PERSPECTIVES

When underlying biology threatens the randomization principle — initial gout flares of urate-lowering therapy

Hyon K. Choi[ID], Yuqing Zhang and Nicola Dalbeth[ID]

Abstract | Flare is the dominant feature of gout and occurs because of inflammatory response to monosodium urate crystals; prevention of gout flares should be the major goal of gout care. However, a paradoxical increase in the risk of flare following initiation of urate-lowering therapy presents considerable challenges for proving the expected long-term benefits of flare prevention in clinical trials. Nevertheless, excluding from enumeration flares that occur in the initial post-randomization period (which can last several months to 1 year) can threaten the core benefits of randomization: the characteristics of the remaining participants can differ from those who were randomized, introducing potential bias from confounding (both measured and unmeasured); participants who drop out or die are excluded from the analysis, introducing potential selection bias; and, finally, ignoring initial flares underestimates participants' experience during the trial. This Perspective discusses these issues and recommends measures that will allow for high-level evidence that preserves the randomization principle, to satisfy methodological scrutiny and generate robust evidence-based guidelines for gout care.

The gout flare is the dominant presentation of gout and occurs because of an inflammatory response to monosodium urate (MSU) crystals. The intense pain and impact of the gout flare mean that it is central to the patient experience of gout, and prevention of gout flares should be the major goal of effective gout management. However, the risk of flare paradoxically increases in the period after initiation of urate-lowering therapy (ULT), presenting considerable challenges related to proving the expected flare-prevention benefits of ULT over the long term in randomized trials. Nevertheless, excluding from enumeration flares that occur in the initial post-randomization period of randomized controlled trials (RCTs), which has been done in all RCTs to date, can threaten the core benefits of the randomization principle. For instance, the characteristics of participants remaining in the RCT after this initial period (lasting several months to 1 year in RCTs) might be different from the characteristics of those who

were randomized, introducing potential bias from measured and unmeasured confounding; moreover, participants who drop out or die during the initial period cannot be included in the analysis, introducing potential selection bias; finally, ignoring initial flares underestimates the burden of gout flares experienced by participants over the entire trial.

In this Perspective, we discuss several measures to accommodate this characteristic biology of paradoxical gout flares while preserving the benefits of randomization, including careful planning for entire-period analyses (as opposed to analyses of a specified post-randomization period), effective flare prophylaxis, sufficient trial duration, maximum efforts and mechanisms for participant retention, use of adherence-adjusted per-protocol analysis (in addition to intention-to-treat (ITT) analysis) and collection of high-quality longitudinal data to predict non-adherence. Implementation of these measures in gout RCTs will lead to high-level evidence of ULT

effects for flare prevention to generate robust evidence-based guidelines for gout care.

Flares are central to gout

The gout flare is the most common and dominant presentation of gout, and occurs because of the activation of the innate immune system in response to MSU crystals[1,2]. The gout flare is experienced as the rapid onset of acute joint inflammation, with severe pain and associated tenderness, swelling, warmth and erythema. The patient experience of the gout flare is multidimensional, as it affects activities of daily living (including difficulty with walking, self-care, driving and sleeping); social and family life (by restricting social participation, employment, independence and intimacy); and psychological health (contributing to irritability, anxiety, fear, depression, isolation and financial worry)[3]. The intensity of the pain and the impact of the gout flare mean that it is central to a patient's experience of gout, and prevention of gout flares should be the major goal for effective gout management. Nevertheless, high-quality data from trials with gout flares as a primary end point remain scarce, which has contributed to conflicting guidelines on gout care for primary care (American College of Physicians)[4] and for rheumatology, as reviewed elsewhere[5–8]. Rheumatology guidelines emphasize a treat-to-target serum urate approach (for example, serum urate concentration <6 mg/dl, a urate sub-saturation point)[9,10]; however, citing the absence of evidence, serum urate is not even measured during ULT in the vast majority of patients with gout in primary care practice, where >90% of gout care occurs[11].

