Clinical miscount of involved joints denotes the need for ultrasound complementation in usual practice for patients with rheumatoid arthritis

G. Murayama, M. Ogasawara, T. Nemoto, Y. Yamada, S. Ando, K. Minowa, T. Kon, K. Tada, M. Matsushita, K. Yamaji, N. Tamura, Y. Takasaki

Department of Internal Medicine and Rheumatology, Juntendo University School of Medicine, Tokyo, Japan

Abstract

Objective

Ultrasound (US) examination can visualise and clarify involved joints anatomically in patients with rheumatoid arthritis (RA), and it enables physicians to verify the accuracy of clinical assessments of involved joints. Here, we studied the practical "miscount"– calculated by subtracting US-determined involved joint count from clinically determined involved joint count – and analysed possible contributing factors for increased miscount.

Methods

The study population consisted of 137 patients with RA. Physical joint examination was performed by 3 assessors with different levels of experience in rheumatology, followed by US joint examination. Clinical and US examinations were performed on 28 joints (proximal interphalangeal, metacarpophalangeal, wrist, elbow, shoulder, and knee on both sides). Miscount was calculated for all patients, and multivariate analysis was conducted on possible contributing factors for miscount, including age, sex, body mass index, disease duration, Steinbrocker stage, erythrocyte sedimentation rate (ESR), C-reactive protein level, patient global assessment (GA), evaluator GA, matrix metalloproteinase-3 level, and power Doppler (PD) score.

Results

A high variability in concordance rate among the joint sites was observed among the 3 assessors. The average miscount was 1.07 (SD, 5.19; range, 18 to -11). ESR and patient GA were determined as significant contributing factors for false-positive miscount, whereas PD score and age were significant factors for false-negative miscount.

Conclusion

In addition to the condition of the involved joint distribution and the assessor's clinical examination skills, the patients' background can also lead to increased miscount. Assessors should be blinded to patients' background information, and US complementation should be included in usual clinical joint examinations.

Key words

rheumatoid arthritis, joint count, ultrasound

*Go Murayama, MD**
*Michihiro Ogasawara, MD, PhD**
Takuya Nemoto, MD
Yusuke Yamada, MD
Seiichiro Ando, MD, PhD
Kentaro Minowa, MD, PhD
Takayuki Kon, MD, PhD
Kurisu Tada, MD, PhD
Masakazu Matsushita, MD, PhD
Ken Yamaji, MD, PhD
Naoto Tamura, MD, PhD
Yoshinari Takasaki, MD, PhD

**These authors made an equal contribution to this study.*

Please address correspondence to:
Dr Michihiro Ogasawara,
Department of Internal Medicine and Rheumatology,
Juntendo University School of Medicine,
2-1-1 Hongo, Bunkyo-ku,
Tokyo 113-8421, Japan.
E-mail: miogasaw@juntendo.ac.jp

Received on October 17, 2012; accepted in revised form on January 4, 2013.

Competing interests: none declared.

Introduction

In clinical examination of rheumatoid arthritis (RA), the count and degree of swollen and/or tender joints are the most important basic clinical features, and should be assessed carefully. These measures are included in the new classification criteria (1); in such composite measures as the disease activity score (DAS), simplified disease activity index (SDAI), and clinical disease activity index (CDAI); in the Boolean remission definition (2-5); and in the recommendations for achieving optimal therapeutic outcomes in RA, which are referred to as "treatment to target" (6), which is important for early diagnosis, disease activity evaluation, treatment decision, therapy evaluation, and prognosis for bone destruction (7, 8).

Inter- and intraobserver variability in the clinical assessment of joint swelling and/or tenderness have been reported (9). However, clinical joint assessment methods still are being used without any modification, which could lead to incorrect diagnosis or over- or underestimation of disease activity.

In recent years, the utilisation of ultrasound (US) has become common in rheumatology, which allows easier visualisation and clarification of involved joints anatomically, and enables physicians to determine the accuracy of clinical judgment of swelling and/or inflammation (10-17). There has been no study showing practical "miscount"– calculated by subtracting US-determined involved joint count from clinically determined involved joint count. Furthermore, no study has conducted multivariate analysis to determine the possible contributing factors for increased miscount.

In this study, we investigated the practical clinical joint examination skills of 3 assessors with different levels of experience in rheumatology. We determined the practical miscount, and analysed the factors possibly contributing to increased miscount.

