Serum Urate as a Surrogate Outcome for Gout Flares: Where Do We Stand Today?

Lisa K. Stamp 1,*, Robin Christensen 2,3 and Melanie B. Morillon 2,3

1 Department of Medicine, University of Otago Christchurch, Christchurch 8011, New Zealand
2 Section for Biostatistics and Evidence-Based Research, The Parker Institute, Bispebjerg and Frederiksberg Hospital, 2000 Copenhagen, Denmark; robin.christensen@regionh.dk (R.C.); melanie.birger.morillon@rsyd.dk (M.B.M.)
3 Research Unit of Rheumatology, Department of Clinical Research, University of Southern Denmark, Odense University Hospital, 5000 Odense, Denmark
* Correspondence: lisa.stamp@cdhb.health.nz

Abstract: In gout research, serum urate has been widely accepted as the primary endpoint in clinical trials of urate-lowering therapies by both the FDA and EMA for many years. However, for serum urate to be a meaningful outcome measure, it should reflect at least one important patient-centered clinical outcome, such as gout flares. The relationship between achieving a pre-specified “target” serum urate and a corresponding improvement in patient-centered outcomes has been difficult to show due to variation in reporting of both serum urate and gout flares in clinical trials; a paradoxical rise in gout flares after starting urate-lowering therapy and a delay after achieving the pre-specified target serum urate before gout flares settle coupled with the relatively short duration of the trials. However, recent evidence from individual-level patient data from two, two-year randomized controlled trials clearly shows that achieving target urate is associated with a subsequent reduction and cessation of gout flares. In this review, we examine the evidence supporting serum urate as a surrogate outcome for gout flares, the methods, and the challenges of showing the validity of surrogacy.

Keywords: serum urate; surrogate; gout flare

1. Introduction

The primary endpoint is the outcome measure of interest that is used to determine the success of a clinical trial, whereas an outcome, in general, refers to any measure used to assess the impact of the intervention on the participants. To date, most clinical trials of urate-lowering therapies (ULT), including phase III trials for FDA and EMA drug registration, have used serum urate (SU) as the most important outcome measure (i.e., the primary endpoint). The use of a laboratory-based measure, such as SU, as the major outcome, has clear advantages in that it is apparently objectively measured, is not susceptible to performance bias, and there is ample evidence that urate-lowering therapies reduce SU when people are adherent [1,2]. The use of a laboratory-based outcome measure provides additional benefits in that it allows for shorter duration and smaller and less costly clinical trials [3].

When designing clinical trials, the primary endpoint is a specific and predefined outcome measure that is of primary interest in a clinical trial. Thus, the primary endpoint is just one of the several possible outcomes measured in a clinical trial. For SU to be a meaningful primary endpoint, it must be a validated “surrogate outcome”—a measure or biomarker used as a substitute or proxy for a clinically meaningful endpoint, such as gout flares. Contemporary gout management has been based on achieving a pre-specified target SU of <0.36 mmol/L (6 mg/dL) [4,5]. The relationship between achieving this pre-specified “target” SU and improvement in patient-centered clinical outcomes, such as a reduction in gout flares, has been difficult to show [6]. The purpose of this paper is to review the
evidence for SU, using the currently recommended target, as a surrogate marker for gout flares and why it is challenging to prove surrogacy, given the available data.

2. Definitions

A biomarker can be defined as a “characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention” [7]. The Outcome Measures in Rheumatology (OMERACT) initiative developed a preliminary schema for validation of soluble biomarkers in rheumatoid arthritis, psoriatic arthritis, and spondyloarthritis [8,9] which have been adapted for use in gout [10] (Table 1).

Table 1. The OMERACT Essential Criteria for validation soluble biomarkers adapted for serum urate in gout (Adapted from [10]).