Initiating ULT triggers flares

Although long-term ULT leads to the prevention of gout flares (through dissolution of deposited MSU crystals), the frequency of gout flares increases at the start of ULT. This common, paradoxical pattern of initial worsening, which can last for months (FIG. 1), followed by improvement of the same disease end point, is unique in modern rheumatology therapeutics, although it is often underappreciated and poorly explained to patients, contributing to premature discontinuation of ULT[12].
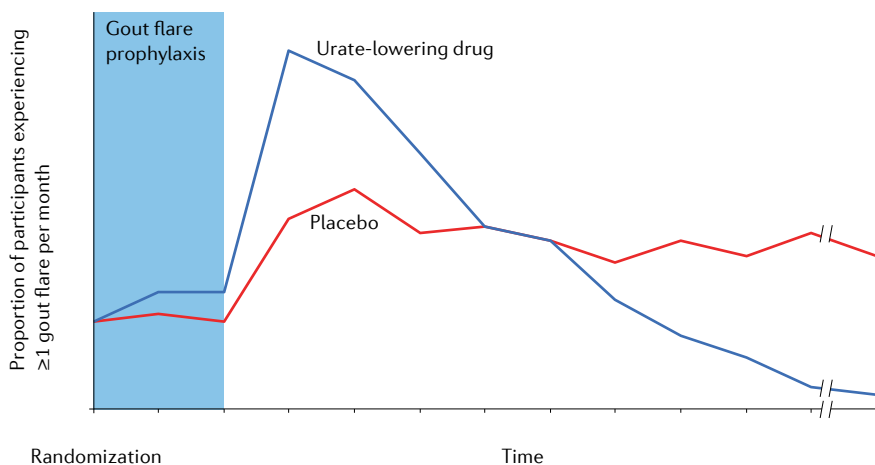
Fig. 1 | **Gout flare trends after initiation of a potent urate-lowering agent in a hypothetical placebo-controlled randomized controlled trial.** The risk of flares in the urate-lowering drug group increases after the initial anti-inflammatory prophylaxis phase of the trial (for example, 3 months) dissipates. This paradoxical worsening is followed by a substantially lower risk of flares over time. By contrast, the placebo group is expected to have a similar (or higher) level of flares over time, once the initial anti-inflammatory prophylaxis effect discontinues.

This phenomenon was recognized from the initial reports of ULT, with Yue and Gutman reporting in their early descriptions of allopurinol from the 1960s "The most troublesome problem we encountered with allopurinol therapy was the precipitation of acute gouty arthritis. The incidence of acute attacks provoked by allopurinol must be considered excessive"[13]. Gout flares occur in up to three-quarters of patients in the first

6 months of allopurinol treatment without anti-inflammatory prophylaxis[14].

Some investigators have termed gout flares occurring soon after initiation of ULT 'mobilization flares', reflecting that these flares are thought to occur as MSU crystals are shed or mobilized from intra-articular deposits when the serum urate level falls, leading to interactions between crystals and resident synoviocytes and initiation of the acute inflammatory response[15,16].

Increased frequency of gout flare has been reported with all currently approved urate-lowering drugs, and occurs more often in the setting of more rapid and intensive reductions in serum urate[17,18]. For this reason, gradual dose escalation of ULT and anti-inflammatory prophylaxis is recommended for the first 3–6 months of ULT[19]. The most common strategy for anti-inflammatory prophylaxis is low-dose daily colchicine, which reduces the frequency and severity of gout flares, and the likelihood of recurrent flares in those starting allopurinol[14]. In an RCT of patients with gout starting allopurinol, colchicine prophylaxis reduced the number of flares over 6 months (0.5 in the colchicine

Table 1 | **Design and gout flare reporting in pivotal clinical trials of urate-lowering therapy for gout**

Trials	Trial summary	Trial duration (dropout rate)	Flares as outcome measure	Anti-inflammatory prophylaxis	Gout flare outcome reporting and time periods	Ref.
FACT	RCT comparing fixed-dose febuxostat and fixed-dose allopurinol	52 weeks (35%)	Secondary end point	Low-dose colchicine or naproxen for 8 weeks	The proportion of participants requiring treatment for acute gout flares from weeks 9 to 52 (specified) The proportion of participants requiring treatment for acute gout flares from weeks 49 to 52 also reported	24
C0405 and C0406	Placebo-controlled RCT of pegloticase	6 months (30%)	Secondary end point	Low-dose colchicine or NSAIDs for 6 months (entire trial period)	The proportion of participants with gout flare (gout flare incidence) during months 1–3 and 4–6 of the trial The number of flares per participant during months 1–3 and 4–6 of the trial	26
CLEAR-1[a]	Placebo-controlled RCT of lesinurad (in combination with allopurinol)	12 months (25%)	Secondary end point	Low-dose colchicine or NSAID for 5 months	Mean rate of gout flares requiring treatment from the end of month 6 to the end of month 12	33
Doherty et al. 2018	RCT comparing nurse-led gout care with usual care	2 years (9% vs 21%)[b]	Secondary end point	Timing not standardized, low-dose colchicine used in 3/255 (1%) of participants in the nurse-led group	Frequency of gout flares during years 1 and 2	25
CSP594	Non-inferiority RCT comparing febuxostat and allopurinol using a treat to serum urate target approach	72 weeks (20%)	Primary end point	Low dose colchicine, NSAIDs, or glucocorticoids for 48 weeks	Primary end point: proportion of participants experiencing one or more flares during weeks 49 to 72 Secondary end point: rate of gout flares (events/person-years) during the entire period as well as each phase of the trial: titration (weeks 0 to 24), maintenance (weeks 25 to 48) and observation (weeks 49 to 72)	27

Trials in this table represent the major phase III trials for urate-lowering therapy approved by the FDA since 2009 or large (>500 participants) investigator-initiated strategy trials. RCT, randomized controlled trial. [a]US-based trial; same design used for CLEAR-2 multinational trial[41]. [b]9% in nurse-led gout care and 21% in usual care.