Patients and methods

The study population consisted of 137 patients with RA (108 women, 29 men) between the ages of 22 and 85 years, who presented to Juntendo University Hospital in Tokyo, Japan (Table I). Diagnosis of RA was based on the criteria established in 1987 (18). First, a physical examination was performed to assess the presence of joint swelling and/or tenderness; the examinations were performed by 1 of 3 rheumatologists (assessor no. 1, GM; assessor no. 2, TN; assessor no. 3, MO) who were blinded to the US findings. These assessors had different levels of experience in rheumatology: assessor no. 1, 2 years; assessor no. 2, 8 years; and assessor no. 3, 16 years and Japan College of Rheumatology–certified. The physical examinations were followed by a US examination (ProSound Alpha7 with UST-5411, 10-13 MHz transducer; Hitachi Aloka Medical, Ltd., Tokyo, Japan). Clinical and US examinations were performed on 28 joints (proximal interphalangeal [PIP], metacarpophalangeal [MCP], wrist, elbow, shoulder, and knee on both sides). The US examination was carried out within 30 minutes of each clinical evaluation by 1 of 3 rheumatologists experienced with US (assessor no. 1, GM; assessor no. 2, TN; assessor no. 3, MO) in a darkened room with a total time of less than 60 minutes, and the results were recorded after achieving a consensus between the assessors.

Synovial effusion and/or hypertrophy were identified as abnormal hypoechoic material within joint recesses, tendon sheaths, or bursa, and was graded on a semiquantitative gray scale (GS) from 0 to 3 (0, absence; 1, mild; 2, moderate; and 3, marked) (19-22). Synovial blood flow was evaluated by power Doppler (PD) in each of the intra-articular and periarticular synovial sites. PD parameters were adjusted to the lowest permissible pulse repetition frequency to maximise sensitivity, while color gain was set just below the level at which color noise appeared underlying bone. Intra-articular PD signals were graded on a semiquantitative scale from 0 to 3 (0: absence, no synovial flow; 1: mild, ≤3 isolated signals; 2: moderate, >3 isolated signals or confluent signal in less than half of the synovial area; and 3: marked, signals in more than half of the synovial area) (19-22). PD signals

in periarticular synovial sites, tendon sheaths, and bursa were also graded on a semiquantitative scale from 0 to 3 (0: absence, no synovial flow; 1: mild; 2: moderate; and 3: marked) (19-22). In each patient, the GS score represented the sum of the GS grade for all 28 joint sites, whereas the PD score represented the sum of the PD grade for all 28 joint sites. Since the maximum GS/PD score for each joint site was 3, the maximum possible score for each patient was 84. For each joint site, ≥GS 2 or ≥PD 1 was defined as a US-determined involved joint site.

We calculated each of the 3 assessors' sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and kappa coefficient of clinical examination for US-determined involved joint sites. We also calculated miscount for each assessor by subtracting the US-determined involved joint count from the clinically determined involved joint count, and then compared the miscount among the 3 assessors.

Finally, the possible contributing factors for miscount were studied with multivariate analysis. Clinically obtained information, such as age, sex, body mass index (BMI), disease duration, Steinbrocker stage, erythrocyte sedimentation rate (ESR), C-reactive protein (CRP) level, patient global assessment (GA), evaluator GA, matrix metalloproteinase-3 (MMP-3) level, and PD score were used as candidates for contributing factors. Clinical joint counts assessed as swelling/tender joint counts were excluded in advance since those were included in the formula of miscount. Clinical joint count-related composite measures were also excluded. An analysis of contributing factors for miscount was performed separately for 2 groups: when miscount was ≥0 (false-positive miscount; clinical joint count ≥ US joint count) and when miscount was ≤0 (false-negative miscount; clinical joint count ≤ US joint count).

The study was conducted in accordance with the Declaration of Helsinki and was approved by the Institutional Review Board at Juntendo University. Informed consent was obtained from all the study participants.

Table I. Patient characteristics.

