

Stewart Sarah (Orcid ID: 0000-0002-9318-5627)
 Gaffo Angelo L. (Orcid ID: 0000-0001-7365-7212)
 Stamp Lisa Katrina (Orcid ID: 0000-0003-0138-2912)
 Dalbeth Nicola (Orcid ID: 0000-0003-4632-4476)

Gout flare resolution, physical activity and sleep

The association of gout flare resolution with physical activity and sleep: a prospective remote-monitoring study

Running head: Gout flare resolution, physical activity and sleep

Authors: Sarah Stewart^{1,2}, PhD; Boston Girven¹; Mia Clarke¹, BHSc (Hon); Irene Zeng³, PhD; Angelo Gaffo⁴, MD, MsPH, FACP; Lisa K Stamp^{5,6}, MBChB, FRACP, PhD; N Lawrence Edwards⁷, MD, MACR, MACP; David Rice^{1,8}, PhD; Tania Ka'ai⁹, PhD; Karl Rudolph¹, Julie Collis¹⁰, PhD; Nicola Kayes¹¹, PhD; Nicola Dalbeth^{2,12}, MBChB, MD, FRACP

¹School of Allied Health, Faculty of Environmental Sciences, Auckland University of Technology, Auckland, New Zealand; ²Department of Medicine, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand; ³Clinical Sciences Management, Faculty of Health and Environmental Sciences, Auckland University of Technology, Auckland, New Zealand; ⁴The University of Alabama at Birmingham and the Birmingham VA Medical Center, Alabama, United States; ⁵Department of Medicine, University of Otago, Christchurch, Christchurch, New Zealand; ⁶Health New Zealand, Te Whatu Ora Waitaha, New Zealand; ⁷University of Florida, Florida, United States; ⁸Department of Anaesthesiology and Perioperative Medicine, Health New Zealand Te Whatu Ora Waitematā, Auckland, New Zealand; ⁹Te Ipukarea Research Institute, Auckland University of Technology, Auckland, New Zealand; ¹⁰Department of Allied Health, School of Health Sciences, Swinburne University, Melbourne, Australia; ¹¹Centre for Person Centred Research, Auckland University of Technology, Auckland, New Zealand; ¹²Health New Zealand, Te Whatu Ora Te Toka Tumai, New Zealand

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the [Version of Record](#). Please cite this article as doi: [10.1002/acr.80194](https://doi.org/10.1002/acr.80194)

This article is protected by copyright. All rights reserved.

Gout flare resolution, physical activity and sleep

Corresponding author: A/Prof Sarah Stewart, School of Allied Health, Faculty of Environmental Sciences, Auckland University of Technology, 90 Akoranga drive, Northcote, Auckland, New Zealand. Email: sarah.stewart@aut.ac.nz

Funding statement: This study was supported by an Arthritis NZ Matepona Aotearoa Summer Student Scholarship (SS94) and an Emerging Researcher First Grant from the Health Research Council of New Zealand (23/451/A).

Word count: 4190/4200

ABSTRACT

Objectives: This study aimed to determine whether gout flare resolution is associated with physical activity and sleep. These findings were intended to inform the future development of a definition for flare resolution.

Methods: Data from 81 participants with gout who experienced ≥ 1 flare during a 6-month prospective observational study were included. Daily activity (step count) and sleep (total sleep time and deep sleep percentage) were collected remotely using wrist-worn trackers. Daily flare occurrence was captured using smartphone diaries. Activity/sleep outcomes during each flare episode were compared to those during three resolution timeframes: T1: the first day following the flare; T2: the consecutive window immediately following the flare and equal in length to the flare; and T3: the consecutive window equal in length to the flare and immediately following a washout window (equal in length to the flare). Adjusted differences in medians between flare and resolution timeframes were estimated using linear quantile mixed models.

Results: Daily step count was lower during flares compared with all resolution timeframes (all $P \leq 0.033$). T3 provided the best overall model fit, suggesting that resolution is most clearly distinguished from the flare when using this timeframe. No significant differences between flare and resolution were observed for the sleep outcomes for any resolution timeframe.

Conclusions: Physical activity improves soon after flare cessation, with the clearest distinction between flare and resolution observed over a longer post-flare recovery period, whereas sleep metrics do not differentiate flare from resolution states. These findings support physical activity as a useful objective domain for gout flare resolution.

SIGNIFICANCE AND INNOVATIONS

- The first remote-monitoring study to evaluate whether objectively measured physical activity and sleep differ between gout flares and predefined post-flare resolution timeframes
- Daily step count improves soon after flare cessation, with the clearest distinction between flare and resolution observed over a longer post-flare recovery period, supporting physical activity as a useful objective domain for gout flare resolution
- Wearable-derived total sleep time and deep sleep percentage have limited ability to distinguish flare from resolution states

Gout flares are intensely painful inflammatory events driven by innate immune recognition of deposited monosodium urate (MSU) crystals within articular and peri-articular tissues.¹⁻³

They are characterised by rapid-onset pain, erythema, warmth, and swelling, causing significant impairment of physical function, daily activities, and psychological well-being.⁴⁻⁶

Sleep is also adversely affected by pain, including difficulty initiating and maintaining sleep.^{4,5} The unpredictability of flares can generate stress and make it difficult for people with gout to plan activities and maintain regular sleeping patterns.^{1,4,5}

Recent studies have sought to standardise the definition of a gout flare.⁶ A validated definition has been developed requiring three of four criteria: patient-reported joint swelling, joint warmth, pain at rest greater than 3 on a 0-10 numeric rating scale (NRS), and patient-reported flare status.⁷ However, measuring gout flares remains challenging, as existing methodological approaches typically rely on binary classifications (i.e., flare presence or absence) and do not capture fluctuations in pain or other outcomes, such as function or sleep, across the duration of the flare.⁸ While gout flares usually resolve over days to weeks,^{1,9,10} there is no validated definition of flare resolution. This limits comparisons between trials assessing short-term gout flare treatments, where flare resolution (rather than flare onset) may be the clinically relevant endpoint for people with gout. Importantly, the intercritical phase can retain features of low-grade inflammation despite symptom improvement,¹¹⁻¹³ further underscoring the need to characterise when people with gout truly return to baseline.