Table 2 | **Flare outcomes during pivotal clinical trials of urate-lowering therapy for gout**

Trials	Flare outcomes in the early trial period	Flare outcomes in the late trial period	Flare outcomes for the entire trial period	Ref.
FACT (Becker et al. 2005)	Day 1 to week 8: more participants with gout flare in the febuxostat 120 mg group than in the allopurinol group (36% vs 21%; $P < 0.001$)	Weeks 9–52: no difference in the proportion of participants with gout flare in the febuxostat 120 mg group compared with the allopurinol group (70% vs 64%) Weeks 49–52: no difference in the proportion of participants with gout flare in the febuxostat 120 mg group compared with the allopurinol group (6% vs 11%)	Not reported	[24]
C0405 and C0406 (Sundy et al. 2011)	Months 1–3: higher gout flare incidence in the bi-weekly pegloticase group than in the placebo group (75% vs 53%; $P = 0.02$) Months 1–3: more gout flares in the bi-weekly pegloticase group than in the placebo group (mean 2.3 vs 1.2; $P = 0.001$)	Months 4–6: lower gout flare incidence in the bi-weekly pegloticase group than in the placebo group (41% vs 67%; $P = 0.007$) Months 4–6: fewer gout flares in the bi-weekly pegloticase group than in the placebo (mean 0.8 vs 1.3; $P = 0.06$)	Not reported	[26]
CLEAR-1 (Saag et al. 2017)	Not reported	Months 7–12: no difference in the lesinurad 400 mg group and the placebo group (mean 0.51 vs 0.58; $P = 0.61$)	Not reported	[33]
Doherty et al. 2018	Year 1: more participants in the nurse-led group than in the usual care group experienced ≥2 gout flares (54% vs 40%, risk ratio 1.36 (95% CI 1.05–1.77))	Year 2: fewer participants in the nurse-led group than in the usual care group experienced ≥2 gout flares (8% vs 24%, risk ratio 0.33 (95% CI 0.19–0.57))	Not reported	[25]
CSP594 (O'Dell et al. 2022)	Weeks 0–48: proportion of participants experiencing one or more gout flares not reported Secondary end point: weeks 0–24: no difference in gout flare rate in the allopurinol group and the febuxostat group (2.09 vs 2.25 flares per person years; rate ratio 0.93 (95% CI 0.81–1.06)) Secondary end point: weeks 25–48: no difference in gout flare rate in the allopurinol group and the febuxostat group (1.60 vs 1.59 flares per person years; rate ratio 1.00 (95% CI 0.85–1.18))	Primary end point: weeks 49–72: fewer participants with gout flares in the allopurinol group than in the febuxostat group (36.5% vs 43.5%; risk ratio −7 (95% CI −∞ to −1.2); $P < 0.001$ for non-inferiority of allopurinol) Secondary end point: weeks 49–72: fewer gout flares in the allopurinol group than in the febuxostat group (1.48 vs 2.02 flares per person years; rate ratio 0.73 (95% CI 0.63–0.86))	Proportion of participants experiencing one or more gout flares over the entire study period not reported Gout flare rate over the entire study period reported in the manuscript: fewer gout flares in the allopurinol group than in the febuxostat group (1.73 vs 1.97 flares per person-years, rate ratio 0.88 (95% CI 0.81–0.96))	[27]

Trials in this table represent the major phase III trials for urate-lowering therapy approved by the FDA since 2009 or large investigator-initiated strategy trials (>500 participants). Where more than one dose was tested, data for the highest dose are presented.

group and 2.9 in the placebo group; $P = 0.008$[14]. Low-dose NSAIDs can also be used as anti-inflammatory prophylaxis[20]. Anti-IL-1 therapies, such as canakinumab and rilonacept, have also been shown in clinical trials to reduce gout flares at the time of initiating ULT[21,22], but are not approved for this indication.

How are flares assessed in trials?
The Outcome Measures in Rheumatology (OMERACT) group has recognized the central importance of the gout flare in its core set of outcome domains for long-term studies of gout[23]. However, pivotal phase III trials of ULT in the modern era of drug approval and major strategy trials have segmented out flare analyses to focus on flare reporting later in the course of treatment, when the risk of flares has subsided (summarized in TABLE 1). Some of these trials have reported an increased risk of flares with the investigational product in the early stages of the trial[24–26], but the approach of reporting the flare experience

over the entire period of the trial has not been adopted (TABLE 2). The comparative efficacy trial of febuxostat and allopurinol published in early 2022 is the first ULT trial to report gout flares as the primary end point[27]. However, this primary end point (the proportion of participants experiencing one or more flares) only covered the third phase of the study (weeks 49 to 72), after urate-lowering had been established, but not for the entirety of the trial or for the first phase (weeks 0 to 24) or second phase (weeks 25 to 48) individually. Nevertheless, flare rate (as opposed to risk proportion), one of the trial's pre-specified secondary end points, was reported over the entire period as well as for each of the three phases[27].