	All patients
n	137
Female, n (%)	108 (78.8)
Age, y*	54 (22–85)
Body weight, kg*	51 (36–91)
Body mass index, kg/m²*	20 (14–38)
Disease duration, months*	63.8 (3–619)
Steinbrocker's STAGE III/IV, n (%)	62 (45.3)
MTX use, n (%)	95 (69.3)
Biologics use, n (%)	36 (26.3)
Steroid use, n (%)	48 (35.0)
RF positive, n (%)	78 (56.9)
MMP-3 positive, n (%)	90 (65.7)
ESR, mm/h*	21 (2–155)
CRP, mg/dL*	0.2 (0–19.7)
Swollen joint count of 28 joints*	3 (0–24)
Tender joint count of 28 joints*	1 (0–26)
Patient GA*	4 (0–10)
Evaluator GA*	4 (0–9)
DAS*	3.5 (0.4–8.2)
CDAI*	13 (0–56)
SDAI*	14 (0–63.5)
DAS remission (<2.6), n (%)	37 (27)
CDAI remission (≤2.8), n (%)	22 (16)
SDAI remission (≤3.3), n (%)	25 (18.2)
Boolean remission, n (%)	16 (11.6)

*Median (range).
CDAI: clinical disease activity index; CRP: C-reactive protein; DAS: disease activity score in 28 joints; ESR: erythrocyte sedimentation rate; GA: global assessment in cm (0-10); MMP-3: matrix metalloproteinase-3; MTX: methotrexate; RF: rheumatoid factor; SDAI: simplified disease activity index.

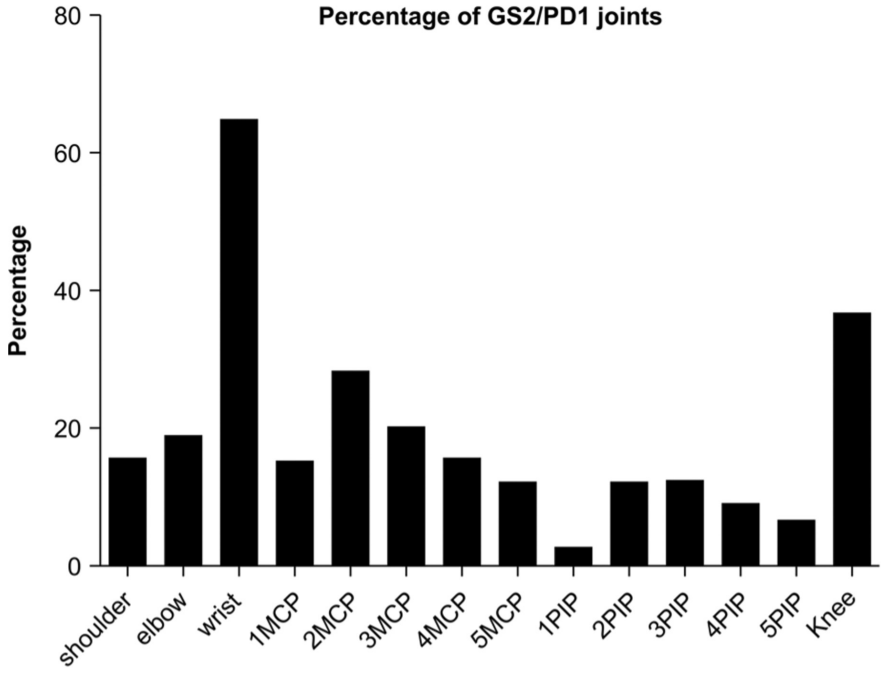

Fig. 1. The percentage of US-determined involved joints (≥GS2/PD1) for each joint site.

Results

Patient demographics

Average disease duration was 93.8 (SD, 96.8) months. Of the 137 patients included in this study, 45.3% were classified as having Steinbrocker stage III or IV, and rheumatoid factor was present in 78 patients (56.9%). MTX was prescribed to 69.3% of the patients, biologics to 26.3%, and glu-

Table II. Sensitivity, specificity, and predictive values for US-determined involved joints.