As part of a broader program of work to develop a definition for gout flare resolution, recent qualitative work has shown that resolution is experienced as a return to normality across pain and function, including the ability to sleep comfortably, with emotional and inflammatory features also contributing.¹⁴ Further evaluation of objective mobility and sleep as potential markers of resolution is therefore needed. Emerging approaches integrating smartphone-based remote tracking and wearable devices demonstrate promise in assessing flare-related

outcomes.^{15,16} Proof-of-concept work in gout demonstrates lower daily step counts during flares,¹⁶ suggesting that device-derived measurements may provide objective markers of flare dynamics and recovery with lower patient burden than traditional data collection methods.

This analysis aimed to determine whether changes in (i) physical activity and/or (ii) sleep were associated with gout flare resolution using objectively measured activity and sleep data collected during a prospective six-month observational study.

PATIENTS AND METHODS

Study design

This analysis used data collected from 90 people with gout who participated in a six-month prospective observational cohort study. The primary aim of that study was to examine physical activity before and during flares relative to matched flare-free periods and will be reported separately. In recognition of the disproportionately severe experience of gout for Māori and Pacific Peoples in Aotearoa New Zealand,¹⁷ and to align with culturally responsive research practice in this setting, the original study was developed with input from a Māori cultural advisor and academic (TK, affiliated with Ngāti Porou and Ngāi Tahu iwi [tribes]) and a Māori patient research partner (KR, affiliated with Tai Poutini and Ngā Puhi iwi). Their contributions to protocol development and data collection procedures supported the culturally safe and feasible delivery of remote activity and sleep monitoring.

Participants

Eligible participants were aged ≥ 20 years, had a physician diagnosis of gout and at least two self-reported gout flares in the past six months, had smartphone access, and were willing to wear a wristwatch continuously for six months. Participants with other forms of inflammatory arthritis, or musculoskeletal or neurological conditions (unrelated to gout)

affecting lower-limb function were excluded, as were those who were immobile, or had cognitive impairment. All participants provided written informed consent. Ethical approval for the study was obtained from the Auckland University of Technology (AUT) Ethics Committee (AUTECH ref: 24/72). For the present analysis, participants who did not experience any gout flares during follow-up (n=9, 10.0%) were excluded a priori. This resulted in a final sample of 81 participants whose data were used in the current analysis.

Data collection

Participants attended a baseline visit at either the AUT North Campus (Northcote, Auckland, New Zealand) or the AUT South Campus (Manukau, Auckland, New Zealand) between June 2024 and April 2025. Demographic and clinical data were obtained and participants were fitted with a Garmin vívoFit 4 wrist-worn tracker, designed for continuous 24/7 wear, with ~1+ year of battery life and step and sleep tracking. Participants' devices were configured to display only the time on the watch face to minimise behavioural reactivity to activity/sleep metrics. The GarminConnect app was installed and tested on participants' smartphones to support remote data capture over follow-up, and the MyCap app was installed as an electronic daily gout flare diary.

Participants were asked to wear the tracker continuously for six months and sync it to Garmin Connect at least monthly. Tracker data were automatically exported to Fitabase (research platform) for storage and subsequent analysis. Participants were prompted daily by the MyCap app to report whether they were experiencing a gout flare that day. Participants were advised to continue their usual pain-relief/anti-inflammatory medications as required during flares. When a flare was reported, the app prompted additional daily questions for the duration of the flare, including pain during activity and at rest (0-10 NRS), presence of swollen joint(s), and presence of warm joint(s), in addition to the use of any pain-relief/anti-

inflammatory medication. These daily flare items enabled classification of which flares met the validated Gaffo-definition.⁷

Data preparation

Physical activity data: Daily step counts were derived from the Garmin tracker via Fitabase exports. To ensure data quality, only days in which the device was worn for at least 80% of a possible 64 15-minute epochs, representing 16 waking hours (and excluding 8 hours during which the participant will be assumed to be sleeping) were included. The 16-hour period provided a consistent denominator for evaluating daytime wear completeness and was not intended to represent each participant's actual waking duration.¹⁶ Days that did not meet this pre-specified wear-time validity criterion were treated as missing for steps and excluded from the activity analyses. Because missingness in wearable-derived activity data may arise from a combination of device non-wear, syncing or technical failures, or participant behaviour, its mechanism could not reliably be established from the available information. Therefore, missing values were not imputed.

Sleep data: Daily sleep parameters were derived from the Garmin tracker via Fitabase exports (using the "Most Complete Version" export option in Fitabase). Sleep epochs are detected by Garmin's proprietary sleep detection algorithm and reported in minutes within the participant's sleep window (based on usual sleep/wake times entered at device set-up). For the purpose of this analysis, two sleep outcomes were constructed: total sleep time (minutes), calculated as the sum of daily deep sleep minutes and daily light sleep minutes; and deep sleep percentage (%), calculated as deep sleep minutes divided by (deep + light sleep minutes), expressed as a percentage. These sleep outcomes were selected as they capture complementary aspects of sleep (overall sleep quantity and a restorative component of sleep architecture) that are biologically relevant to inflammation and pain sensitivity.^{18,19} Nights without a Garmin sleep record (for example, due to device non-wear during sleep,

failure to detect sleep within the configured sleep window, or unsuccessful syncing/export) were treated as missing and excluded pairwise from sleep analyses.