These analytic approaches[27] are in contrast to those used in trials of other interventions with time-dependent (non-proportional) effects (TABLE 3). For example, similar time-dependent trade-offs arise with the use of initially intrusive interventions, such as

transplantation, surgeries or other invasive procedures, which can have immediate adverse effects — even mortality, initially — but subsequently lead to benefits among those who survive. For example, the 2014 ASTIS RCT of autologous haematopoietic stem cell transplantation for systemic sclerosis produced crossing survival curves for both death and organ failure, as expected, but the primary end point (event-free survival) was reported over the entire period (median follow-up 5.8 years) as well as at year 1, year 2 and year 4, starting from randomization[28] (TABLE 3). Other trials have similarly reported all adverse effects and benefits that occur during the entire period since randomization[29,30].

To accommodate the expected lag in biological effect of the COVID-19 vaccination, two trials[31,32] ascertained the primary end point (COVID-19 infection) from a pre-specified time after randomization (7 or 14 days after the second dose of vaccine) (TABLE 3) in a per-protocol analysis, which excluded ~3%

Table 3 | Clinical trials with post-randomization landmark time or time-dependent relative risks

Trial	Trial summary	Trial duration	Primary outcome measure	Outcome reporting	Ref.
ASTIS (van Laar et al. 2014)	RCT of ASCT vs cyclophosphamide for diffuse systemic sclerosis	5.8 years	Event-free survival (death or persistent major organ failure)	Events during the entire follow-up period and also at year 1, year 2 and year 4 Outcomes initially worse with ASCT and then better; survival curves crossed at year 2	[28]
SCOT (Sullivan et al. 2018)	RCT of myeloablative ASCT vs cyclophosphamide for severe scleroderma	54 months (primary) and up to 72 months	Global rank composite score (including death and major events)	ITT and 'per-protocol'[a]: event-free survival of the entire follow-up period. ASCT had longer-term benefits in events and mortality, but treatment-related mortality was 6% in the ASCT group vs 0% in the cyclophosphamide group	[29]
MEDIC (Skou et al. 2015)	RCT of TKR vs non-surgical treatment	12 months	Change in OA outcome scale scores (0–100) over 12 months	OA outcome scale score over 12 months improved with TKR Adverse effects over 12 months were worse in the TKR group	[30]
Polack et al. 2020	RCT of BNT162b2 mRNA vaccine vs placebo	Varied owing to ethical concern, although designed to be up to 2 years	COVID-19 onset ≥7 days after second dose of vaccine	Risk of COVID-19 after dose 1, between doses 1 and 2, 7 days after dose 2, and ≥7 days after dose 2 (primary)	[31]
Baden et al. 2021	RCT of mRNA-1273 vaccine vs placebo	Varied; median follow-up duration of ≥2 months, per FDA	COVID-19 onset ≥14 days after the second dose of vaccine	Risk of COVID-19 any time after randomization, between randomization and 14 days after dose 1, 14 days after dose 1 to dose 2, dose 2 to 14 days after dose 2, and ≥14 days after dose 2 (primary)	[32]

ASCT, autologous stem cell transplantation; ITT, intention-to-treat; OA, osteoarthritis; RCT, randomized controlled trial; TKR, total knee replacement. [a]Defined as participants who received a transplant or completed nine or more doses of cyclophosphamide.

to 7% of the ITT population as randomized. Both trials also conducted ITT analyses for the same end points over the entire period of follow-up, starting from the point of randomization (TABLE 3), the results of which were consistent with the primary analyses. Thus, these vaccine trials adopted similar landmark analysis strategies to the gout trials discussed above (TABLES 1,2), although the duration of the unaccounted period was shorter in the vaccine trials (4–6 weeks), resulting in lower dropout rates; additionally, there was no initial paradoxical worsening in the intervention group, and the same primary end points were reported over the entire period after randomization (using ITT analysis).

