A randomized controlled trial for improving patient self-assessment of synovitis in rheumatoid arthritis with education by ultrasonography: the RAEUS Study

Peter P. Cheung1,2,3, Manjari Lahiri1,2, Gim Gee Teng1,2, Anita Y. N. Lim1,2, Tang Ching Lau1,2, Aisha Lateef1,2, Anselm Mak1,2, Laure Gossec4 and Lyn March3

Abstract

Objective. Patients can potentially monitor disease activity of RA through self-assessed swollen joints (clinical synovitis), but reliability is poor. The objective is to evaluate the use of education by US feedback on the ability of patients to assess for clinical synovitis in RA.

Methods. We performed a 6 month, single-centre, randomized controlled trial on patients with established RA to study the effect of education on self-assessment of joints that included initial brief patient training on tender (TJC) and swollen (SJC) joint counts followed by US feedback every 3 months vs standard care without education. Patient and physician independently performed 28-joint counts at each visit. Outcome variables included the percentage of patients with good agreement with physician-derived swollen joints [prevalence-adjusted bias-adjusted kappa (PABAK) >0.6] as well as agreement in the SJC (Bland and Altman 95% limits of agreement), feasibility/patient satisfaction survey and disease activity at 6 months.

Results. Of the 101 randomized patients, 95 were included (51 in the education arm and 44 in the standard care arm). At 6 months there was a significant difference in the proportion of patients with swollen joint PABAK >0.6 in the education arm compared with standard care (98 vs 85%, P = 0.02). Limits of agreement for the SJC difference between physician and patients were reduced only in the education arm. The training method is considered feasible, with 94% of patients reporting it as useful. A trend of higher rates of disease remission (28-joint DAS <2.6) in the education arm vs standard care (47% vs 29%, P = 0.07) was seen.

Conclusion. A short course of education with US feedback may be helpful in educating patients to assess for clinical synovitis.


Key words: synovitis, rheumatoid arthritis, joint counts, self-report, ultrasonography.

Rheumatology key messages

- US can be used as an adjunctive training tool to improve patient assessment of swollen joints (clinical synovitis).
- A training programme to educate and improve patient self-assessment of clinical synovitis is acceptable and feasible for patients.

Introduction

RA is a chronic inflammatory disease predominantly affecting joints, with clinical synovitis, traditionally evaluated through the presence of joint swelling as part of a joint count assessment. Joint count is regarded as the gold standard in clinical disease activity assessment [1] and is included in the core data set of disease activity indices such as the 28-joint DAS (DAS28) [2].
Not all clinicians routinely perform a formal joint count in daily practice [3] due to time constraints, lack of resources for regular visits and also the belief that formal joint counts are not required when patients are clinically stable. However, the importance of joint assessment has been demonstrated in patients who continue to have radiographic progression despite being in clinical remission (DAS28 <2.6) when there is residual joint swelling [4]. Due to the need for regular disease activity monitoring in order to treat to target [5], there has been renewed interest in patient self-assessed joint counts [6–12]. Self-assessment of disease activity between clinic visits may assist patients in achieving and remaining in sustained clinical remission.

However, the reliability of patient-reported joint counts has been shown to be poor, in particular, the swollen joint count (SJC) [13]. Studies on the role of training in patient-reported joint counts are limited and the specific training methodologies are not well described and are of short duration without adequate longitudinal follow-up [10, 12, 13]. To date, novel effective ways of improving patient-reported joint counts have not been evaluated.

Ultrasonography has been increasingly used as an adjunctive bedside tool in the rheumatology clinic, especially for assessing the disease activity of RA [14, 15]. Although training and availability of equipment continue to be an issue, US feedback has been shown to be potentially useful in improving the clinical examination of swollen joints by physicians [16]. Hence this bedside tool may also have a positive impact on patient-reported swollen joints. The aim of this study was to longitudinally evaluate the use of education using US feedback on the ability of patients with RA to detect clinical synovitis (swollen joints) using the physician as the reference standard.

Methods

Study design
This was a 6 month, prospective, randomized controlled trial [RA Self-assessment Through Education and Ultrasonography (RAEUS) study].

