tophus Measurement as an Outcome Measure for Clinical Trials of Chronic Gout: Progress and Research Priorities

NICOLA DALBETH, FIONA M. McQUEEN, JASVINDER A. SINGH, PATRICIA A. MacDONALD, N. LAWRENCE EDWARDS, H. RALPH SCHUMACHER Jr, LEE S. SIMON, LISA K. STAMP, TUHINA NEOGI, ANGELO L. GAFFO, PUJA P. KHANNA, MICHAEL A. BECKER, and WILLIAM J. TAYLOR

ABSTRACT. Despite the recognition that tophus regression is an important outcome measure in clinical trials of chronic gout, there is no agreed upon method of tophus measurement. A number of methods have been used in clinical trials of chronic gout, from simple physical measurement techniques to more complex advanced imaging methods. This article summarizes methods of tophus measurement and discusses their properties. Physical measurement using Vernier calipers meets most aspects of the Outcome Measures in Rheumatology (OMERACT) filter. Rigorous testing of the complex methods, particularly with respect to reliability and sensitivity to change, is needed to determine the appropriate use of these methods. Further information is also required regarding which method of physical measurement is best for use in future clinical trials. The need to develop and test a patient-reported outcome measure of tophus burden is also highlighted. (J Rheumatol 2011;38:1458–61; doi:10.3899/jrheum.110272)

Key Indexing Terms:
GOUT TOPHUS OUTCOME MEASURE

The tophus is a pathognomonic feature of chronic gout. This lesion represents a chronic inflammatory response to monosodium urate (MSU) crystals deposited most often within the subcutaneous (SC) tissues or the joint[1]. Tophi have been implicated in the pathogenesis of joint damage in chronic gout and are strongly associated with disability in this disease[2,3]. Thus, regression of tophi is an important therapeutic outcome for patients with chronic tophaceous gout. In the recent Delphi exercise regarding outcome measures for clinical trials, tophus regression was identified as one of the core domains for studies of chronic gout[4]. This domain was also endorsed by plenary voting at the 2009 Outcome Measures in Rheumatology meeting (OMERACT 9)[5].

Despite recognition that tophus regression is an important outcome measure in clinical trials of chronic gout, there is no agreed upon method of tophus measurement[6]. This article summarizes the methods that have been used in chronic gout studies and discusses the properties of these methods. Discussion from OMERACT 10 breakout groups on tophus measurement is summarized, and results of plenary voting are shown. Finally, the future research agenda is highlighted.

TOPHUS MEASUREMENT

Methods

The 8 methods of tophus measurement reported to date are listed in Table 1. These methods can be separated into 2

From the Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand; Birmingham Veterans Affairs (VA) Medical Center and University of Alabama, Birmingham, AL, USA; Takeda Global Research and Development, Deerfield, IL; Department of Medicine, University of Florida, Gainsville, FL; University of Pennsylvania and VA Medical Center, Philadelphia, PA; SDG LLC, Cambridge, MA, USA; Department of Medicine, University of Otago, Christchurch, New Zealand; Boston University School of Medicine, Boston, MA; Division of Rheumatology, David Geffen School of Medicine, University of California, Los Angeles, CA; Rheumatology Section, University of Chicago, Chicago, IL, USA; and Department of Medicine, University of Otago, Wellington, New Zealand.

Supported by resources and use of facilities at the Birmingham VA Medical Center, Alabama, USA (Dr. Singh). Dr. Khanna is the recipient of an American College of Rheumatology REF Clinical Investigator Award (Quality of Life and Healthcare Utilization in Chronic Gout). Dr. Neogi is supported by NIAMS 5K23AR55127. Dr. Dalbeth has received consultant fees from Takeda and Novartis. Dr. Singh has received speaker honoraria from Abbott; research and travel grants from Allergan, Takeda, Savient, Wyeth, and Amgen; and consultant fees from Savient, URL Pharmaceuticals, and Novartis. Dr. Simon has served on the Board of Directors for Savient Pharmaceuticals and as a consultant for Takeda. Ms MacDonald is an employee of Takeda. Dr. Becker has received consultant fees from Takeda, Savient, BioCryst, URL Pharmaceuticals, Ardea, and Regeneron.