Landmark analysis can bias results

The randomization in clinical trials (RCTs), when done properly and with a sufficiently large sample size, guarantees that the potential confounders, known or unknown, are evenly distributed between the comparison groups, providing a powerful advantage over observational studies. This advantage is valid at the time of randomization (the index date) and can be sustained when all events are counted after randomization during the entire trial period, without biased follow-up (for example, from participants dropping out or switching treatment). However, all gout trials to date, including the pivotal trials in TABLE 1, started counting flare events (as a primary end point in one trial[27] and a secondary end point in the other trials[24–26,33]) at a specific time that was substantially after the time of randomization (that is, a post-randomization landmark time), and they also had notable dropout rates (TABLE 1). Such post-randomization analysis (also called landmark analysis)[34] resets the trial 'clock', by moving the index date from the time of randomization to the landmark time. As a result, the characteristics of participants who survive and are retained to the landmark date could be different from the characteristics of those who were randomly assigned to a treatment group on the index date[34,35], introducing potential confounding bias (including both measured and unmeasured confounders). Furthermore, participants who are not retained to the landmark date (for example, owing to dropout or death) cannot be included in the analysis, introducing potential selection bias. Finally, this approach ignores the outcome (that is, gout flare) that occurs between the index date and the landmark date, resulting in underestimation of the risk of flares that participants experience over the entire study period[34,35], as stated above.

ITT can underestimate effect

ITT analysis of events during the entire study period guarantees a valid estimate of the effect of the treatment on the outcome, provided that there is no treatment misclassification and no selection bias. However, in the context of notable non-adherence to treatment assignment, including loss-to-follow-up, particularly in long-term RCTs (which is often the case with gout trials) (TABLE 1), ITT analysis would, in general, underestimate the effect of treatment. As such, if the outcome of interest concerns undesirable events, such as gout flare or safety (for example, toxicity), the results could be incorrectly interpreted as lack of evidence of harm[36]. A better approach to account for non-adherence in this context (including loss to follow-up) is an adherence-adjusted per-protocol analysis. By predicting non-adherence using appropriate statistical methods[37], this approach enables the investigators to assess the effect that would have been observed if all participants (as randomized) had received their assigned treatment during the study period[37] (BOX 1). This adherence-adjusted per-protocol analysis should not be confused with conventional on-treatment or as-treated analysis, which jeopardizes the central purpose of randomization, unlike ITT or adherence-adjusted per-protocol analysis (BOX 1). Nevertheless, adherence-adjusted per-protocol analysis relies on available prognostic factors to predict the risk of non-adherence, necessitating pre-planned collection of high-quality longitudinal data including health care utilization, comorbidities and medication use[38] (TABLE 4).

Solutions and associated issues

To accommodate the characteristic biology of flares in gout while retaining the advantages of randomization, future RCTs of ULT with gout flares as end points would be well served by several considerations (TABLE 4). To take full advantage of the RCT design and avoid potential biases that could interfere with identifying causal relationships, it would be desirable to include flares over the entire trial period (starting from randomization) as the primary outcome. As a minimum, investigators should include in their report entire-period data for the occurrence of flares. Analysis of a pre-specified specific period after initial flare early in the course of treatment (that is, landmark analysis)[34] should consider including measures to appropriately address expected non-adherence and dropout by the landmark time and to adjust for potential confounders, as the intervention and comparison groups are no longer the same as the groups that were randomized. These issues can threaten the validity of an RCT, particularly when there are notable dropouts or switching of treatments. In other words, although the data are generated from an RCT, the study ends up having the vulnerabilities of observational studies and the severe loss of the advantages of an RCT, which can provide misleading data under the guise and perceived weight of the RCT label. To overcome these issues, we recommend analyses of pre-specified periods that start counting flares from the time of randomization, although the follow-up time can be stopped at different time points of interest to demonstrate the lagged effect of ULT after the expected initial worsening (TABLE 4).

One difficulty to consider is that the effect of ULT on flares analysed over the entire study period could be diluted by flares that occur after ULT initiation. Also complicating outcome analysis is that non-adherence could arise as a consequence of the adverse effect of a flare following ULT initiation. These difficulties can be mitigated by the use of highly effective anti-inflammatory flare prophylaxis and gradual dose escalation of ULT[17]. The duration of prophylaxis has varied substantially in gout trials to date, ranging from nearly zero prophylaxis[25] to prophylaxis for up to 11 months[27] (TABLE 1). We recommend prophylaxis for a duration of 3–6 months as a minimum, as recommended in the ACR guideline for the management of gout[19].

Trial duration is another important consideration, as the clinical benefits of ULT for flare outcomes are usually observed after more than 1 year of therapy[39]. The trial must therefore be long enough to overcome the initial worsening of flares, particularly with potent ULT, although longer trials tend to suffer from higher rates of non-adherence and dropout than shorter ones. Nevertheless, as with all RCTs and even more so in gout trials (which tend to have notable dropouts), investigators should maximize efforts to avoid dropouts by the use of intense retention strategies, including repeated engagement of patients by research staff such as nurses, wherever feasible and appropriate. For example, in a UK trial the dropout rate in a group receiving nurse-led intervention was less than 10% over 2 years[25].