Gout flare data: Daily flare status and flare features were obtained from MyCap flare diaries. Flare days were classified in two ways (run as separate/sensitivity analyses): (1) daily adapted Gaffo flare classification: days meeting ≥ 3 of 4 patient-reported criteria: current self-reported flare, current pain at rest $> 3/10$, presence of ≥ 1 swollen joint, and presence of ≥ 1 warm joint;⁷ and (2) self-reported flare days: any day the participant reported being in a flare. The first classification was a daily adaptation of the validated Gaffo flare definition; unlike the original definition, in which pain at rest referred to the maximum pain experienced over the preceding week, all criteria in the present study were assessed daily and pain at rest referred to current symptoms. A flare episode was defined as the maximal sequence of consecutive flare days under the relevant definition. All eligible flare episodes experienced by participants across the follow-up period were included in the analysis.

Flare “resolution” timeframes (T1–T3): Gout flare resolution was defined as a ‘timeframe’ that followed a Gaffo-defined gout flare episode (or a self-reported gout flare episode as the comparator in the sensitivity analyses). Three resolution timeframes were constructed around each flare episode, during which activity and sleep data were considered representative of gout flare resolution (**Figure 1**):

- **T1:** The first non-flare day that immediately followed the final flare day and did not overlap with a subsequent flare.
- **T2:** The consecutive non-flare window - equal in duration (days) to the flare episode - that immediately followed the flare and did not overlap with subsequent flares.
- **T3:** A delayed non-flare window intended to reduce carryover from residual gout flare symptoms and/or anti-inflammatory treatment (**Figure 1**). A washout period

equal in duration to the flare episode was first defined as the consecutive non-flare days immediately following the final flare day. The T3 resolution window was then defined as the consecutive non-flare period of equal duration to the preceding flare episode. The T3 window was included only when it did not overlap with subsequent flares or washout periods.

To control for confounding effects in comparisons of physical activity and sleep outcomes, and to avoid overweighting longer episodes, physical activity and sleep measured across multiple flare days were defined as the episode-level mean across those days. This approach was selected to estimate differences between episode-level flare and resolution states rather than day-by-day recovery trajectories. Episodes lacking valid, non-overlapping windows were excluded for that timeframe. The timeframes were designed a priori by the research team (with extensive experience in gout flare measurement) to not only ensure comparability of flare periods, but also to minimise carryover/contamination from the index flare and any flare medication use, and to reflect clinically meaningful post-flare recovery.

Statistical analysis

To address the study objectives, associations between gout flare status (flare vs resolution) and (1) physical activity, and (2) sleep, across 6 months of remote follow-up were evaluated. Each analysis was performed using two different flare definitions: first using the Gaffo flare definition and the second using self-reported flare as a sensitivity definition.

Flare episodes and timeframes (independent variables): The flare period and three post-flare resolution timeframes were constructed as described above and illustrated in **Figure 1**.

Separate paired episode-level datasets were created for **T1**, **T2**, and **T3**. Each dataset contained the episode-level mean outcome during the flare and the corresponding outcome from the relevant valid, non-overlapping post-flare day or window. This approach was

selected to compare clinically interpretable flare and resolution states rather than day-by-day recovery trajectories.

Covariates/other factors: Covariates were selected a priori based on their clinical relevance and potential to confound associations between flare status and activity or sleep outcomes. All models adjusted for participant-level factors (baseline serum urate, age group, ethnicity group, body mass index, and gout disease duration) and flare episode duration (number of days in flare). Step count models also adjusted for lower-limb involvement, defined as involvement of a lower-limb joint on at least one day during the flare. For the single-day timeframe (**T1**), models additionally adjusted for day type (workday vs non-workday) and an episode-level indicator of any anti-inflammatory medication use during the flare episode (any flare day with medication). For window timeframes (**T2**, **T3**), models adjusted for the proportion of workdays within the flare window and within the post-flare window.

Anti-inflammatory medication exposure was represented by the flare-episode indicator (as above). In supplementary sensitivity analyses of the indicator of any medication use during the flare was replaced by medication use on the first flare day to examine whether accounting for early treatment influenced the results.

Modelling strategy: Data were organized as paired episode-level datasets using a long format (each participant contributed multiple rows based on the number of gout flare episodes). For each flare episode, one of the matched resolution days/windows was included next to the episode flare for all outcomes and other variables. All analyses were independently conducted by two researchers (an experienced researcher (SS) and a biostatistician (IZ)), and results were cross-checked for consistency. Analyses were conducted in R (v. 2025.09.01) using the *lqmm* package for linear quantile mixed models.^{20,21} This method was adopted because standard linear regression models (Gaussian errors), as per the initial analysis plan, produced non-normal residuals for all outcomes. To retain estimates on the original outcome

scale and avoid changing the outcome distribution family by using transformations, median ($\tau = 0.5$) quantile mixed modelling was used to provide robust and easily interpretable adjusted differences in the conditional median. All eligible flare episodes experienced by participants across the 6-month follow-up period contributed to the analysis. As each participant could contribute multiple flares, all analyses included a participant-specific random intercept to account for within-person repeated flare measures. Analyses were performed using timeframe-specific complete cases, requiring both flare states (flare and resolution) to be present for each pair. A complete-case approach was used for outcome data, with analyses restricted to observed values and no imputation of missing outcomes. Consequently, the number of participants and flare episodes contributing to each analysis varied according to the availability of data. Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC) were extracted for each model to determine which timeframe (**T1-T3**) better defined resolution (i.e., more clearly distinguished resolution from flare comparator observations), while balancing model fit. Primary estimates were reported as adjusted differences in the conditional median. Two-sided p-values and 95% CIs were derived using Wald z tests (normal approximation) based on the estimated covariance matrix of fixed effects.