Patients
Patients with established RA according to the 1987 ACR criteria [17] between the ages of 21 and 80 years from the outpatient clinic of the National University Hospital, Singapore, from November 2011 to June 2012 were included. Local ethics approval was obtained from the Domain Specific Review Board (National Healthcare Group, Singapore) and patients gave written informed consent.

Allocation of patients
A randomization sequence was generated, stratified by age (<50 years old and ≥50 years old) and disease activity (DAS28 <3.2 and DAS28 ≥3.2), and placed in opaque envelopes numbered chronologically, stapled together so that each patient recruited consecutively was allocated according to the envelope. At baseline, prior to randomization, all patients self-assessed 28 joints (shoulders, elbows, wrists, MCP joints, PIP joints and knees) for tenderness and swelling, followed by a blinded assessment by a physician (P.C. or M.L.). Study randomization was performed after physician joint count assessment at baseline to either (i) the active group [education of self-assessment with US feedback (EUS)] or (ii) the standard care group (no education of self-assessment and no US feedback). Selection and allocation biases were minimized with no patients excluded after the baseline physician joint count.

Self-assessment education with US feedback
Patients randomized to the active arm received an explanation by the physician on how to perform joint counts [tender joint count (TJC) and SJC] that took 5–10 min. This was followed by a US assessment of the 28 joints by a physician (P.C.) to demonstrate to the patients joints with or without active synovitis on power Doppler US (PDUS) as part of the training. Although not mandatory, patients were encouraged to practise self-assessment of their joints at home each week, such as at the time of their MTX dose. Adjustments in treatment were made by the physician based on clinical assessment, but the US results could be used if there was doubt about ongoing disease activity by clinical assessment alone.

Patients in the EUS group returned 3 months later and again self-assessed their joints, followed by an independent blinded physician assessment and then US feedback. The technique of joint count assessment and education on the location of active inflammation by PDUS was repeated during this visit, which took 15–20 min. Adjustment of treatment based on clinical assessment was made, if appropriate, and US results could be used if clinical assessment alone was not conclusive.

Standard care arm
At baseline, physician joint count and US assessment were similarly performed, but education and feedback on the location of active inflammation by PDUS was not provided. Neither did patients receive any formal training on how to perform joint counts. Treatment was adjusted by the study investigators based on the overall disease of the patient; however, as in the active arm, investigators were allowed to adjust treatment based on US results if disease activity was still in doubt based on clinical assessment alone. Patients were reviewed 3 months later by their treating rheumatologist, where US results performed at baseline could be reviewed as part of the standard consultation and treatment was adjusted if appropriate. However, patients were not asked to assess their own joints and did not receive further US scanning at this visit.

All patients in both arms returned at 6 months for a repeat evaluation of their joints similar to that performed at baseline. During the 6 month study period, treatments were changed according to the rheumatologist’s discretion and not based on a set protocol.

Collected data
Baseline patient and disease characteristics included age, sex, ethnicity, educational background, disease duration
and treatment with DMARDs. Other outcome variables included patient global assessment and functional impairment using the modified HAQ [18] as well as physical and mental components of health using the 12-item Short Form Health Survey (SF-12) [19]. The DAS28 was calculated using the TJC and SJC obtained by the physicians and patient, respectively.

Ultrasonography

A GE LOGIQ e US machine (GE Healthcare, Little Chalfont, UK) was used according to the protocol proposed by Backhaus et al. [20] on the same joints included in the 28-joint count. Each joint was scored for B-mode and PDUS on a validated semi-quantitative scale from 0 to 3, with the assistance of an atlas [21-23]. US semi-quantitative grading in B-mode was as follows: grade 0, absence, no synovial thickening; grade 1 = mild, minimal synovial thickening filling the angle between the periarticular bones without bulging over the line linking the tops of the bones; grade 2 = moderate, synovial thickening bulging over the line linking the tops of the periarticular bones but without extension to at least one bone diaphysis; grade 3, marked, synovial thickening bulging over the line linking the tops of the periarticular bones with and extension to at least one of the bone diaphyses.

PDUS signals were graded accordingly: grade 0 = absence, no synovial flow; grade 1 = mild, single-vessel signals; grade 2 = moderate, confluent signals in less than half of the synovial area; grade 3 = marked, signals in more than half of the synovial area (Fig. 1).