Assessor no. 1					Assessor no. 1				
≥GS2/PD1	Sensitivity	Specificity	PPV	NPV	≥GS1/PD1	Sensitivity	Specificity	PPV	NPV
Shoulder	0.50	0.93	0.67	0.88	Shoulder	0.29	0.92	0.67	0.69
Elbow	0.00	0.85	0.00	0.96	Elbow	0.10	0.82	0.25	0.61
Wrist	0.54	0.90	0.90	0.55	Wrist	0.51	0.92	0.93	0.47
1MCP	0.00	0.96	0.00	0.88	1MCP	0.06	0.97	0.33	0.80
2MCP	0.50	0.83	0.50	0.83	2MCP	0.42	0.83	0.55	0.74
3MCP	0.33	0.84	0.21	0.91	3MCP	0.29	0.85	0.36	0.81
4MCP	0.25	0.97	0.33	0.96	4MCP	0.08	0.97	0.33	0.85
5MCP	0.00	0.97	0.00	0.92	5MCP	0.09	0.99	0.50	0.87
1PIP	0.50	1.00	1.00	0.99	1PIP	0.33	1.00	1.00	0.97
2PIP	0.71	0.81	0.29	0.96	2PIP	0.75	0.82	0.35	0.96
3PIP	0.75	0.82	0.20	0.98	3PIP	0.75	0.82	0.20	0.98
4PIP	0.75	0.94	0.43	0.98	4PIP	0.50	0.94	0.43	0.95
5PIP	0.67	0.96	0.40	0.98	5PIP	0.67	0.96	0.40	0.98
Knee	0.50	0.85	0.78	0.61	Knee	0.36	1.00	1.00	0.11
Average	**0.43**	**0.90**	**0.41**	**0.88**	Average	**0.37**	**0.91**	**0.52**	**0.77**
Assessor no. 2					Assessor no. 2				
≥GS2/PD1	Sensitivity	Specificity	PPV	NPV	≥GS1/PD1	Sensitivity	Specificity	PPV	NPV
Shoulder	0.50	0.97	0.50	0.97	Shoulder	0.17	0.97	0.50	0.86
Elbow	0.38	0.87	0.38	0.87	Elbow	0.23	0.88	0.63	0.56
Wrist	0.70	0.94	0.95	0.67	Wrist	0.68	0.97	0.97	0.62
1MCP	0.00	0.91	0.00	0.95	1MCP	0.00	0.91	0.00	0.89
2MCP	0.68	0.89	0.65	0.90	2MCP	0.53	0.94	0.85	0.76
3MCP	0.67	0.86	0.44	0.94	3MCP	0.46	0.91	0.72	0.77
4MCP	0.63	0.87	0.33	0.96	4MCP	0.47	0.91	0.60	0.85
5MCP	0.60	0.92	0.33	0.97	5MCP	0.40	0.93	0.44	0.92
1PIP	0.00	0.96	0.00	0.99	1PIP	0.00	0.96	0.00	0.98
2PIP	0.73	0.89	0.50	0.95	2PIP	0.73	0.89	0.50	0.95
3PIP	0.50	0.84	0.25	0.94	3PIP	0.44	0.84	0.25	0.92
4PIP	1.00	0.88	0.44	1.00	4PIP	0.89	0.89	0.50	0.98
5PIP	1.00	0.98	0.50	1.00	5PIP	1.00	0.98	0.50	1.00
Knee	0.54	0.96	0.88	0.79	Knee	0.35	0.94	0.88	0.55
Average	**0.57**	**0.91**	**0.44**	**0.92**	Average	**0.45**	**0.92**	**0.52**	**0.83**
Assessor no. 3					Assessor no. 3				
≥GS2/PD1	Sensitivity	Specificity	PPV	NPV	≥GS1/PD1	Sensitivity	Specificity	PPV	NPV
Shoulder	0.67	0.84	0.55	0.90	Shoulder	0.60	0.83	0.55	0.86
Elbow	0.56	0.83	0.60	0.81	Elbow	0.48	0.86	0.73	0.67
Wrist	0.68	0.72	0.85	0.50	Wrist	0.66	0.71	0.86	0.42
1MCP	0.33	0.94	0.64	0.82	1MCP	0.22	0.93	0.64	0.68
2MCP	0.62	0.73	0.51	0.81	2MCP	0.45	0.72	0.73	0.43
3MCP	0.59	0.75	0.49	0.82	3MCP	0.41	0.74	0.71	0.45
4MCP	0.43	0.89	0.57	0.82	4MCP	0.28	0.89	0.74	0.53
5MCP	0.57	0.92	0.60	0.91	5MCP	0.30	0.91	0.70	0.67
1PIP	0.50	0.91	0.11	0.99	1PIP	0.25	0.91	0.22	0.93
2PIP	0.55	0.76	0.24	0.92	2PIP	0.53	0.77	0.32	0.89
3PIP	0.88	0.72	0.42	0.96	3PIP	0.86	0.76	0.53	0.95
4PIP	0.70	0.81	0.32	0.96	4PIP	0.58	0.81	0.32	0.93
5PIP	0.58	0.84	0.37	0.93	5PIP	0.53	0.85	0.42	0.90
Knee	0.71	0.74	0.53	0.87	Knee	0.48	0.81	0.84	0.43
Average	**0.60**	**0.81**	**0.48**	**0.86**	Average	**0.47**	**0.82**	**0.59**	**0.69**