Baseline-standardised sensitivity analysis: A sensitivity analysis examined whether the differences in daily step count between each flare and its resolution timeframes were retained after standardising step count to participants' baseline activity. This was represented by each flare's intercritical-gout reference state, defined a priori as the third non-flare day prior to flare onset (day -3) (**Figure 1**). This day was selected to minimise contamination by immediate pre-flare (prodromal) changes in the 1-2 days preceding a flare, while also avoiding overlap with prior flare episodes and any previously defined post-flare resolution or washout windows. This analysis was restricted to daily step count using the Gaffo-defined

flare classification, because this was the only outcome/definition combination showing clear differences between flares and resolution timeframes in the primary analyses. For each eligible episode, daily step count during flares and the corresponding resolution timeframe were expressed as the % difference from day -3. The same linear quantile mixed modelling approach was used as above, with the exception of participant-level covariates, which were excluded because standardisation to day -3 was intended to account for the between-person variation in usual activity.

RESULTS

Participants

Table 1 summarises baseline characteristics of the 81 participants included in this analysis. Participants were predominantly male and ethnically diverse, with a substantial proportion with high body mass index and cardiometabolic comorbidity. Tophi were present in 11% of participants and mean number of gout flares in the preceding six months was 4 (range 2-30). Urate-lowering therapy was used by just under two-thirds.

Physical activity, sleep, and gout flare outcomes

Across 14,904 participant-days of follow up during the 6-month monitoring period, 78.5% of days met the prespecified $\geq 80\%$ validity criterion. **Table 2** summarises daily step counts, nightly sleep characteristics and gout flare outcomes across the observation period. Self-reported flares were more common than Gaffo-defined flares, most flares were mono-articular, and as expected, pain ratings during flares tended to be higher for Gaffo-defined flares than for self-reported flares. The median (IQR) daily step count, total sleep time, and deep sleep percentage across intercritical gout, flare episodes and resolution timeframes for

Gaffo-defined flares is shown in **Figure 2** and for self-reported flares, in **Supplementary File 1**.

Physical activity and gout flare resolution

Across the physical activity analyses, 61 to 75 participants contributed 175 to 474 flare episodes, depending on the follow-up timeframe and flare definition used (**Supplementary File 2**). Using the Gaffo flare definition, daily step count was lower during flares compared with post-flare resolution for all three predefined timeframes (**T1-T3**) (**Table 3**). When applying the self-reported flare definition, there were no significant differences in daily step count between flares and any resolution timeframes. In some physical activity models, higher BMI, older age, and lower-limb involvement were associated with lower daily step count, independent of any differences in step count between flare and resolution states (full covariate effects are presented in **Supplementary File 3**). For steps, the resolution timeframe that provided the best overall model fit (lowest AIC and BIC) was **T3** for both Gaffo-defined flares (AIC=6802; BIC=6868) and self-reported flares (AIC=12048; BIC=12123), suggesting that resolution is most clearly distinguished from flare states when using this timeframe.

Baseline-standardised sensitivity analysis: These analyses included 53 to 56 participants contributing 123 to 142 flare episodes with complete D-3, flare, and resolution data (**Supplementary File 2**). When expressing the flare and resolution step counts as the % difference from the episode-specific day -3 value, median daily step count was 9.2% below day -3 during flares, 3.3% below at **T1**, 5.0% below at **T2**, and 2.6% above baseline at **T3**. In the models, only **T3** differed significantly from flares, with a 14.3% higher daily step count during resolution (95% CI 3.2- 25.4; $P=0.015$), whereas **T1** and **T2** were not significantly different from flares (**Supplementary File 4**). Considered alongside the primary analyses, these results suggest that although T1 and T2 resolution timeframes may capture early

improvement in steps, T3 may be the most meaningful resolution window when considering resolution as a return to an individual's baseline steps.

Sleep and gout flare resolution

Across the sleep analyses, 62 to 78 participants contributed 176 to 477 flare episodes, with sample sizes varying according to follow-up timeframe and flare definition (**Supplementary File 2**). Across both flare definitions, there were no statistically significant differences in total sleep time or deep sleep percentage for any of the flare-resolution timeframe contrasts (**T1-T3**) (**Table 4**). Although there were no differences in sleep outcomes between flare and resolution states, higher BMI was commonly associated with shorter total sleep time, particularly across the Gaffo-defined flare models. Full covariate results are provided in **Supplementary File 3**. For both total sleep time and deep sleep %, **T3** consistently provided the lowest AIC and BIC for Gaffo-defined flares (total sleep time AIC=4219, BIC=4285; deep sleep % AIC=3006, BIC=3071) and self-reported flares (total sleep time AIC=7215, BIC=7290.0; deep sleep % AIC = 5194, BIC=5270), suggesting this time frame provided the clearest definition of resolution relative to flare state observations for sleep outcomes.

In supplementary sensitivity analyses replacing any medication use with medication use on the first flare day, the primary findings were unchanged (**Supplementary File 5**).