In terms of analytical approach, we recommend a priori specification of the statistical analysis plan for adherence-adjusted per-protocol analysis over the entire trial period starting at randomization, in addition to ITT analysis. The adherence-adjusted per-protocol analysis will be directly relevant in accounting for expected dropouts and non-adherence during the trial while accommodating the initial flare phase and the pre-specified partial period analyses. For example, in a 2022 trial that used gout flare as the primary end point[27], the dropout rate by the time the investigators enumerated the primary analysis end point (end of the third phase) was 20%. Furthermore, for the adherence-adjusted per-protocol analysis to effectively account for dropouts, high-quality longitudinal data should be collected by planning ahead (or be available through linked electronic medical records) (TABLE 4). Finally, we recommend using rates (number of events per person-time) as the primary end point, as opposed to the proportion of participants experiencing one or more events (or risk estimate), as the latter would be difficult to implement, particularly given that flares will be frequent during the initial months of ULT, overwhelming the first-event analysis. To that end, Poisson distribution would reflect gout flare events well by accommodating the event counts in rates (TABLE 4).

Other end points in ULT trials

Although the central importance of the gout flare for long-term gout trials is recognized[23], the current practice of using serum urate concentration as the primary end point in pivotal trials for the approval of new ULT drugs for gout care[40] is likely to continue, as long as the effect of the treat-to-serum-urate-target approach is firmly established with the determination of clinically meaningful, quantitative improvement in serum urate levels over the entire duration of treatment. To that end, it would be desirable to quantify the value of such an approach for clinical end points in comparison with alternative strategies in a high-quality RCT. Once this effect and its magnitude are clearly established, future studies could rely on serum urate response as a powerful surrogate for gout flare risk, similar to the way in which serum cholesterol levels came to be used as a surrogate end point in the development and approval of cardiovascular drugs (after several large trials confirmed its strength as a surrogate for 'hard' end points) . This approach might ultimately reduce the cost of future development programmes for urate-lowering drugs. Furthermore, serum urate levels start improving, usually within days of starting ULT, without initial worsening (in contrast to flare end points)[40], avoiding many of the issues we discuss above. In terms of other core clinical end points of gout in the evaluation of ULT, tophus burden reduction

---

Box 1 | **Adherence-adjusted per-protocol analysis and conventional on-treatment or as-treated analysis**

- Adherence-adjusted per-protocol analysis uses inverse probability weights to account for non-adherence to treatment. The denominator of the inverse probability weight is the probability that a participant adhered to their assigned treatment, obtained from logistic regression. The predictors in the logistic regression model consist of the baseline covariates and post-randomization time-varying covariates[34]. By accounting for non-adherence[34], this method allows for the assessment of the effect that would have been observed if all participants (as randomized) had received their assigned treatment during the study period[34].

- By contrast, traditional naive 'per-protocol analysis' (or complete set) only includes participants/ person-time that adhere to assigned treatment by excluding those who deviate from the protocol. If non-adherence is not random, a treatment effect is susceptible to confounding or selection bias (as discussed in the main text). Similarly, the conventional as-treated or on-treatment approach defines a participant's treatment status according to the treatment they received regardless of their randomized allocation. This approach is vulnerable to confounding bias because it ignores the random assignment of treatment.

Table 4 | **Recommendations for future ULT trials with gout flare end points**

Recommendations	Purpose
Plan on reporting entire-period flare results (primary period of interest) as well as pre-specified partial periods (starting from the time of randomization) (secondary period of interest)	To accommodate gout-specific biology while retaining the advantage of the RCT design
Implement effective flare prophylaxis during the initial period of ULT, when risk of flare is paradoxically increased	To minimize dilution of the effect of the intervention in analysis of the entire trial period
Conduct long-term trials to overcome the initial worsening of flares in the intervention group	To avoid false-negative results while quantifying the clinical benefits and risks of ULT, given that ULT is a long-term care medication
Design and carry out the trial to minimize dropouts	To maximize the validity of the RCT in both entire-period and partial-period analyses
Specify a priori the statistical analysis plan for adherence-adjusted per-protocol analysis[a] for the entire period, as well as pre-specified partial-period analyses, in addition to ITT analysis	To account for dropouts and treatment adherence
Collect high-quality longitudinal data, including health care utilization, comorbidities and medication use	To effectively predict and account for treatment adherence
Consider using flare rate as the primary end point, as opposed to flare risk (proportional), and employ Poisson regression models	To best accommodate for recurrent events of gout, the time-dependent (non-proportional) risk and the paradoxical increase in flare risk after ULT initiation
Collect data on the use of prophylaxis and medication for acute gout care	To serve as a secondary end point and as a key variable to account for non-adherence and censoring events for primary flare end point

ITT, intention-to-treat; RCT, randomized controlled trial; ULT, urate-lowering therapy. [a]See BOX 1 for further explanation of adherence-adjusted per-protocol analysis.

does not involve initial worsening in the same way as flare end points, although patient-reported-outcome and quality-of-life measures would be partially affected by the initial worsening of flares[40]. As such, these clinical end points would also be better served by long-term trials, such as those of 2 years' duration.