GS: grey scale; MCP: metacarpophalangeal; NPV: negative predictive value; PD: power Doppler; PIP: proximal interphalangeal; PPV: positive predictive value.

cocorticoids to 35%. The percentage of patients in clinical remission differed, based on the index and definition: DAS (37/137; 27%), CDAI (22/137; 16%), SDAI (25/137; 18.2%), or Boolean (16/137; 11.6%) (Table I). Some of the patients in clinical remission still had subclinical synovitis (defined here as PD score >0): 57% (21/37; average PD score, 4.2±3.8), 45% (10/22, average PD score, 3.2±2.3), 52% (13/25; average PD score, 3.0±2.0), and 50% (8/16; average PD score, 2.6±1.5), respectively.

The percentage of US-determined involved joints for each joint site are shown in Fig. 1. US-determined involved joints were particularly ob-

Table III. Kappa coefficients.

	Assessor no. 1		Assessor no. 2		Assessor no. 3	
	≥GS2/PD1	≥GS1/PD1	≥GS2/PD1	≥GS1/PD1	≥GS2/PD1	≥GS1/PD1
Shoulder	0.36	0.22	0.31	0.15	0.32	0.30
Elbow	-0.04	0.02	0.31	0.22	0.38	0.35
Wrist	0.39	0.34	0.57	0.55	0.36	0.30
1MCP	-0.06	0.04	-0.07	-0.10	0.34	0.18
2MCP	0.33	0.27	0.56	0.51	0.32	0.13
3MCP	0.14	0.16	0.44	0.41	0.30	0.12
4MCP	0.25	0.08	0.35	0.41	0.33	0.15
5MCP	-0.04	0.12	0.38	0.35	0.50	0.25
1PIP	0.66	0.49	-0.02	-0.03	0.12	0.13
2PIP	0.33	0.40	0.49	0.54	0.13	0.16
3PIP	0.26	0.26	0.22	0.19	0.35	0.42
4PIP	0.51	0.41	0.53	0.55	0.27	0.23
5PIP	0.47	0.47	0.56	0.74	0.29	0.30
Knee	0.37	0.27	0.45	0.29	0.23	0.21
Average	**0.28**	**0.25**	**0.36**	**0.34**	**0.30**	**0.23**

GS: grey scale; MCP: metacarpophalangeal; PD: power Doppler; PIP: proximal interphalangeal.

served in the wrist and knee (64.4% and 36.3%, respectively).

Sensitivity, specificity, and kappa coefficients for each joint site (Tables II and III; Fig. 2abc)

Regarding sensitivity, assessor no. 3 showed the highest average sensitivity among the 3 assessors, and did not show extremely low sensitivity for any of the joint sites. Depending on the joint site, the sensitivity was highly different among the assessors, and they showed especially varied sensitivity for several joint sites (Fig. 2a). Except for assessor no. 3, the other 2 assessors' sensitivity was 0 in the elbow, 5MCP, 1MCP, and 1PIP. Thus, they might not have assessed swelling/tenderness in those joint sites, or they might have misdiagnosed involved joints as normal joints.

Regarding specificity, for the wrist, 2MCP, 3MCP, 2PIP, 3PIP, and knee, assessor no. 3 showed lower specificity than the other 2 assessors (Fig. 2b). This indicates that the assessor tended to estimate those normal joints as involved joints.

Concordance rate (the kappa coefficient) between clinically determined involved joints and US-determined involved joints was calculated for each joint site. (Table III; Fig 2c). Assessor no. 2 showed the highest average kappa coefficient of the 3 assessors; however, for 1MCP and 1PIP, the kappa coefficient was extremely low. Assessor No. 1 showed the lowest average kappa coefficient, which was particularly low in the elbow, 1MCP, and 5MCP.

US revealed high variability in the kappa coefficient among joint sites; thus, clinical joint examinations show advantages and disadvantages, depending on the joint site. Therefore, it was suggested that the condition of involved joint distribution and the assessor's clinical examination skills could be related to the degree of miscount.

Miscount (Fig. 3ab)

Miscount was calculated for all study patients (Fig. 3a). The average miscount was 1.07 (SD, 5.19; range, 18 to -11); clinical joint count was slightly higher than US joint count.