DISCUSSION

The findings from this analysis show that objectively measured activity improves after a flare resolves, with the clearest distinction between flare and resolution observed over a longer post-flare recovery period. In contrast, objectively measured sleep did not distinguish

between flare and flare-resolution states. These findings identify step count as a potentially responsive objective domain and inform the timing of its assessment in the future development of a multidimensional flare resolution definition.

The baseline-standardised sensitivity analysis supported the view that resolution is better understood as movement back towards an individual's non-flare baseline rather than a binary transition on the first non-flare day. When flare and resolution were adjusted for this intercritical/baseline activity, only the longer delayed recovery window remained significantly different from flare, suggesting that a clearer return towards usual non-flare activity is more evident over a longer recovery period. This aligns with prior wearable device work showing lower step counts during flare periods compared with non-flare periods,¹⁶ and with recent qualitative work in which return of function and mobility is described as an early and important marker of resolution, often preceding complete resolution of inflammatory symptoms.¹⁴

Importantly, daily step count varied substantially across participants, and older age and higher BMI were consistently associated with lower activity irrespective of flare status, highlighting marked between-person differences in usual activity level. Together with the sensitivity analysis showing the clearest baseline-standardised comparison between flare and resolution states in the later recovery window, this suggests that activity-based definitions of flare resolution may be more meaningful when interpreted relative to an individual's non-flare reference state rather than a single universal threshold.

Differences in step counts between flare and resolution periods were observed only when flares were defined with the stricter Gaffo definition, and not when flares were based on self-report alone. This suggests that the two flare definitions may identify somewhat different flare episodes, with the Gaffo definition potentially capturing episodes associated with greater disruption of physical activity. However, these results should not be interpreted as

evidence that one flare definition is more clinically meaningful than the other, as self-reported flares may include episodes that are important to people with gout but have less impact on objective activity measures.

Although gout flares can disrupt sleep and people with gout frequently report poor sleep quality and daytime sleepiness,²²⁻²⁴ the examined sleep metrics did not differentiate flare from resolution states. Sleep problems in gout may not be confined to flare periods, but may also reflect chronic sleep disorders, psychological distress, or ongoing joint pain, reducing their ability to discriminate flare from non-flare states. These results are consistent with previous device-derived work showing no statistically significant differences in total sleep duration between flare and non-flare periods,¹⁶ and with qualitative research in which sleep is not described as a prominent feature in defining flare resolution, compared with pain and function.¹⁴ The null findings may also reflect limitations of the sleep metrics examined, as total sleep time reflects quantity rather than quality, deep sleep percentage is only a partial marker of restorative sleep, and consumer-grade wearables may both stage sleep imprecisely²⁵ and overestimate sleep duration by misclassifying pain-related immobility as sleep.²⁶ More detailed sleep metrics (such as sleep efficiency, awakenings, or time awake after sleep onset) may prove more sensitive for detecting flare-related sleep disturbance in future studies.

This analysis has some limitations. First, the sample was almost entirely male, which limits generalisability to women with gout and precludes meaningful sex-stratified analyses. The participation requirements may also have underrepresented people with gout who had lower digital literacy or limited access to technology, which also limits generalisability of these findings beyond similar populations. Although accounting for early medication use did not materially alter the findings, variation in medication type, dose, and duration may still have influenced the speed and pattern of resolution. Variation in flare characteristics may also have

influenced the extent to which flares disrupted mobility or sleep. Although lower-limb involvement was associated with lower step count in some models, this finding was not consistent and residual variation in the specific joints involved, number of affected joints and pain severity may have contributed to variability between flare and resolution states. The complete-case approach, together with the requirement for valid, non-overlapping resolution/intercritical windows may have excluded some flare episodes (particularly in participants with frequent or closely spaced flares). In addition, missingness in wearable-derived data is methodologically complex,^{27,28} and alternative assumptions or analytic strategies may have produced different estimates. Finally, the intercritical reference state used in the sensitivity analysis was based on a single day (day -3) per episode, which may have been more susceptible to day-to-day variation in activity unrelated to gout than a multi-day baseline measure. Furthermore, episode-level means prevented longer flares from contributing disproportionately but did not capture daily recovery trajectories, which warrant future time-varying analyses. Although T3 provided the clearest distinction between flare and resolution among the three predefined timeframes, it was a clinically informed timeframe rather than a validated definition of post-flare recovery. Further research is needed to characterise recovery trajectories and determine the timing and duration of a clinically meaningful post-flare recovery period.

Despite these limitations, this analysis has important strengths. Firstly, the ethnically diverse sample, including a representative proportion of Māori and Pacific participants, supports the applicability of these results to those most affected by gout in Aotearoa New Zealand. Given that Māori and Pacific Peoples experience higher flare frequency and greater activity limitation,²⁹ the observed associations with activity and pain likely reflect resolution domains which are clinically meaningful to these groups. This study was also strengthened by the prospective ascertainment of flares which minimised recall bias, and the use of objective,

continuously measured activity and sleep data which provided a low-burden assessment of real-world behaviour. Finally, the use of pre-specified resolution timeframes that align closely with clinically interpretable phases of flare resolution strengthened internal validity and interpretability.

In conclusion, daily step count improves quickly after a gout flare resolves, but the clearest distinction between flare and resolution was observed over a longer post-flare recovery period. In contrast, wearable-derived total sleep time and deep sleep percentage do not reliably differentiate flare from resolution states. These findings suggest that flare resolution is dynamic, and that definitions of resolution may be strengthened by considering person-specific recovery of activity measured over several days rather than relying on a single post-flare timepoint.

