

# Patient-reported flares were correctly predicted by an algorithm using Machine-Learning statistics on activity tracker data on steps, in a longitudinal 3-month study of 170 patients with rheumatoid arthritis (RA) or axial spondyloarthritis (axSpA). The ActConnect study

Laure Gossec 1, Frédéric Guyard 2, Didier Leroy 2, Thomas Lafargue 2, Michel Seiler 3, Charlotte Jacquemin 1, Anna Molto 4, Jérémie Sellam 5, Violaine Foltz 1, Frédérique Gandjbakhch 1, Christophe Hudry 4, Stéphane Mitrovic 1, Bruno Fautrel 1, Hervé Servy 6

1. UPMC Univ Paris 06; Pitié Salpêtrière hospital, APHP, Paris, France. 2. Orange IMT, Nice France. 3. Orange Healthcare, Paris, France. 4. PRES Sorbonne Paris-Cité; Cochin Hospital, APHP, Paris, France. 5. INSERM UMRS\_938, UPMC Univ Paris 06, St-Antoine Hospital, APHP, Paris, France. 6. Sanoïa, e-Health services, Géménos, France.

## BACKGROUND

- The natural history of rheumatoid arthritis (RA) and axial spondyloarthritis (axSpA) comprises periods of low disease activity and flares.
- Flares are of interest to patients and clinicians since they reflect disease activity fluctuations, and may be associated with long-term outcomes. (1,2)
- However, there are few data linking patient-reported flares to quantifiable outcomes and objective measures.
- Physical activity and in particular walking may be influenced by flares. It can be objectively and precisely measured using activity trackers.
- We recently published the ActConnect study (3): a 3-months longitudinal study of patients with either RA or axSpA where patient-reported flares were assessed weekly and patients wore a connected activity tracker. We found flares were related to a moderate decrease in physical activity, since weeks with flares led to a relative decrease in physical activity of 12-21%, i.e., an absolute decrease of 836-1462 steps/day. (3). At the group level and on amalgamated data using usual statistics, the link between flares and physical activity was weak and it was not possible to determine modifications in physical activity patterns which could adequately predict patient-reported flares.
- Recently, there has been an explosion of Big Data and of Machine-Learning Statistics.(4) Machine-Learning Statistics allow multiple analyses of large datasets and make the best use of the available data, with minimal data amalgamation.

## OBJECTIVE

- The objective was to assess longitudinally the association between patient-reported flares and activity-tracker-provided continuous flows of steps per minute
- By reanalysing the ActConnect dataset using Machine-Learning Statistics.

## METHODS

- This is a reanalysis of the French prospective multi-center observational ActConnect study
- Inclusion criteria were: definite axSpA (ASAS criteria) or RA (ACR/EULAR criteria), owning a smartphone, and internet access
- All patients were provided with an activity tracker (Withings® Activité Pop watch) connected by Bluetooth and were instructed to wear it every day for 3 months. Thus, over 3 months, physical activity was assessed continuously by number of steps/minute, using a consumer grade activity tracker. The Withings® tracker records the number of steps per minute. Ninety consecutive days from the first Monday following activation of the device were collected. No instruction about physical activity was given to the participants.
- In parallel, flares were self-assessed weekly from home, using a specific flare question: « your disease flared up since the last assessment ? », with a categorical response according to: no flare, flare lasting 1 to 3 days (short flare) or flare lasting more than 3 days (persistent flare). Here flares were analysed as a binary information (yes/no).

- The ActConnect study collected physical activity information (steps) at the minute level, during 3 months for 170 patients. In this analysis of the ActConnect dataset, machine-learning statistical methods using selective (multiclass) naive Bayesian classification were applied.
- Physical activity data were first normalized at the patient level using each patient's mean and standard deviation of steps for a similar timeframe without flares.
- Then the data were analysed by multiclass Bayesian methods with a Machine Learning software belonging to Orange: Khiops® (5). The software was instructed to find the best predictive model of patient-reported flares.
- Sensitivities and specificities of the machine-generated models of physical activity were calculated when predicting flares (gold standard=patient-reported flare), based on steps and kappas were computed.
- Several sensitivity analyses were performed using different physical activity timeframes (ranging from 1 hour to 24 hours), and the analyses were performed 10 times on different validation sets randomly chosen from the dataset.
- More information on the statistics are given in [Poster 1234](#)

## RESULTS

Among the 170 patients included in the study, 155 (82 RA and 73 axSpA patients) were analyzed. (Table 1)  
This corresponds to 1339 weekly flare assessments and 224,952 hours physical activity assessment timeframes, ie close to 13.5 million points.