N. Dalbeth, MD, FRACP, Associate Professor; F.M. McQueen, MD, FRACP, Professor, Faculty of Medical and Health Sciences, University of Auckland; J.A. Singh, MBBS, MPH, Associate Professor, Birmingham VA Medical Center and University of Alabama; P.A. MacDonald, BSN, NP, Associate Director of Clinical Science, Takeda Global Research and Development; N.L. Edwards, MD, Professor of Medicine, Department of Medicine, University of Florida; H.R. Schumacher Jr, MD, Professor of Medicine, University of Pennsylvania and VA Medical Center; L.S. Simon, MD, Professor of Medicine, SDG, LLC; L.K. Stamp, PhD, FRACP, Associate Professor, Department of Medicine, University of Otago; T. Neogi, MD, PhD, FRCPC, Associate Professor, Boston University School of Medicine; A.L. Gaffo, MD, MSPH, Assistant Professor, Birmingham VA Medical Center and University of Alabama; P.P. Khanna, MD, MPH, Clinical Instructor, Division of Rheumatology, David Geffen School of Medicine, University of California; M.A. Becker, MD, Professor Emeritus of Medicine, Rheumatology Section, University of Chicago; W.J. Taylor, Associate Professor, Department of Medicine, University of Otago.

Address correspondence to Dr. Dalbeth.
E-mail: n.dalbeth@auckland.ac.nz

Table 1. Quantitative methods of tophus measurement.

Method	Category	Original Study
Counting the total number of subcutaneous tophi	Simple	Becker[8]
Tape measurement of subcutaneous tophus area	Simple	Schumacher[11]
Vernier calipers for measurement of subcutaneous tophus diameter	Simple	Perez-Ruiz[12]
Digital photography for measurement of subcutaneous tophus area	Complex	Maroli[13]
Ultrasonography for measurement of tophus diameter and volume	Complex	Perez-Ruiz[14]
Magnetic resonance imaging for measurement of tophus volume	Complex	Schumacher[15]
Conventional computed tomography (CT) for measurement of tophus volume	Complex	Dalbeth[16]
Dual energy CT for measurement of tophus volume	Complex	Choi[17]

groups: simple methods of physical measurement that allow assessment of SC tophi, and more complex methods requiring specialized equipment or technology. Most complex methods involve advanced imaging techniques and allow for assessment of both SC and intraarticular tophi. The complex imaging methods may have greater sensitivity and validity, as changes within bone and joints are also identified. A detailed guide and atlas have recently been compiled to allow standardization of these methods, for use in future clinical trials[7].

Properties

A systematic review has been completed to identify quantitative methods of tophus measurement and to summarize the properties of the various methods of tophus measurement in detail according to the OMERACT filter[7]. The findings of this review were used as working documents in the tophus breakout sessions at the OMERACT 10 meeting. The following is a summary of discussion in the breakout groups regarding properties of the various methods of tophus measurement.

Counting the total number of SC tophi. This method, which involves counting all visible tophi at each study visit, has been used in clinical trials of febuxostat[8,9]. Sensitivity to change and between-group discrimination have been demonstrated[9,10], but reliability data are not available. Further, there are no available data regarding construct and criterion validity (such as how this measure relates to other measures of tophus size, and whether lesions counted as tophi contain MSU crystals). The breakout groups noted that although this method is feasible and simple, sensitivity to change may be limited in the early phase of a study, when tophus size has been reduced but tophi have not yet disappeared.

Tape measurement of SC tophus area. This method involves identification and serial measurement of an index tophus. The length and width axes are measured using a tape measure, and these values are then multiplied to provide an area measurement[11]. This method has been used in clinical trials of febuxostat[8,9,10]. Tape measurement of SC tophus area has high reliability and is sensitive to change[9,11], but did not differentiate between allopurinol- and febuxostat-treated groups, possibly because there was no difference[8]. Further, there are no available data regarding construct and criterion validity. The breakout groups also commented that measurement error may be greater using this method, compared with other physical methods such as measurement in a single axis using Vernier calipers, as 2 separate measurements are recorded. This issue has not been addressed in validation studies to date.

Vernier caliper measurement of SC tophus diameter. This method involves identification and serial measurement of an index tophus. The longest diameter of the index tophus is measured using Vernier calipers[12]. The properties of this method according to the OMERACT filter are summarized in Table 2. This method was considered by the breakout groups to fulfil most aspects of the OMERACT filter, although it should be noted that this method has not been used as an outcome measure in a randomized controlled trial in chronic gout. Voting at the OMERACT 10 plenary session confirmed the conclusions of the breakout groups, with 56/68 (83%) respondents agreeing that tophus maximum diameter measurement by Vernier caliper meets the OMERACT filter for truth, discrimination, and feasibility.

Digital photographic assessment of SC tophus size. This method involves standardized digital photography and computer-assisted measurement of SC tophi by a central reader. This method has been used in the phase 3 clinical trials of pegloticase with change reported as categories of response[13]. Sensitivity to change and between-group discrimination have been demonstrated, but reliability data and validity data are not yet available. The breakout groups agreed that this was likely to be a highly reliable method due the storage of the images in an electronic record, which allows for central reading and cross-checking of the data. However, comparison with the simple physical methods was considered important for further validation.