Acknowledgements: We would like to thank Arthritis NZ Mateponapona Aotearoa, Akoranga Integrated Health Clinic, Turuki Health Care, and Counties Manukau Middlemore Hospital for their assistance with recruitment for this study. OpenEvidence was used as an adjunctive tool to assist in identifying potentially relevant literature. All AI-generated outputs and suggested references were independently reviewed and verified against original sources by the authors.

References

1. Dalbeth N, Gosling AL, Gaffo A, Abhishek A. Gout. *The Lancet*. 2021;397(10287):1843-1855.:[https://doi.org/10.1016/S0140-6736\(21\)00569-9](https://doi.org/10.1016/S0140-6736(21)00569-9)
2. Bursill D, Taylor WJ, Terkeltaub R, et al. Gout, Hyperuricemia, and Crystal-Associated Disease Network Consensus Statement Regarding Labels and Definitions for Disease Elements in Gout. *Arthritis Care & Research*. 2019;71(3):427-434. doi:10.1002/acr.23607
3. Gaffo AL, Schumacher HR, Saag KG, et al. Developing a provisional definition of flare in patients with established gout. *Arthritis & Rheumatism*. 2012;64(5):1508-1517. doi:10.1002/art.33483
4. Stewart S, Guillen AG, Taylor WJ, et al. The experience of a gout flare: a meta-synthesis of qualitative studies. *Seminars in Arthritis and Rheumatism*. Aug 2020;50(4):805-811. doi:10.1016/j.semarthrit.2020.06.001
5. Singh JA. Any sleep is a dream far away: A nominal group study assessing how gout affects sleep. *Rheumatology (United Kingdom)*. 2018;57(11):1925-1932. doi:10.1093/rheumatology/kex535
6. Sutherland C, Gaffo AL. Investigating gout flares: beyond a definition. *Current opinion in rheumatology*. 2024;36(4):309-313. doi:10.1097/BOR.0000000000001024
7. Gaffo AL, Dalbeth N, Saag KG, et al. Brief Report: Validation of a Definition of Flare in Patients With Established Gout. *Arthritis Rheumatol*. Mar 2018;70(3):462-467. doi:10.1002/art.40381
8. Stewart S, Dalbeth N, Gaffo A. The challenge of gout flare measurement. *Best practice & research Clinical rheumatology*. 2021;35(4):101716. doi:10.1016/j.berh.2021.101716

9. Mikuls TR. Gout. *The New England journal of medicine*. 2022;387(20):1877-1887. doi:10.1056/NEJMcp2203385
10. Neogi T, Jansen TLTA, Dalbeth N, et al. 2015 Gout classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Annals of the Rheumatic Diseases* 2015. p. 1789-1798.
11. Pascual E, Batlle-Gualda E, Martínez A, Rosas J, Vela P. Synovial fluid analysis for diagnosis of intercritical gout. *Ann Intern Med*. Nov 16 1999;131(10):756-9. doi:10.7326/0003-4819-131-10-199911160-00007
12. Ea H-K, Kischkel B, Chirayath TW, et al. Systemic inflammatory cytokine profiles in patients with gout during flare, intercritical and treat-to-target phases: TNFSF14 as new biomarker. *Annals of the Rheumatic Diseases*. 2024/07/01/ 2024;83(7):1-12. <https://doi.org/10.1136/ard-2023-225305>
13. Sandoval-Plata G, Morgan KM, Doherty M, Abhishek A. Inter-Critical Gout, Not Hyperuricemia or Asymptomatic Urate Crystal Deposition, Is Associated with Systemic Inflammation. *Gout, Urate, and Crystal Deposition Disease*. 2025;3(3):11. doi:10.3390/gucdd3030011
14. Stewart S, Horne A, Gaffo A, Stamp LK, Edwards NL, Dalbeth N. A return to normality: A descriptive qualitative interview study exploring the patient experience of gout flare resolution. *Arthritis Care & Research*. 2026 (accepted) doi: 10.1002/acr.80089
15. Pouls BPH, Bekker CL, Gaffo AL, Bemt BJFvd, Flendrie M. Tele-monitoring flares using a smartphone app in patients with gout or suspected gout: a feasibility study. *Rheumatology Advances in Practice*. 2021;5(3):1-8. doi:10.1093/rap/rkab100

16. Elmagboul N, Coburn BW, Foster J, et al. Physical activity measured using wearable activity tracking devices associated with gout flares. *Arthritis Research & Therapy*. 2020;22(1):181. doi:10.1186/s13075-020-02272-2
17. Phipps-Green AJ, Merriman ME, Topless R, et al. Twenty-eight loci that influence serum urate levels: analysis of association with gout. *Ann Rheum Dis*. 2016;75(1):124-30. doi:10.1136/annrheumdis-2014-205877
18. Besedovsky L, Lange T, Haack M. The sleep-immune crosstalk in health and disease. *Physiol Rev*. 2019;99(3):1325-1380. doi:10.1152/physrev.00010.2018
19. Kourbanova K, Alexandre C, Latremoliere A. Effect of sleep loss on pain—new conceptual and mechanistic avenues. systematic review. *Frontiers in Neuroscience*. 2022;16. doi:10.3389/fnins.2022.1009902
20. Geraci M. Linear Quantile Mixed Models: The lqmm Package for Laplace Quantile Regression. *Journal of Statistical Software*. 2014;57(13):1 - 29. doi:10.18637/jss.v057.i13
21. Geraci M, Bottai M. Linear quantile mixed models. *Statistics and Computing*. 2014;24(3):461-479. doi:10.1007/s11222-013-9381-9
22. Stewart S, Guillen AG, Taylor WJ, et al. The experience of a gout flare: a meta-synthesis of qualitative studies. *Semin Arthritis Rheum*. 2020;50(4):805-811. <https://dx.doi.org/10.1016/j.semarthrit.2020.06.001>
23. Coulshed A, Nguyen AD, Stocker SL, Day RO. Australian patient perspectives on the impact of gout. Multicenter Study. *Int J Rheum Dis*. 2020;23(10):1372-1378. <https://dx.doi.org/10.1111/1756-185X.13934>
24. Singh JA. The impact of gout on patient's lives: a study of African-American and Caucasian men and women with gout. *Arthritis research & therapy*. 2014;16(3):R132. doi:10.1186/ar4589