<table>
<thead>
<tr>
<th>Essential Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Truth and Discrimination</strong></td>
</tr>
<tr>
<td>• The assay for measurement of SU is reproducible according to reliability analysis.</td>
</tr>
<tr>
<td>• The effects of sources of variability on SU concentrations in appropriate controls are known for the following core variables—age, sex, ethnicity, circadian rhythms, body mass index, renal/hepatic function, fasting/non-fasting, medications, and cardiovascular disease.</td>
</tr>
<tr>
<td>• SU demonstrates independent association with clinical and patient centered endpoints. The following key clinical/patient centered endpoints in chronic gout are: number of gout flares, tophus regression, dissolution of crystals, radiographic damage and patient function and quality of life [11].</td>
</tr>
<tr>
<td><strong>Feasibility</strong></td>
</tr>
<tr>
<td>• The assay for measurement of SU is internationally standardized (availability of reference standards) and is readily accessible if used for clinical practice.</td>
</tr>
<tr>
<td>• Stability of SU at room temperature, in frozen specimen, after repeat freeze/thaw cycles, and after long-term storage has been documented</td>
</tr>
</tbody>
</table>

A surrogate end point may be defined as “a laboratory measurement or a physical sign used as a substitute for a clinically meaningful endpoint that measures directly how a patient feels, functions or survives” [12].

A biomarker can only be considered a consistent and reliable surrogate outcome if there is evidence supporting a causal relationship between change in the (surrogate) outcome measure and a corresponding change in the patient-important outcome. Thus, the biomarker must be recognized as being in the causal pathway of the disease process, and an intervention’s entire effect on the clinical outcome of interest should be fully captured by a change in the biomarker. Unfortunately, the terminology around surrogate outcomes and biomarkers is often used interchangeably. However, it is important to recognize the key difference, namely that a surrogate can act as a replacement for one or more clinical outcomes, which is not necessarily the case for a biomarker. The question is, then, are we able to establish a causal relationship and present an approach to critical appraisal of studies using a surrogate outcome (e.g., serum urate) rather than the patient-important outcome (e.g., gout flares)?

Several methods for evaluating whether a biomarker is truly a surrogate outcome that can replace a clinical endpoint have been proposed. Bucher et al. [13] suggested the following three questions must be answered in the affirmative:

1. Is there a strong, independent, consistent association between the biomarker (potential surrogate) and the clinical patient-important outcome (that the surrogate could potentially replace)?
2. Is there evidence from randomized trials in the same drug class that improvement in the surrogate outcome has consistently led to improvement in the target outcome?
3. Is there evidence from randomized trials in other drug classes that improvement in the surrogate outcome has consistently led to improvement in the target outcome?
Another framework was presented by the Biomarker Surrogate Evaluation Schema (BSES) based on four domains: study design, target outcome, statistical evaluation, and generalizability [14,15] (Table 2). For the surrogate to be validated using the BSES, it requires high-quality statistical evidence from multiple randomized trials, a patient-centered reversible outcome, and evidence of effect across different drugs with different mechanisms of action.

Table 2. The Biomarker Surrogate Evaluation Schema (BSES) 3.

<table>
<thead>
<tr>
<th>Biomarker-Surrogate Domains</th>
<th>Score</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study design</td>
<td>0</td>
<td>Biological plausibility and lower quality clinical studies, e.g., cross-sectional observational studies</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Rank 0 and at least 2 good quality prospective observational cohort studies measuring S and T</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Rank 1 and at least 2 high quality adequately powered RCTs measuring S and T</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Rank 1 and all, and at least 5 high quality adequately powered, RCTs measuring S and T</td>
</tr>
<tr>
<td>Target outcome</td>
<td>0</td>
<td>Target is reversible disease-centered biomarker of harm</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Target is irreversible disease-centered biomarker of harm</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Target is patient-centered endpoint of reversible organ morbidity or clinical burden of disease or clinical harm</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Target is patient-centered endpoint of irreversible organ morbidity or clinical burden of disease or severe irreversible clinical harm or death</td>
</tr>
<tr>
<td>Statistical Evaluation of biosurrogate—Target (B-T) domain</td>
<td>0</td>
<td>Poor: Does not meet the criteria for Rank 1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Fair: RCT $R^2_{\text{trial}} \geq 0.2$ AND STEP $\geq 0.1$ AND $R^2_{\text{ind}} \geq 0.2$ OR cohort data $R^2_{\text{ind}} \geq 0.4$</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Good: RCT $R^2_{\text{trial}} \geq 0.4$ AND STEP $\geq 0.2$ AND $R^2_{\text{ind}} \geq 0.4$</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Excellent: RCT $R^2_{\text{trial}} \geq 0.6$ AND STEP $\geq 0.3$ AND $R^2_{\text{ind}} \geq 0.6$ (without data subdivision)</td>
</tr>
<tr>
<td>Generalizability</td>
<td>0</td>
<td>No clinical or pharmacologic evidence</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Clinical OR pharmacologic evidence</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Clinical AND pharmacologic evidence</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Consistent Clinical RCT AND pharmacologic RCT evidence</td>
</tr>
</tbody>
</table>