Conclusions

The period of increased flare risk occurring after ULT initiation, which lasts for months, presents considerable challenges in proving the expected flare-prevention benefits of ULT in the long term. Excluding flare outcomes that occur in the initial post-randomization period can threaten the randomization property that allows for causal conclusions to be drawn from RCTs. To accommodate this rare biological phenomenon with the randomization property intact, we recommend careful planning for entire-period analyses, adequate trial duration, effective flare prophylaxis, maximum retainment efforts/mechanisms, adherence-adjusted per-protocol analysis (in addition to ITT), and high-quality longitudinal data collection for adherence prediction. These approaches will allow for high-level evidence of the

ULT effect that preserves the randomization principle, which will satisfy methodological scrutiny and generate solid evidence-based guidelines for optimal gout care.

Hyon K. Choi [ID][1,2][✉], *Yuqing Zhang*[1,2] *and Nicola Dalbeth* [ID][3]

[1]*Division of Rheumatology, Allergy, and Immunology, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA.*

[2]*The Mongan Institute, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA.*

[3]*Department of Medicine, University of Auckland, Auckland, New Zealand.*

✉*e-mail: hchoi@mgh.harvard.edu*

1. Faires, J. S. & McCarty, D. J. Acute arthritis in man and dog after intrasynovial injection of sodium urate crystals. *Lancet* **280**, 682–685 (1962).
2. Martinon, F., Petrilli, V., Mayor, A., Tardivel, A. & Tschopp, J. Gout-associated uric acid crystals activate the NALP3 inflammasome. *Nature* **440**, 237–241 (2006).
3. Stewart, S. et al. The experience of a gout flare: a meta-synthesis of qualitative studies. *Semin. Arthritis Rheum.* **50**, 805–811 (2020).
4. Qaseem, A., Harris, R. P. & Forciea, M. A. Management of acute and recurrent gout: a clinical practice guideline from the American College of Physicians. *Ann. Intern. Med.* **166**, 58–68 (2017).
5. Dalbeth, N. et al. Discordant American College of Physicians and international rheumatology guidelines for gout management: consensus statement of the Gout, Hyperuricemia and Crystal-Associated Disease Network (G-CAN). *Nat. Rev. Rheumatol.* **13**, 561–568 (2017).
6. McLean, R. M. The long and winding road to clinical guidelines on the diagnosis and management of gout. *Ann. Intern. Med.* **166**, 73–74 (2017).
7. Neogi, T. & Mikuls, T. R. To treat or not to treat (to target) in gout. *Ann. Intern. Med.* **166**, 71–72 (2017).
8. FitzGerald, J. D., Neogi, T. & Choi, H. K. Editorial: do not let gout apathy lead to gouty arthropathy. *Arthritis Rheumatol.* **69**, 479–482 (2017).
9. Khanna, D. et al. 2012 American College of Rheumatology guidelines for management of gout. Part 1: systematic nonpharmacologic and pharmacologic therapeutic approaches to hyperuricemia. *Arthritis Care Res.* **64**, 1431–1446 (2012).
10. Richette, P. et al. 2016 updated EULAR evidence-based recommendations for the management of gout. *Ann. Rheum. Dis.* **76**, 29–42 (2017).
11. Krishnan, E., Lienesch, D. & Kwoh, C. K. Gout in ambulatory care settings in the United States. *J. Rheumatol.* **35**, 498–501 (2008).
12. Latif, Z. P., Nakafero, G., Jenkins, W., Doherty, M. & Abhishek, A. Implication of nurse intervention on engagement with urate-lowering drugs: a qualitative study of participants in a RCT of nurse led care. *Joint Bone Spine* **86**, 357–362 (2019).
13. Yue, T. F. & Gutman, A. B. Effect of allopurinol (4-hydroxypyrazolo-(3,4-d)pyrimidine) on serum and urinary uric acid in primary and secondary gout. *Am. J. Med.* **37**, 885–898 (1964).
14. Borstad, G. C. et al. Colchicine for prophylaxis of acute flares when initiating allopurinol for chronic gouty arthritis. *J. Rheumatol.* **31**, 2429–2432 (2004).
15. Schumacher, H. R. Jr & Chen, L. X. The practical management of gout. *Clevel. Clin. J. Med.* **75**, S22–S25 (2008).
16. Perez-Ruiz, F. Treating to target: a strategy to cure gout. *Rheumatology* **48**, ii9–ii14 (2009).
17. Yamanaka, H. et al. Stepwise dose increase of febuxostat is comparable with colchicine prophylaxis for the prevention of gout flares during the initial phase of urate-lowering therapy: results from FORTUNE-1, a prospective, multicentre randomised study. *Ann. Rheum. Dis.* **77**, 270–276 (2018).
18. Becker, M. A., MacDonald, P. A., Hunt, B. J., Lademacher, C. & Joseph-Ridge, N. Determinants of the clinical outcomes of gout during the first year of urate-lowering therapy. *Nucleosides Nucleotides Nucleic Acids* **27**, 585–591 (2008).
19. FitzGerald, J. D. et al. 2020 American College of Rheumatology guideline for the management of gout. *Arthritis Rheumatol.* **72**, 879–895 (2020).
20. Wortmann, R. L., Macdonald, P. A., Hunt, B. & Jackson, R. L. Effect of prophylaxis on gout flares after the initiation of urate-lowering therapy: analysis of data from three phase III trials. *Clin. Ther.* **32**, 2386–2397 (2010).
21. Schlesinger, N. et al. Canakinumab reduces the risk of acute gouty arthritis flares during initiation of allopurinol treatment: results of a double-blind, randomised study. *Ann. Rheum. Dis.* **70**, 1264–1271 (2011).
22. Mitha, E. et al. Rilonacept for gout flare prevention during initiation of uric acid-lowering therapy: results from the PRESURGE-2 international, phase 3, randomized, placebo-controlled trial. *Rheumatology* **52**, 1285–1292 (2013).
23. Schumacher, H. R. et al. Outcome domains for studies of acute and chronic gout. *J. Rheumatol.* **36**, 2342–2345 (2009).
24. Becker, M. A. et al. Febuxostat compared with allopurinol in patients with hyperuricemia and gout. *N. Engl. J. Med.* **353**, 2450–2461 (2005).
25. Doherty, M. et al. Efficacy and cost-effectiveness of nurse-led care involving education and engagement of patients and a treat-to-target urate-lowering strategy versus usual care for gout: a randomised controlled trial. *Lancet* **392**, 1403–1412 (2018).
26. Sundy, J. S. et al. Efficacy and tolerability of pegloticase for the treatment of chronic gout in patients refractory to conventional treatment: two randomized controlled trials. *JAMA* **306**, 711–720 (2011).
27. O'Dell, J. R. et al. Comparative effectiveness of allopurinol and febuxostat in gout management. *NEJM Evid.* **1**, EVIDoa2100028 (2022).
28. van Laar, J. M. et al. Autologous hematopoietic stem cell transplantation vs intravenous pulse cyclophosphamide in diffuse cutaneous systemic sclerosis: a randomized clinical trial. *JAMA* **311**, 2490–2498 (2014).
29. Sullivan, K. M. et al. Myeloablative autologous stem-cell transplantation for severe Scleroderma. *N. Engl. J. Med.* **378**, 35–47 (2018).