Ultrasonography measurement of tophus diameter and volume. This method involves identification and serial measurement of an index tophus using ultrasonography, and has been assessed in a longitudinal observational study of patients receiving urate-lowering therapy[14]. Measurements using ultrasonography correlate highly with magnetic resonance imaging (MRI) measurements, and intra and interobserver reliability is good–excellent (intraclass correlation coefficient > 0.82). Sensitivity to change has been demonstrated, but this method was not able to discriminate between different urate-lowering therapies, possibly because there was no differ-

Table 2. Properties of the Vernier caliper method according to the OMERACT filter.

	Findings	Data source	Study
Feasibility	Low cost, short acquisition time, high patient acceptability	Observational cross-sectional study	Dalbeth[16]
Construct validity	High correlation between computed tomography and Vernier caliper measurement ($r = 0.91$, $p < 0.001$)	Observational cross-sectional study	Dalbeth[16]
Reliability	Intraobserver intraclass correlation coefficient 1.0 (95% CI 0.99, 1.0); interobserver intraclass correlation coefficient 0.99 (95% CI 0.97, 0.99)	Observational cross-sectional study	Dalbeth[16]
Change sensitivity	All index tophi completely resolved with urate-lowering therapy after mean 21 months; effect size 1.83	Non-randomized parallel treatment study	Perez-Ruiz[12]
Between-group discrimination	Faster resolution in benzbromarone treatment group vs allopurinol treatment group (velocity of regression 1.21 mm/month vs 0.57 mm/month, $p < 0.01$)	Non-randomized parallel treatment study	Perez-Ruiz[12]

ence[14]. The breakout groups considered that this method might provide a useful compromise between the simple and complex methods of tophus measurement, particularly in the context of increasing availability of ultrasonography within the rheumatology clinic. However, operator variability was considered a potential problem with this method in large multicenter studies.

Other imaging methods of assessing tophus size. MRI, conventional computed tomography (CT), and dual-energy computed tomography (DECT) are potentially useful methods of tophus size measurement[15,16,17]. As with ultrasonography, these methods have the capacity to measure both SC and intraarticular tophi. The breakout groups commented that these methods have the benefit of storage of raw data for central reading and cross-checking. However, these methods are expensive, time-consuming, and require specialized equipment and technical expertise. None of these methods has been assessed in prospective studies to date. Several groups have raised concerns regarding the reliability of MRI for tophus size assessment[14,15]. CT measurement of tophus volume is highly reliable and this imaging modality has the additional advantage of excellent resolution of bone erosion[2,16]. However, the use of radiation and the time required to complete the volume assessments may limit feasibility of this imaging technique. The breakout groups noted that the reliability of DECT is yet to be reported, but that this method has high face validity and might potentially be the gold standard method against which other methods could be compared and validated.

Tophus Measurement as Outcome Measure: Further Considerations

The breakout groups agreed that change in tophus size should be measured in all clinical trials of chronic gout, but that the complexity of the method used may vary depending on the research question. Reporting of raw data rather than categories of response was considered ideal in order to maintain statistical power and discriminative ability. Although tophi have been shown to be strongly associated with joint damage and disability in chronic gout, little information is currently available about the patient perceptions of tophi. Further, changes other than size, such as softening or improved mobility of a joint affected by tophus, may be of relevance to the patient but are not identified by the current methods. The breakout groups concluded that a patient-reported outcome (PRO) measure of tophus burden/impact should be developed. This conclusion was supported by plenary voting at the OMERACT 10 meeting; 61/70 (87%) respondents agreed that a PRO of tophus burden should be a priority for the research agenda.

RESEARCH PRIORITIES

The central question for further studies of tophus measurement is whether extra information obtained from more complex methods of measurement merits the additional cost, time, and complexity. Rigorous testing of the complex methods, particularly with respect to reliability and sensitivity to change, is needed to determine the appropriate use of these methods. Further information is required regarding simple methods of physical measurement. In particular, it is not clear whether these methods perform differently, and which would be optimal for use in future clinical trials. Ideally, a simple method is needed, for use in large clinical trials, that has been validated against one or more of the complex methods in a smaller but rigorous validation exercise. Finally, there is a need to develop and test a PRO measure of tophus burden. This tool may include aspects of tophus improvement such as softening, in addition to change in size. Validation of this tool against other PRO, such as the Health Assessment Questionnaire or the Medical Outcome Study Short-Form 36, would be required. Comparison between this and other methods of tophus measurement (both physical and imaging methods) would ensure it is employed as a simple, easy to use PRO of tophus measurement in clinical trials.