Miscount was also compared among the 3 assessors (Fig. 3b): assessor no. 1, 0.08 (SD, 3.85); assessor no. 2, 1.15 (SD, 4.77); and assessor no. 3, 1.79 (SD, 6.13). Depending on the assessor, the distribution of miscount showed a different tendency (Fig. 3b); however, the average miscount was not significantly different among the 3 assessors (Kruskal-Wallis test, $p>0.05$). Therefore, it was suggested that variability in each assessor's joint examination skills, and not the difference in skills among the assessors, was a reason for miscount.

Contributing factors for miscount

To further investigate the reasons for miscount, we analysed the patients' background for possible contributing factors. As categorical variables, sex (Mann-Whitney U-test, $p=0.3735$; Fig. 4a) and Steinbrocker stage (Kruskal-Wallis test, $p=0.5679$; Fig. 4b) were analysed, and they showed no significant contributions to false-positive or false-negative miscount. Treatment difference (biologics, MTX, and/or steroid use) also did not show apparent contribution to miscount (data not shown). For continuous variables, we conducted paired correlation coefficient analysis between miscount and each possible contributing factor (Table IV). Only statistically significant factors were used in subsequent multiple correlation analyses. In the case of false-positive miscount, ESR (Fig. 4c) and patient GA (Fig. 4d) were determined as significant contributing factors. It was suggested that when patients showed high ESR and/or large patient GA, false-positive miscount tended to be estimated larger. In the case of false-negative miscount, PD score (Fig. 4e) and age (Fig. 4f) were determined as significant contributing factors. It was suggested that when patients showed high PD score and/or were older, false-negative miscount tend to be estimated larger. Therefore, it was suggested that not only the condition of involved joint distribution and the assessor's clinical examination skills but also patients' backgrounds could be a cause of increased miscount.

Discussion

Physical examination skills for US-determined involved joints were investigated in 3 rheumatologists with different levels of experience in rheumatology. Overall, the average concordance rate (kappa coefficient) was not much different; however, several joint sites showed high variability in the kappa coefficient among the 3 assessors. In particular, in assessors with less experience in rheumatology, there was a high variability in the kappa coefficient among joint sites, and an especially low kappa coefficient was observed in several joints, suggesting weak skills for those particular joint sites. It was thought that these miscounts must be observed usually, and so the condition of involved joint distribution and the