The physician (P.C.) who performed the US is experienced in the detection of synovitis and had achieved good interobserver reliability for both acquisition and interpretation of B-mode and PDUS [intraclass correlation coefficient (ICC) 0.75 (95% CI 0.66, 0.84) and ICC 0.59 (95% CI 0.48, 0.72), respectively] in a multicentre standardization exercise using the same machine [24].

Primary outcome

The primary outcome was the proportion of patients achieving good agreement, defined by prevalence-adjusted bias-adjusted kappa (PABAK) >0.6, with physician-derived swollen joints as the gold standard at 6 months. The strength of agreement was based on the following: PABAK =0.0, poor; 0–0.2, slight; 0.21–0.4, fair; 0.41–0.6, moderate; 0.61–0.8, substantial and 0.81–1, excellent [25].

Secondary outcomes

Secondary outcomes included comparison of the mean PABAK between the two groups at 6 months, as well as agreement of the TJC and SJC by the Bland and Altman method between patient and physician. Agreement of the DAS28 score change between patient- and physician-derived DAS28 from baseline to 6 months was evaluated, in addition to a feasibility and patient satisfaction survey from the EUS arm according to a 5-point Likert scale.

Disease activity outcome measures were also compared between the two groups at 6 months.

Statistical methods

A total sample size of 100 patients was calculated to detect a difference of 30% in patient and physician agreement (proportion of patients with PABAK >0.6) in the SJC between the two groups at 6 months with 80% power and a two-sided significance level of 5%, allowing for 10% dropout. Statistical analysis was carried out using SPSS version 20 (IBM, Armonk, NY, USA). Agreement at the joint level was compared using percentage agreement and the PABAK statistic, which adjusts kappa for skew in the data and for consistent rater bias. PABAK is calculated by the formula (2 × percentage exact agreement) – 1 [26].

All patients who completed at least one assessment post-baseline were included in the analysis, with the last observation carried forward for missing values. The primary outcome analysis compared the proportion of patients achieving a high agreement (PABAK >0.6) at 6 months in each study group using the chi-squared test. Absolute PABAK at 6 months was compared initially using t-tests and then linear regression adjusted for baseline PABAK levels between the two groups. Limits of agreement analysis were used to evaluate the TJC and SJC agreement between patients and physician at the patient level between the two groups at baseline and 6 months [27]. The Bland and Altman plot is a graphical representation of this analysis using a scatterplot of the difference between physician- and patient-derived joint count (the vertical axis) vs the average of the joint count value derived by physician and patient (horizontal axis). Three horizontal reference lines are superimposed on the scatterplot: one line at the mean difference between physician- and patient-derived joint counts, along with lines to mark the upper and lower limits of the 95% CI. Lastly, multiple linear regression analysis was used to explore factors potentially predictive of agreement at 6 months. These factors included level of disease activity at 6 months as well as level of formal education, age, disease duration and sex.

Results

Baseline patient characteristics

One hundred and six patients with RA were referred in the study period and 101 patients met the inclusion criteria and were randomly allocated, resulting in 51 patients in the intervention arm with self-assessment assisted by education and US feedback and 50 patients in the standard care arm. Overall, 91% of the patients (48 in the EUS arm and 44 in the standard care arm) completed the study (Fig. 2). Using the last observation carried forward method for missing results at the end of the study, 95 patients were included in the analysis (51 in the EUS arm and 44 in the standard care arm).

There were no clinically important group differences at baseline (Table 1). Patients had a mean age of 54.1 years (s.d. 12.8), the majority were women [n = 77 (81%)], 73 (78%) were positive for either RF or anti-CCP, 71 (77%) were Chinese and 80 (86%) were never smokers.
The mean disease duration was 6.8 years (S.D. 7.8). Sixty-two patients (67%) had an education level of at least secondary school and above. Patients had moderate disease activity at baseline, with a mean physician DAS28 of 3.4 (S.D. 1.2). The patient-derived DAS28 was higher, at 3.6 (S.D. 1.2), due to higher patient-reported TJCIs. On the other hand, physician SJCs were higher than patient SJCs. Only 26% of patients were in DAS28 remission.