The two key similarities across these schemata are the need for an independent association between the surrogate and the patient-centered outcomes with more than one class of drug.

3. Serum Urate as a Biomarker

SU obviously fulfills the definition of being a biomarker in that it is a characteristic that is objectively measured and evaluated as an indicator of pharmacologic response to a therapeutic intervention, namely urate-lowering therapy. SU is also on the causal pathway for gout and hyperuricemia; that is, elevated SU above the physiological point of saturation is critical for gout to develop [16]. SU has been shown to fulfill all of the OMERACT essential criteria for a soluble biomarker, with the exception of its effects on patient-centered outcome measures [10].

4. The Gout Flare as an Important Patient-Centered Outcome in Gout

The gout flare, which is characterized by hot, swollen, exquisitely painful joint(s), reflects the acute inflammatory response to intra- or peri-articular monosodium urate crystals, which form in the presence of hyperuricemia. The gout flare is central to the patient’s experience of gout, with the severe pain affecting nearly every aspect of life, including sleep, work, and daily activities, as well as family and social activities [17,18]. As such, gout flare prevention is a key aspect of gout management for both people with gout and their healthcare providers.

5. Unique Challenges in Showing Serum Urate Is a Surrogate Outcome for Gout Flares

There are a number of factors that make showing a relationship between a pre-specified SU target or a change in SU with gout flares challenging, including the way both SU and gout flares are reported, the paradoxical increase in gout flare after commencing urate-lowering therapy and the latent period between reduction in SU and reduction in gout flares.
5.1. SU Reporting

Although SU is well established as a mandatory core outcome domain in “chronic gout” studies [19] and is the primary endpoint for the majority of ULT studies, there is no standardized way of reporting SU across clinical trials. A recent review identified 10 different ways in which SU was reported, which fit into three broad categories namely, percentage of individuals at target SU, mean SU in each group and change in SU in each group [20,21].

5.2. Gout Flare Reporting

Like SU, gout flares are considered a mandatory core outcome domain in “chronic gout” studies [19]. However, there is wide variation in the methods used to define, measure, and report gout flares [21,22]. The majority of studies use “self-reported gout flare” or “self-reported gout flare requiring treatment” as the definition of a gout flare. Although there is a validated gout flare definition that requires fulfillment of three or more of the following criteria: patient-defined flare, pain at rest score of >3 on a 0–10 numerical rating scale, presence of at least one swollen joint, and presence of at least one warm joint [23,24], it has not been widely used to date. Reporting of gout flares is also highly variable, although broadly reported as mean number of gout flares per randomized group and the proportion of participants with a certain number of flares [20–22].

5.3. Relationship between Change in Serum Urate, Gout Flares and Trial Design

It is widely recognized that gout flares can be precipitated when people with gout commence ULT. This paradoxical effect is thought to be due to the mobilization of total body urate stores and fluctuations in SU. For this reason, anti-inflammatory prophylaxis to prevent gout flares for the first six months of commencing urate-lowering therapy is recommended [4,5]. Furthermore, it is clear that it can take many months for gout flares to reduce and cease after achieving target SU. Thus, although short-term clinical trials, i.e., 6–12 months, are sufficient to show an effect of urate-lowering therapy on SU, they are of insufficient duration to show an effect on gout flares. Clinical trials of at least 18–24 months are required in order to detect a reliable reduction in gout flares. Given that SU is currently the primary endpoint in most clinical trials, there are very few of these longer-duration trials that provide adequate data to examine the relationship between SU and gout flares.