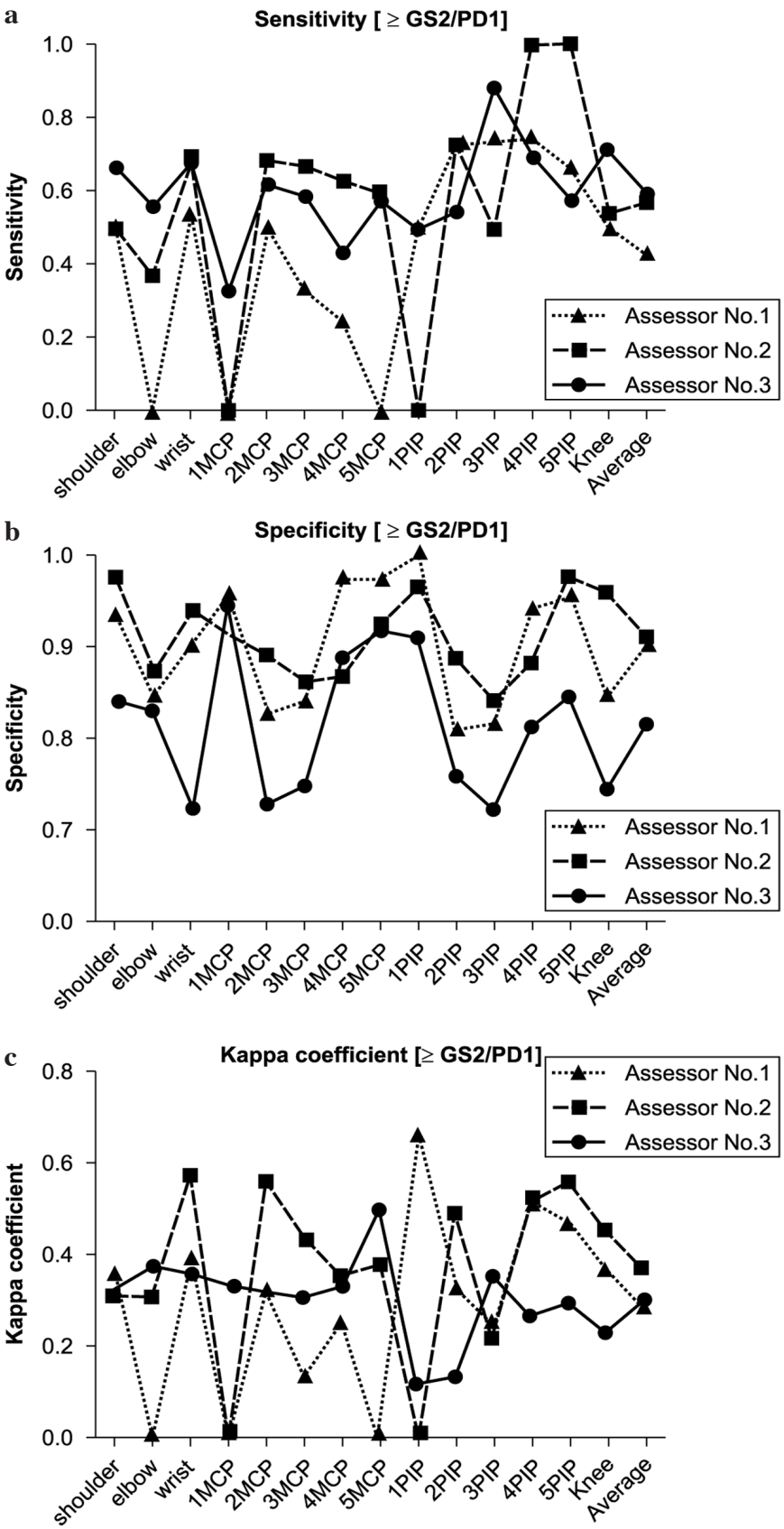

Fig. 2. a. Sensitivity of clinical examination for US-determined involved joint (≥GS2/PD1);
b. Specificity of clinical examination for US-determined involved joint (≥GS2/PD1);
c. Kappa coefficient between clinically determined involved joint and US-determined involved joint (≥GS2/PD1).

assessor's clinical examination skills could be related to the degree of miscount and wrong estimation of disease activity. In clinical practice, for correct diagnosis and estimation of RA disease activity, it is important to confirm the assessor's clinical joint examination skills and train assessors by utilising US (23).

In this study, we determined the contributing factors for miscount among the 3 assessors. ESR and patient GA were significantly related to false-positive miscount and were independent contributing factors. On the basis of these results, in clinical practice, we suggest that blood test results (including ESR) and patient complaints might lead to wrong estimation of involved joint count. Therefore, joint assessors should be blinded to patients' background information in usual practice. On the other hand, with respect to false-negative miscount, PD score and age were found to be independent contributing factors. In patients with multiple active synovitis, it was thought that assessors could not find involved joints fully due to low sensitivity, resulting in the tendency to underestimate the involved joint count. Regarding age, since it is thought that elderly patients with RA are more difficult to treat due to adverse effects and complications, most receive conservative treatment. Therefore, it is possible that these circumstances cause assessors to underestimate the involved joint count. These newly determined contributing factors equally indicate that clinical joint examination should be conducted with the physician blinded to patients' background information in order to avoid bias.

Recent advanced treatment of RA has emphasised the need for early diagnosis and more correct monitoring of disease activity (24, 25). Joint swelling and/or tenderness is the central surrogate marker in classification criteria and disease activity judgment. Moreover, joint swelling has been focused on recently because it was found to be a more important contributing factor to bone erosion than CRP level (7, 8). However, difference in clinical joint examination skills among the 3 assessors, particularly in several joints, and patients' back-

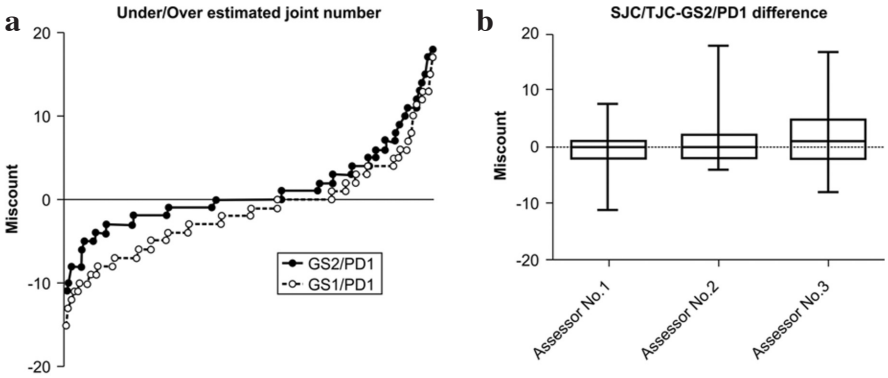

Fig. 3. a. Distribution of miscount calculated by subtracting US-determined involved joint count from clinically determined involved joint count.
b. Comparison of each assessor's miscount distribution.