At baseline, 81% of patients were on MTX, with 43% on low-dose steroids. More than 70% were on combination DMARD therapy and one patient was on a biologic DMARD.

Level of agreement between patient- and physician-derived swollen joints (clinical synovitis)

The proportion of patients that achieved a swollen joint PABAK >0.6 at 6 months was significantly higher for the education arm (from 84% at baseline to 98% by 6 months) compared with the standard arm (from 90% at baseline to 85% by 6 months) (P = 0.024). For swollen joints, the mean PABAK between patient and physician increased from 0.84 (S.D. 0.23) to 0.87 (S.D. 0.14) from baseline to 6 months in the education arm, compared with a reduction in the mean PABAK from 0.89 (S.D. 0.19) to 0.83 (S.D. 0.37) for the standard arm. When adjusted for the PABAK at baseline, there was no significant difference between the groups at 6 months [β = −0.108 (95% CI −0.177, 0.058), P = 0.32].

**Agreement of joint counts between patients and physician at the patient level**

Agreement of joint counts between patients and physician using the Bland and Altman limits of agreement confirmed the previous results (see supplementary Table S1, available at *Rheumatology* Online, and Fig. 3). The SJC difference between physician and patients at 6 months

---

**Fig. 1 PDUS synovitis semi-quantitative grading**

Grade 0 = absence, no synovial flow;

Grade 1 = mild, single vessel signals

Grade 2 = moderate, confluent signals in less than 50% of the synovial area

Grade 3 = marked, signals in 50% or more of the synovial area

Reproduced with the permission of Hammer et al. and BMJ Publishing Group Ltd, from Hammer et al. [23].
improved in the EUS arm, with the 95% limits of agreement decreasing from \((-4, 6)\) to \((-3, 4)\) compared with worsening in the standard arm from \((-5, 5)\) at baseline to \((-10, 5)\) at 6 months. Large differences remained in the TJC between patients and physician in both groups at 6 months. In particular, the difference in the TJC worsened in the standard care group at 6 months.

Feasibility of self-assessment of clinical synovitis using US feedback
Self-assessment education using US feedback was well accepted by the patients randomized to the education arm: 94% felt it was useful and 72% admitted it made them more aware of their disease activity. In addition, 66% felt that self-assessment of joints was useful in monitoring disease activity.

Agreement in DAS28 score change between patients and physician
A scatterplot of the patient vs physician DAS28 score change from baseline to 6 months (Fig. 4) showed that there was general agreement in the direction of change (i.e. improvement or worsening of disease activity). It is important to note that worsening in the DAS28 (i.e. an