6. What Is the Current Evidence That SU Is a Valid Surrogate for Gout Flares?

For the aforementioned reasons, providing statistical evidence that SU is a valid surrogate for gout flares using currently available data has been challenging. A systematic literature review and meta-regression analysis published in 2018 identified only 10 RCTs, of which the longest was 24 months in duration, including a total of 6100 people [6]. Meta-regression analysis did not support any statistically significant association between the difference in proportions of individuals who achieved SU < 6 mg/dL and the observed flare risk ratio (Figure 1).

When subsequent analyses were undertaken to account for study duration, the delay between achieving target SU and reduction in gout flares by using a subset of six studies, there was an association between the proportion of individuals achieving SU target (<6 mg/dL) and the risk of having gout flares (relative to “patient-years”) and the duration of ULT was inversely associated with the proportion of patients experiencing a gout flare (Figure 2A,B, respectively).

Using these data, SU failed to meet the BSES criteria for a valid surrogate based on the statistical evaluation criteria.

Further analyses using individual-level patient data from two two-year RCTs of allopurinol [1,2] for the management of gout have also been undertaken [25]. Although these studies were not specifically designed to answer the question of whether SU was a surrogate for gout flares, the two-year study duration facilitated analysis based on
achieving SU urate responder status in months 6–12 and gout flares in the subsequent 12-month period.

![Figure 1. Relationship between proportions of individuals achieving target SU and flare rate (with permission from [6]).](image)

![Figure 2. Association between (A) proportion achieving SU < 6 mg/dL and gout flares per person years and (B) time of flares assessment after commencing ULT and proportion with flares (with permission from [6]).](image)

In this analysis, significantly fewer SU responders (defined as individuals who on average achieved a SU < 6 mg/dL [0.36 mmol/L] based on data at 6, 9, and 12 months post-baseline) had a gout flare than did SU non-responders between 12 and 24 months (91 [27%] of 343 vs. 156 [64%] of 245; adjusted odds ratio [OR] 0.29 [95% CI 0.17 to 0.51], p < 0.0001). The mean number of flares per participant per month between 12 and 24 months was also significantly lower in SU responders than in SU non-responders (adjusted mean difference –0.41 [95% CI –1.77 to –1.04], p < 0.0001). This association was independent of the original randomized treatment allocation [25]. These results support the validity of SU as a surrogate outcome for gout flares. However, to date, we have been unable to access data from urate-lowering therapies with a different mechanism of action than allopurinol of sufficient duration to fulfill the Bucher criteria.

7. Where Do We Stand Today?

The recent individual patient data from two allopurinol RCTs show a clear relationship between achieving target urate and gout flare reduction [25]. However, the association between lowering SU and the reduction in gout flares needs to be replicated using data from ULT in other classes, namely uricosurics (e.g., probenecid, benzbromarone, lesinurad)
or recombinant uricases (e.g., pegloticase) in order for the Bucher criteria to be fulfilled and SU to be validated as a surrogate for gout flares.

8. What More Do We Need

Standardized reporting of outcomes in clinical trials of urate-lowering therapies in gout is required. Although OMERACT has endorsed outcome domains, including serum urate, gout flares, and tophus burden [19], the lack of standardized ways of reporting these measures in clinical trials, as outlined in Section 5, impedes quantitative evidence synthesis. Further work needs to be undertaken to develop agreed reporting standards for these important outcomes for gout clinical trials. In addition, reporting of clinical trials using surrogate outcomes needs to be improved, including specific discussion about the validity of the surrogate. Given that there is a move to develop specific extensions to the SPIRIT and CONSORT checklists, known as SPIRIT-SURROGATE and CONSORT-SURROGATE, to be used to report randomized trial protocols and results that use a surrogate primary endpoint [25], the gout community need to work together to provide the evidence that SU is a valid surrogate for gout flares and other patient-important outcomes used in clinical trials of urate-lowering therapies.

Funding: This research received no specific external funding. Section for Biostatistics and Evidence-Based Research, the Parker Institute is supported by a core grant from the Oak Foundation (OCAY-18-774-OFIL).

Conflicts of Interest: Author 1 has received grant funding from the Health Research Council of New Zealand, Royalties from Up-To-Date and consulting fees from Pharmac. The other authors declare no conflict of interest.

References


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.