25. Budig M, Stoohs R, Keiner M. Validity of two consumer multisport activity tracker and one accelerometer against polysomnography for measuring sleep parameters and vital data in a laboratory setting in sleep patients. *Sensors*. 2022;22(23):9540.
26. Feehan LM, Geldman J, Sayre EC, et al. Accuracy of Fitbit Devices: Systematic Review and Narrative Syntheses of Quantitative Data. *JMIR mHealth and uHealth*. 2018;6(8):e10527. doi:10.2196/10527
27. Braem CIR, Yavuz US, Hermens HJ, Veltink PH. Missing data statistics provide causal insights into data loss in diabetes health monitoring by wearable sensors. *Sensors*. 2024;24(5). doi:10.3390/s24051526
28. Mobley K, Gittner KB, Nielsen KE, et al. Determining a likely mechanism of missingness in repeated measures sleep data from wearable fitness trackers: longitudinal analysis. *JMIR mHealth and uHealth*. 2026;14
29. Dalbeth N, House ME, Horne A, et al. The experience and impact of gout in Māori and Pacific people: a prospective observational study. *Clinical Rheumatology*. 2013;32(2):247-251. doi:10.1007/s10067-012-2110-5

TABLES

Table 1. Baseline participant characteristics

N		81
Age, years, mean (SD)		44.9 (11.5)
Sex, n (%)	Female	1 (1.2%)
	Male	80 (98.8%)
Ethnicity, n (%) ^a	Asian	22 (27.2%)
	European	35 (43.2%)
	Māori	9 (11.1%)
	Pacific Peoples	22 (27.2%)
	Other	2 (2.5%)
Body mass index, kg/m ² , mean (SD)		33.2 (8.1)
Gout disease duration, years, mean (SD)		11.3 (9.8)
Number of gout flares in preceding 6 months, mean (SD)		4.0 (3.5)
Subcutaneous tophi, n (%)		9 (11.1%)
Serum urate, mmol/L, mean (SD)		0.41 (0.13)
Urate lowering therapy use, n (%)		52 (64.2%)
Other medications, n (%)	Colchicine	25 (30.9%)
	NSAIDs	32 (39.5%)
	Analgesics	7 (8.6%)
	Corticosteroids	13 (16.0%)
	Antihypertensives	17 (21.0%)
	Lipid lowering agents	12 (14.8%)
	Antidiabetics	8 (9.9%)
	Antiplatelets/anticoagulants	7 (8.6%)
	Diuretics	1 (1.2%)
Comorbidities, n (%)	Cardiovascular disease	4 (4.9%)
	Type 2 diabetes	8 (9.9%)
	Hypertension	19 (23.5%)
	High cholesterol	15 (18.5%)
	Respiratory disease	11 (13.6%)
	Sleep apnoea	3 (3.7%)

Anxiety or depression	8 (9.9%)
Obesity	46 (56.8%)
Pain at rest due to gout (0-10), mean (SD) ^b	1.5 (2.3)
Pain during activity due to gout (0-10), mean (SD) ^b	1.9 (2.4)
Overall pain (0-10), mean (SD) ^b	2.0 (2.1)
Patient global score (0-10), mean (SD) ^b	2.1 (2.7)

^aSome participants identified with multiple ethnic groups. ^bScores reflect baseline non-flare symptoms (past-week pain and current-day global assessment), not flare-day diary ratings. NSAIDs = non-steroidal anti-inflammatory drugs; B-PSQI = Brief Pittsburgh Sleep Quality Instrument.

Table 2. Descriptive statistics for physical activity, sleep, and gout flare outcomes

		Gaffo-defined	Self-reported
Daily step count, median (IQR)	Intercritical gout	7157 (5492)	6713 (5442)
	Flare episode	6021 (5567)	6593 (5201)
	T1	6629 (6124)	6688 (5935)
	T2	6632 (4974)	6646 (5151)
	T3	7305 (5701)	7092 (5319)
Total sleep time (minutes), median (IQR)	Intercritical gout	433 (150)	443 (115)
	Flare episode	425 (125)	441 (117)
	T1	431 (153)	437 (105)
	T2	431 (136)	437 (130)
	T3	428 (131)	446 (117)
Deep sleep percent (%), median (IQR)	Intercritical gout	40.8 (23.9)	40.8 (22.7)
	Flare episode	40.5 (22.0)	40.9 (20.0)
	T1	39.2 (24.3)	41.3 (23.2)
	T2	41.7 (20.7)	41.1 (19.3)
	T3	40.8 (23.9)	40.8 (22.7)
Observed days with a flare, n (%)		684 (5.4%)	1722 (11.6%)
Total number of individual gout flare episodes, n		281	563
Participants with at least 1 flare, n (%)		69 (85.2%)	81 (100.0%)
Number of flares per participant, mean (SD)		4.1 (4.0)	7.0 (6.4)
Duration of flares (days), mean (SD)		2.4 (2.6)	3.1 (7.9)
Pain during rest (0-10 NRS), mean (SD)		3.4 (2.2)	2.3 (1.9)
Pain during activity (0-10 NRS), mean (SD)		4.1 (2.4)	3.0 (2.2)
Any medication use during flare, n (%)		237 (84.3%)	397 (70.5%)
Medication use on first flare day, n (%)		222 (79.0%)	358 (63.6%)
Location of flare ^a	Ankle	77 (27.4%)	165 (29.3%)
	Big toe	91 (32.4%)	200 (35.5%)
	Elsewhere in foot	81 (28.8%)	178 (31.6%)
	Knee	65 (23.1%)	98 (17.4%)
	Elbow	13 (4.6%)	19 (3.4%)
	Hand/wrist	29 (10.3%)	65 (11.5%)
	Other	3 (1.1%)	9 (1.6%)