*Had at least one assessment post-baseline to enable use of the last observation carried forward method, therefore 51 patients in the active arm and 44 patients in the standard care arm were included in the analysis.
**TABLE 1** Demographic and baseline characteristics of patients included in the main outcome analysis* (n = 95)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total (n = 95)</th>
<th>Education (EUS) arm (n = 51)</th>
<th>Standard care arm (n = 44)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (s.d.), years</td>
<td>54.1 (12.8)</td>
<td>54.6 (13.7)</td>
<td>53.7 (12.1)</td>
</tr>
<tr>
<td>Female, n (%)</td>
<td>77 (81)</td>
<td>38 (75)</td>
<td>39 (89)</td>
</tr>
<tr>
<td>Disease duration, mean (s.d.), years</td>
<td>6.8 (7.8)</td>
<td>7.0 (8.1)</td>
<td>6.5 (6.0)</td>
</tr>
<tr>
<td>Positive for RF or anti-CCP, n (%)</td>
<td>73 (78)</td>
<td>41 (80)</td>
<td>32 (74)</td>
</tr>
<tr>
<td>Non-smokers, n (%)</td>
<td>80 (86)</td>
<td>41 (82)</td>
<td>39 (91)</td>
</tr>
<tr>
<td>Chinese, n (%)</td>
<td>71 (77)</td>
<td>39 (77)</td>
<td>32 (73)</td>
</tr>
<tr>
<td>Education level secondary school and above, n (%)</td>
<td>62 (67)</td>
<td>34 (68)</td>
<td>28 (65)</td>
</tr>
<tr>
<td>Physician TJC, mean (s.d.)</td>
<td>2.1 (3.4)</td>
<td>2.0 (2.6)</td>
<td>2.3 (4.2)</td>
</tr>
<tr>
<td>Physician SJC, mean (s.d.)</td>
<td>1.5 (2.3)</td>
<td>1.8 (2.6)</td>
<td>1.2 (1.9)</td>
</tr>
<tr>
<td>No swollen joints, n (%)</td>
<td>47 (50)</td>
<td>22 (43)</td>
<td>25 (57)</td>
</tr>
<tr>
<td>Patient TJC, mean (s.d.)</td>
<td>4.0 (5.5)</td>
<td>3.6 (4.6)</td>
<td>4.4 (6.4)</td>
</tr>
<tr>
<td>Patient SJC, mean (s.d.)</td>
<td>1.2 (2.4)</td>
<td>1.0 (2.2)</td>
<td>1.3 (2.6)</td>
</tr>
<tr>
<td>PDUS score, mean (s.d.)</td>
<td>2.9 (3.8)</td>
<td>2.8 (3.9)</td>
<td>2.9 (3.7)</td>
</tr>
<tr>
<td>ESR, mean (s.d.), mm/h</td>
<td>28.6 (18.5)</td>
<td>27.2 (16.8)</td>
<td>30.2 (20.4)</td>
</tr>
<tr>
<td>Patient global assessment, mean (s.d.), VAS mm</td>
<td>27.4 (22.6)</td>
<td>27.5 (21.0)</td>
<td>27.4 (24.6)</td>
</tr>
<tr>
<td>Physician DAS28, mean (s.d.)</td>
<td>3.4 (1.2)</td>
<td>3.4 (1.1)</td>
<td>3.3 (1.2)</td>
</tr>
<tr>
<td>DAS28 remission, n (%)</td>
<td>25 (26)</td>
<td>13 (26)</td>
<td>12 (27)</td>
</tr>
<tr>
<td>Patient DAS28, mean (s.d.)</td>
<td>3.6 (1.2)</td>
<td>3.5 (1.2)</td>
<td>3.6 (1.4)</td>
</tr>
<tr>
<td>Modified HAQ, mean (s.d.)</td>
<td>0.16 (0.34)</td>
<td>0.16 (0.27)</td>
<td>0.16 (0.41)</td>
</tr>
<tr>
<td>SF-12 physical, mean (s.d.)</td>
<td>42.6 (9.4)</td>
<td>41.9 (9.0)</td>
<td>43.4 (9.8)</td>
</tr>
<tr>
<td>SF-12 mental, mean (s.d.)</td>
<td>47.8 (9.5)</td>
<td>48.3 (9.8)</td>
<td>47.2 (10.1)</td>
</tr>
</tbody>
</table>

*There are no significant differences between the two groups. DAS28: 28-joint DAS; PDUS: power Doppler US; SF-12: 12-item Short Form Health Survey; SJC: swollen joint count; TJC: tender joint count; VAS: visual analogue scale.

**Fig. 3** Bland and Altman plots of swollen joint count agreement between physician and patients

SJC: swollen joint count.
There were a number of potential limitations to this study. First, physician assessment rather than US was used as the patient’s reference standard for detecting swollen joint agreement, with US used purely as a training tool for patients. This is more reflective of real-life clinical practice, where decisions are based on physician assessment, with bedside US used only in situations where examination is in doubt. Second, patients in the standard care arm were not entirely blinded to the US findings during scanning. However, this is unlikely to have had a major impact on the self-assessment findings, as patients had already assessed their own joints prior to US scanning and there was no direct training or communication of the findings to these patients. If this misclassification bias was significant, then it would be less likely that a significant difference would be detected between the two groups. It was also impossible to know whether the self-assessment teaching or the US feedback alone contributed to the improvement in agreement. Lastly, these results were observed in a group of patients with established RA who were considered to have relatively stable disease.

Recent evidence has provided support for using US feedback to improve clinical examination with increased concordance seen between physicians in the detection of clinical synovitis (joint swelling) after feedback of results with US, in particular for joints with PDUS signals [16]. Our study was the first to evaluate this modality in assisting patients to self-assess their own joints for clinical synovitis. It is important to evaluate whether continued self-assessment of disease activity improves treatment adherence in this group of patients.