REFERENCES

1. Dalbeth N, Pool B, Gamble GD, Smith T, Callon KE, McQueen FM, et al. Cellular characterization of the gouty tophus: a quantitative analysis. Arthritis Rheum 2010;62:1549-56.

2. Dalbeth N, Clark B, Gregory K, Gamble G, Sheehan T, Doyle A, et al. Mechanisms of bone erosion in gout: a quantitative analysis using plain radiography and computed tomography. Ann Rheum Dis 2009;68:1290-5.
3. Dalbeth N, Collis J, Gregory K, Clark B, Robinson E, McQueen FM. Tophaceous joint disease strongly predicts hand function in patients with gout. Rheumatology 2007;46:1804-7.
4. Taylor WJ, Schumacher HR Jr, Baraf HS, Chapman P, Stamp L, Doherty M, et al. A modified Delphi exercise to determine the extent of consensus with OMERACT outcome domains for studies of acute and chronic gout. Ann Rheum Dis 2008;67:888-91.
5. Schumacher HR, Taylor W, Edwards L, Grainger R, Schlesinger N, Dalbeth N, et al. Outcome domains for studies of acute and chronic gout. J Rheumatol 2009;36:2342-5.
6. Grainger R, Taylor WJ, Dalbeth N, Perez-Ruiz F, Singh JA, Waltrip RW, et al. Progress in measurement instruments for acute and chronic gout studies. J Rheumatol 2009;36:2346-55.
7. Dalbeth N, Schauer C, MacDonald P, Perez-Ruiz F, Schumacher HR, Hamburger S, et al. Methods of tophus assessment in clinical trials of chronic gout: a systematic literature review and pictorial reference guide. Ann Rheum Dis 2011;70:597-604.
8. Becker MA, Schumacher HR Jr, Wortmann RL, MacDonald PA, Eustace D, Palo WA, et al. Febuxostat compared with allopurinol in patients with hyperuricemia and gout. N Engl J Med 2005;353:2450-61.
9. Becker MA, Schumacher HR Jr, MacDonald PA, Lloyd E, Lademacher C. Clinical efficacy and safety of successful longterm urate lowering with febuxostat or allopurinol in subjects with gout. J Rheumatol 2009;36:1273-82.
10. Schumacher HR Jr, Becker MA, Wortmann RL, MacDonald PA, Hunt B, Streit J, et al. Effects of febuxostat versus allopurinol and placebo in reducing serum urate in subjects with hyperuricemia and gout: a 28-week, phase III, randomized, double-blind, parallel-group trial. Arthritis Rheum 2008;59:1540-8.
11. Schumacher HR Jr, Becker MA, Palo WA, Streit J, MacDonald PA, Joseph-Ridge N. Tophaceous gout: quantitative evaluation by direct physical measurement. J Rheumatol 2005;32:2368-72.
12. Perez-Ruiz F, Calabozo M, Pijoan JI, Herrero-Beites AM, Ruibal A. Effect of urate-lowering therapy on the velocity of size reduction of tophi in chronic gout. Arthritis Rheum 2002;47:356-60.
13. Maroli AN, Waltrip R, Alton M, Herbert S, Baraf B, Huang B, et al. First application of computer-assisted analysis of digital photographs for assessing tophus response: phase 3 studies of pegloticase in treatment failure gout [abstract]. Arthritis Rheum 2009;60 Suppl:S416.
14. Perez-Ruiz F, Martin I, Canteli B. Ultrasonographic measurement of tophi as an outcome measure for chronic gout. J Rheumatol 2007;34:1888-93.
15. Schumacher HR Jr, Becker MA, Edwards NL, Palmer WE, MacDonald PA, Palo W, et al. Magnetic resonance imaging in the quantitative assessment of gouty tophi. Int J Clin Pract 2006;60:408-14.
16. Dalbeth N, Clark B, Gregory K, Gamble GD, Doyle A, McQueen FM. Computed tomography measurement of tophus volume: comparison with physical measurement. Arthritis Rheum 2007;57:461-5.
17. Choi HK, Al-Arfaj AM, Eftekhari A, Munk PL, Shojania K, Reid G, et al. Dual energy computed tomography in tophaceous gout. Ann Rheum Dis 2009;68:1609-12.

Papers presented at the OMERACT 10 Conference, Kota Kinabalu, Borneo, May 4–8, 2010.

Part 1 Disease-specific Outcomes

Part 2 Patient-reported Outcomes

Part 3 Biomarker and Imaging Outcomes

Parts 2 and 3 will appear in the August and September issues.