Any lower-limb joint involvement, n (%)		258 (91.8%)	509 (90.4%)
Joint involvement, n (%)	Mono-articular	217 (77.2%)	415 (73.7%)
	Multi-articular	52 (18.5%)	124 (22.0%)
	Not reported	10 (3.6%)	22 (3.9%)

^aSome gout flares involved more than one location, therefore percentages sum to more than 100%. Intercritical gout was represented by the third eligible non-flare day prior to flare onset (day -3); **T1-T3** are the predefined post-flare timeframes illustrated in **Figure 1**; NRS = Numeric Rating Scale; IQR = interquartile range.

Table 3. Adjusted differences in daily step count between resolution timeframes (T1-T3) and gout flares

Term	Gaffo-defined flares			Self-reported flares		
	Estimate	95% CI	P	Estimate	95% CI	P
T1 vs flare	672.70	285.64-1059.76	0.001	333.67	-23.94-691.29	0.067
T2 vs flare	520.08	68.99-971.18	0.024	202.38	-154.17-558.92	0.226
T3 vs flare	914.73	386.28-1542.38	0.004	538.96	-26.93-1104.85	0.068

Estimates are adjusted differences in the conditional median ($\tau=0.5$) daily step count. Estimates are shown as resolution timeframe minus flare (flare comparator is the episode-level mean across flare days). Positive values indicate higher daily step count in the resolution timeframe while negative values indicate lower daily step count. Models include a participant random intercept and adjust for prespecified covariates (full covariate estimates in **Supplementary File 2**).

Table 4. Adjusted differences in daily sleep outcomes between resolution timeframes (T1-T3) and gout flares

Term	Gaffo-defined flares			Self-reported flares		
	Estimate	95% CI	P	Estimate	95% CI	P
Total sleep time (minutes)						

T1 vs flare	0.62	-14.05- 15.29	0.934	-1.19	-10.94-8.56	0.813
T2 vs flare	4.49	-10.51- 19.50	0.560	-0.14	-10.22-9.94	0.978
T3 vs flare	6.99	-5.73- 19.71	0.287	-0.64	-11.32- 10.03	0.907
Deep sleep percentage (%)						
T1 vs flare	0.34	-1.91-2.58	0.770	0.53	-1.20-2.25	0.552
T2 vs flare	1.18	-1.10-3.47	0.316	0.37	-1.33-2.08	0.669
T3 vs flare	0.52	-1.99-3.03	0.687	0.06	-1.83-1.94	0.954

*Estimates are adjusted differences in the conditional median ($\tau=0.5$) total sleep time (minutes) or deep sleep %. Estimates are shown as resolution timeframe minus flare (flare comparator is the episode-level mean across flare days). Positive values indicate higher total sleep time/deep sleep % during the resolution timeframe while negative values indicate lower total sleep time/deep sleep %. Models include a participant random intercept and adjust for prespecified covariates (full covariate estimates in **Supplementary File 2**).*

Stewart Sarah (Orcid ID: 0000-0002-9318-5627)
Gaffo Angelo L. (Orcid ID: 0000-0001-7365-7212)
Stamp Lisa Katrina (Orcid ID: 0000-0003-0138-2912)
Dalbeth Nicola (Orcid ID: 0000-0003-4632-4476)

Gout flare resolution, physical activity and sleep

FIGURE LEGENDS

Figure 1. The three “resolution” timeframes (T1-T3) constructed around each gout flare episode. The intercritical gout state is represented as the third non-flare day prior to each index flare (day -3). *Note: resolution timeframes are only valid for the index flare if non-overlapping with subsequent gout flare episodes. Intercritical gout days are only valid for the index flare if non-overlapping with prior gout flare episodes, resolution timeframes or washout windows.*

Figure 2. Unadjusted median (IQR) daily step counts (A), total sleep time (minutes) (B), and deep sleep percentage (%) (C) across intercritical gout periods, flare episodes and resolution timeframes for Gaffo-defined flares.

Supporting Information

SUPPLEMENTARY FILE 1. Unadjusted median (IQR) daily step counts (A), total sleep time (minutes) (B), and deep sleep percentage (%) (C) across intercritical gout periods, flare episodes and resolution timeframes for self-reported gout flares.

SUPPLEMENTARY FILE 3. Results from fully-adjusted linear quantile mixed models (LQMM)

SUPPLEMENTARY FILE 4. Baseline-standardised sensitivity analysis

SUPPLEMENTARY FILE 5. Supplementary sensitivity analyses including early medication treatment

 Intercritical gout

 Index flare

 Resolution timeframe

 Washout period

———— Consecutive days —————→