It has been demonstrated that training patients in joint count self-assessment can improve SJC reliability [10, 12]. However, those studies limited training to a short demonstration by the physician or metrologist without further practice, with evaluation made only 3 months after training. Our present study evaluated training in swollen joint detection (compared with the physician) over a longer period of follow-up of 6 months, with the additional use of US as a tool to improve joint assessment. In addition, unlike the previous studies, agreement was evaluated at both the joint and patient level in the same cohort. It was interesting to observe deterioration in the patient–physician swollen joint agreement in the standard care group over 6 months. This explained the statistical significance demonstrated for the primary outcome rather than the significant improvements in agreement in the intervention arm. From the Bland and Altman plots evaluating agreement at the patient level, there were clearly two outliers in the standard care group at 6 months that contributed to the poor agreement, and even when they were removed from the analysis, the results remained the same. Despite this, the Bland and Altman plots for the intervention group clearly demonstrate an improvement in SJC agreement.

Performance bias by the physician was minimized mainly because patients and the physician performed the joint counts independently, that is, patients performed the joint count before consultation with the physician and the physician did not have access to the results of
the patient joint counts during the consultation. It is difficult to interpret why patients in the control arm started off having very good agreement. One potential explanation was the fact there were a large proportion of patients with limited disease and no swollen joints, which would have made the agreement between patients and physician better. Secondly, the treat-to-target approach may have been used less with US and may have contributed to higher disease activity in the control group at the end of the study and therefore, as there were more swollen joints clinically, the agreement got worse. In order to further evaluate the effects of active intervention, the study ideally should have a crossover, with the control group receiving the active intervention for 6 months and the active group receiving no further intervention. This should be able to assess whether the agreement improves and also whether there is a carry-over effect with the intervention.

Level of disease activity remains a significant predictive factor for swollen joint agreement. This has important clinical implications as self-assessment of joints for clinical synovitis is most likely to be useful in patients with relatively stable disease and low disease activity. The good agreement in the change in DAS28 score between patients and physician, in particular, worsening in disease activity, suggests that patients can detect flares when results are incorporated into composite disease indices. This has been supported by recent evidence that patients are able to assist with the detection of flares in their disease by using self-assessed joint counts between clinic visits, which are most reliable when incorporated into composite indices [28].

Feasibility and issues of training for patients are potential barriers to using self-assessment. A nurse-led programme of patient self-assessment of disease activity in France demonstrated that patients can be compliant with home self-assessment of joints, with 89% completing their assessments during the 6 month study [7]. There was also a positive effect of self-assessment, with a change in DMARD therapy that was statistically significant in the active group [7]. However, the feasibility and compliance of self-assessment at home in multi-ethnic Asian cohorts, where patients are reluctant to take an active role in the management of their chronic illnesses, should be evaluated in the future.

Although we observed a trend towards a higher proportion of DAS28 remission in the intervention arm, the study was not designed or powered to evaluate disease outcomes between the two groups. Moreover, treatment adjustments were not according to a set protocol and were at the discretion of the rheumatologist seeing the patient, but the US results could be used if there was doubt regarding disease activity by clinical assessment alone. Hence it is difficult to determine whether improvement in disease activity was due to the intervention alone or to other factors. We suspect an improved acceptance of treatment changes by patients, and possibly increased use of the US results for treatment decisions, may have played a large part in better disease control in the active arm.

In conclusion, US can potentially be useful as a feedback training tool to educate patients and improve their assessment of clinical synovitis. Future evaluation is warranted to determine whether the skills are maintained and whether ongoing self-assessment of disease activity improves overall RA management through treatment adherence and acceptance of treatment escalation strategies.

**Funding:** No specific funding was received from any funding bodies in the public, commercial or not-for-profit sectors to carry out the work described in this article.

**Disclosure statement:** The authors have declared no conflicts of interest.

## Supplementary data

Supplementary data are available at *Rheumatology* Online.
References

3. Pincus T, Segurado OG. Most visits of most patients with rheumatoid arthritis to most rheumatologists do not include a formal quantitative joint count. Ann Rheum Dis 2006;65:820–2.