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 TOTHEUS 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 joint1. Tophi have been implicated in the pathogenesis of joint damage in chronic gout and are strongly associated with disability in this disease2,3. 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 gout4. 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 measurement6. 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...
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.

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. 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. Sensitivity to change and between-group discrimination have been demonstrated, 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. This method has been used in clinical trials of febuxostat. Tape measurement of SC tophus area has high reliability and is sensitive to change, but did not differentiate between allopurinol- and febuxostat-treated groups, possibly because there was no difference. 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. The properties of this method according to the OMERACT filter are summarized in Table 2. This method was considered by the breakout groups to fulfill 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. 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 to 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. 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 1.** Quantitative methods of tophus measurement.

<table>
<thead>
<tr>
<th>Method</th>
<th>Category</th>
<th>Original Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counting the total number of subcutaneous tophi</td>
<td>Simple</td>
<td>Becker8</td>
</tr>
<tr>
<td>Tape measurement of subcutaneous tophus area</td>
<td>Simple</td>
<td>Schumacher11</td>
</tr>
<tr>
<td>Vernier calipers for measurement of subcutaneous tophus diameter</td>
<td>Simple</td>
<td>Perez-Ruiz12</td>
</tr>
<tr>
<td>Digital photography for measurement of subcutaneous tophus area</td>
<td>Complex</td>
<td>Maroli13</td>
</tr>
<tr>
<td>Ultrasonography for measurement of tophus diameter and volume</td>
<td>Complex</td>
<td>Perez-Ruiz14</td>
</tr>
<tr>
<td>Magnetic resonance imaging for measurement of tophus volume</td>
<td>Complex</td>
<td>Schumacher15</td>
</tr>
<tr>
<td>Conventional computed tomography (CT) for measurement of tophus volume</td>
<td>Complex</td>
<td>Dalbeth16</td>
</tr>
<tr>
<td>Dual energy CT for measurement of tophus volume</td>
<td>Complex</td>
<td>Choi17</td>
</tr>
</tbody>
</table>
have been shown to be strongly associated with joint damage. Although tophi categories of response was considered ideal in order to maintain research question. Reporting of raw data rather than categorical 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


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.