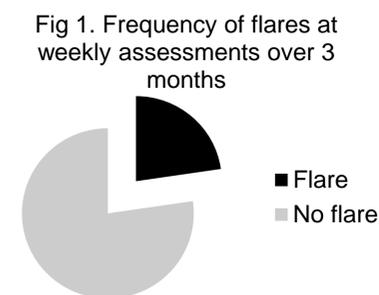
**Table 1: Characteristics of RA and axSpA patients**

	RA (N=82)	axSpA (N=73)
Sex, males, N (%)	14 (17.1)	41 (56.2)
Age, mean (SD)	48,9 (12.6)	41.2 (10.3)
BMI, mean (SD)	24.7 (4.5)	24.6 (4.6)
Disease duration (years), mean (SD)	10.5 (8.8)	10.8 (9.1)
Work status, employed, N (%)	61 (74.4)	61 (83.6)
Studies above high school, N (%)	69 (84.1)	66 (90.4)
Functional comorbidity Index (/18), mean (SD)	1.6 (0.9)	1.4 (0.9)
NSAIDs intake, N (%)	17 (20.7)	44 (60.3)
Glucocorticoid intake, N (%)	19 (23.2)	1 (1.4)
Synthetic DMARDs, N (%)	76 (92.7)	17 (23.3)
Biological therapy, N (%)	37 (45.1)	44 (60.3)
Stability of treatment over the last 3 months, N (%)	59 (72.0)	47 (64.4)
Baseline disease activity, mean (SD)	DAS28: 2.3 (1.2)	BASDAI: 3.2 (2.1)

\* Percentages are calculated on available information

**Patients had long-standing disease and around half of them were receiving a biologic (Table 1). Disease was well-controlled.**

**Flares were frequent: among the 155 patients, 112 (72.2%) patients reported at least one flare over the 3 months follow up. Patients reported having experienced a flare on average at 22.7% of the questionnaires (Fig 1).**



**Physical activity was moderate (Fig 2). Over all the assessments, the mean number of steps/day was 6838 (SD, 4033).**



Link between physical activity and self-reported flares

**Table 2. Prediction of flares by steps/hour using a Machine Learning approach on Khiops**

	Patient-reported flare	No patient-reported flare
Flare according to Khiops ©	88 (21.8%)	10 (2.4%)
No flare according to Khiops ©	4 (0.9%)	301 (74.6%)

**Results for one of the validation subsets (set 8): results are N and % of 491 weekly assessments of flare**

**The Khiops generated model detected correctly both patient-reported flares and absence of flare (Table 2 presents the results on one of the validation sets; the abstract gave bootstrapped global results explaining the slight differences). Mean results over the different validation sets were a sensitivity of 97% and a specificity of 96%. The corresponding positive and negative predictive values were respectively 98% and 89%.**

**The different sensitivity analyses were confirmatory (see also [poster 1234](#)).**

## CONCLUSIONS

- This study confirms the objective consequences of patient-reported flares and the interest of machine-learning statistics on rich and wide datasets.
- Patient-reported flares are associated to less physical activity (measured by activity trackers) in RA and axSpA.
- When analysing the dataset using Machine-Learning Statistics, it was possible to predict with great precision patient-reported flares based on physical activity data.
- Connected activity trackers with Machine Learning processing may be used as continuous indicators of disease activity in RA and axSpA.

## DISCUSSION

- This study demonstrated that patient-reported flares are strongly linked to physical activity and that patterns of physical activity can be used to predict flares with great accuracy.
- This study also demonstrated the importance of the data preparation in large datasets: here the data had to be normalised since when analysing differences between weeks without normalisation at the patient level, the results were inconclusive.
- Multiple analyses were run and were confirmatory, however the relatively small sample size in the present study warrants caution when interpreting the results.

## NOTES

### References

- Bykerk VP, et al. Identifying flares in rheumatoid arthritis: reliability and construct validation of the OMERACT RA Flare Core Domain Set. RMD Open 2016;2:e000225.
- Fautrel B, et al. Validation of flare-ra, a self-administered tool to detect recent or current rheumatoid arthritis flare. Arthritis Rheumatol. Hoboken NJ 2016;doi: 10.1002/art.39850.
- Jacquemin C, et al. Flares assessed weekly in patients with rheumatoid arthritis or axial spondyloarthritis and relationship with physical activity measured using a connected activity tracker: a 3-month study. RMD Open. 2017 29;3(1):e000434
- Khiops software for data mining, PredicSis; accessed 06/01/2017: <https://khiops.predicSis>.

The ActConnect initial study was funded by unrestricted academic grants from Lilly France, BMS France and Pfizer France.  
Orange IMT provided for free the technical infrastructure, the software used to perform Machine Learning statistics and the experts to perform data management and analysis.