30. Skou, S. T. et al. A randomized, controlled trial of total knee replacement. *N. Engl. J. Med.* **373**, 1597–1606 (2015).
31. Polack, F. P. et al. Safety and efficacy of the BNT162b2 mRNA covid-19 vaccine. *N. Engl. J. Med.* **383**, 2603–2615 (2020).
32. Baden, L. R. et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N. Engl. J. Med.* **384**, 403–416 (2021).
33. Saag, K. G. et al. Lesinurad combined with allopurinol: a randomized, double-blind, placebo-controlled study in gout patients with an inadequate response to standard-of-care allopurinol (a US-based study). *Arthritis Rheumatol.* **69**, 203–212 (2017).
34. Dafni, U. Landmark analysis at the 25-year landmark point. *Circ. Cardiovasc. Qual. Outcomes* **4**, 363–371 (2011).
35. Garcia-Albeniz, X., Maurel, J. & Hernan, M. A. Why post-progression survival and post-relapse survival are not appropriate measures of efficacy in cancer randomized clinical trials. *Int. J. Cancer* **136**, 2444–2447 (2015).
36. Hernan, M. A., Hernandez-Diaz, S. & Robins, J. M. Randomized trials analyzed as observational studies. *Ann. Intern. Med.* **159**, 560–562 (2013).
37. Smith, V. A., Coffman, C. J. & Hudgens, M. G. Interpreting the results of intention-to-treat, per-protocol, and as-treated analyses of clinical trials. *JAMA* **326**, 433–434 (2021).
38. Hernan, M. A. & Robins, J. M. Per-protocol analyses of pragmatic trials. *N. Engl. J. Med.* **377**, 1391–1398 (2017).
39. Stamp, L. et al. Serum urate as surrogate endpoint for flares in people with gout: a systematic review and meta-regression analysis. *Semin. Arthritis Rheum.* **48**, 293–301 (2018).
40. Morillon, M. B. et al. Serum urate as a proposed surrogate outcome measure in gout trials: from the OMERACT working group. *Semin. Arthritis Rheum.* **51**, 1378–1385 (2021).
41. Bardin, T. et al. Lesinurad in combination with allopurinol: a randomised, double-blind, placebo-controlled study in patients with gout with inadequate response to standard of care (the multinational CLEAR 2 study). *Ann. Rheum. Dis.* **76**, 811–820 (2017).