Table IV. Analysis for contributing factors.

False positive miscount variable	*r*	Paired correlation *p*-value	Multiple correlation analysis *p*-value
ESR	0.41	0.0001	0.0099*
Patient GA	0.41	0.0001	0.0048**
Evaluator GA	0.34	0.002	0.130
CRP	0.32	0.003	0.884
MMP-3	0.26	0.024	0.875
PD score	0.19	0.09	–
Age	0.08	0.47	–
BMI	0.05	0.60	–
Disease duration	-0.03	0.80	–
False negative miscount variable	***r***	**Paired correlation *p*-value**	**Multiple correlation analysis *p*-value**
PD score	-0.44	< 0.0001	0.0035***
ESR	-0.29	0.01	0.702
Age	-0.29	0.01	0.0372*
MMP-3	-0.27	0.02	0.340
Disease duration	-0.26	0.06	–
Evaluator GA	-0.19	0.09	–
CRP	-0.15	0.20	–
Patient GA	-0.08	0.48	–
BMI	-0.05	0.60	–

*<0.05. **<0.01. ***<0.0001. Miscount = clinical joint count – ultrasound joint count. Positive miscount = clinical > ultrasound joint count. Negative miscount = clinical < ultrasound joint count.
BMI: body mass index; CRP: C-reactive protein; ESR: erythrocyte sedimentation rate; GA: global assessment in cm (0–10); MMP-3: matrix metalloproteinase-3; PD: power Doppler.

ground, such as ESR, patient GA, active synovitis, and age, were shown to increase miscount. As described here, multiple factors can affect the judgment of involved joint count; thus, it should not be used without correction in clinical practice. It is suggested that US complementation should be conducted regardless of the assessor's experience level, particularly to avoid overlooking subclinical synovitis.

This study has some problems and limitation. Our data were derived from a single centre, and the results of miscount were highly dependent on the 3 assessors' examination skills. However, since the 3 assessors are representative of physicians in Japan, who have received similar rheumatology training, our study results might be applicable to other centres. Although osteoarthritis, bone destruction, and joint deformity were considered as possible reasons for miscount, those data were not determined on radiograph or ultrasound for all clinically examined 28 joints, and, therefore, we could not estimate their contributions to miscount in this study. The degree of bone destruction and joint deformity was estimated with only Steinbrocker stage, and showed no relationship to miscount here; however, this cannot fully determine whether bone destruction and joint deformity have no relationship to miscount. Other possible contributing factors for miscount, such as local fat and edema, were difficult to determine and also were not assessed here.

In this study, we found that clinically determined joint counts are mostly incorrect by showing the discrepancy in joint counts between clinical and US examinations. Moreover, we found that patients' background, including ESR, patient GA, PD score, and age, could cause increased miscount, along with the condition of involved joint distribution and the assessor's clinical examination skills. These results show that (1) clinical joint examination is mostly incorrect, and examination skills should be confirmed and trained with the use of US; (2) US complementation is needed in usual clinical joint examinations; and (3) physicians should be blinded to patients' background information when clinical joint examination is conducted.

References

1. ALETAHA D, NEOGI T, SILMAN AJ *et al*.: Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum* 2010; 62: 2569-81.
2. VAN DER HEIJDE DM, VAN 'T HOF MA, VAN RIEL PL *et al*.: Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. *Ann Rheum Dis* 1990; 49: 916-20.
3. SMOLEN JS, BREEDVELD FC, SCHIFF MH *et al*.: A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology* (Oxford) 2003; 42: 244-57.
4. ALETAHA D, NELL VP, STAMM T *et al*.: Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. *Arthritis Res Ther* 2005; 7: R796-806.
5. FELSON DT, SMOLEN JS, WELLS G *et al*.: American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Arthritis Rheum* 2011; 63: 573-86.
6. SMOLEN JS, ALETAHA D, BIJLSMA JW *et al*.: Treating rheumatoid arthritis to target: rec-

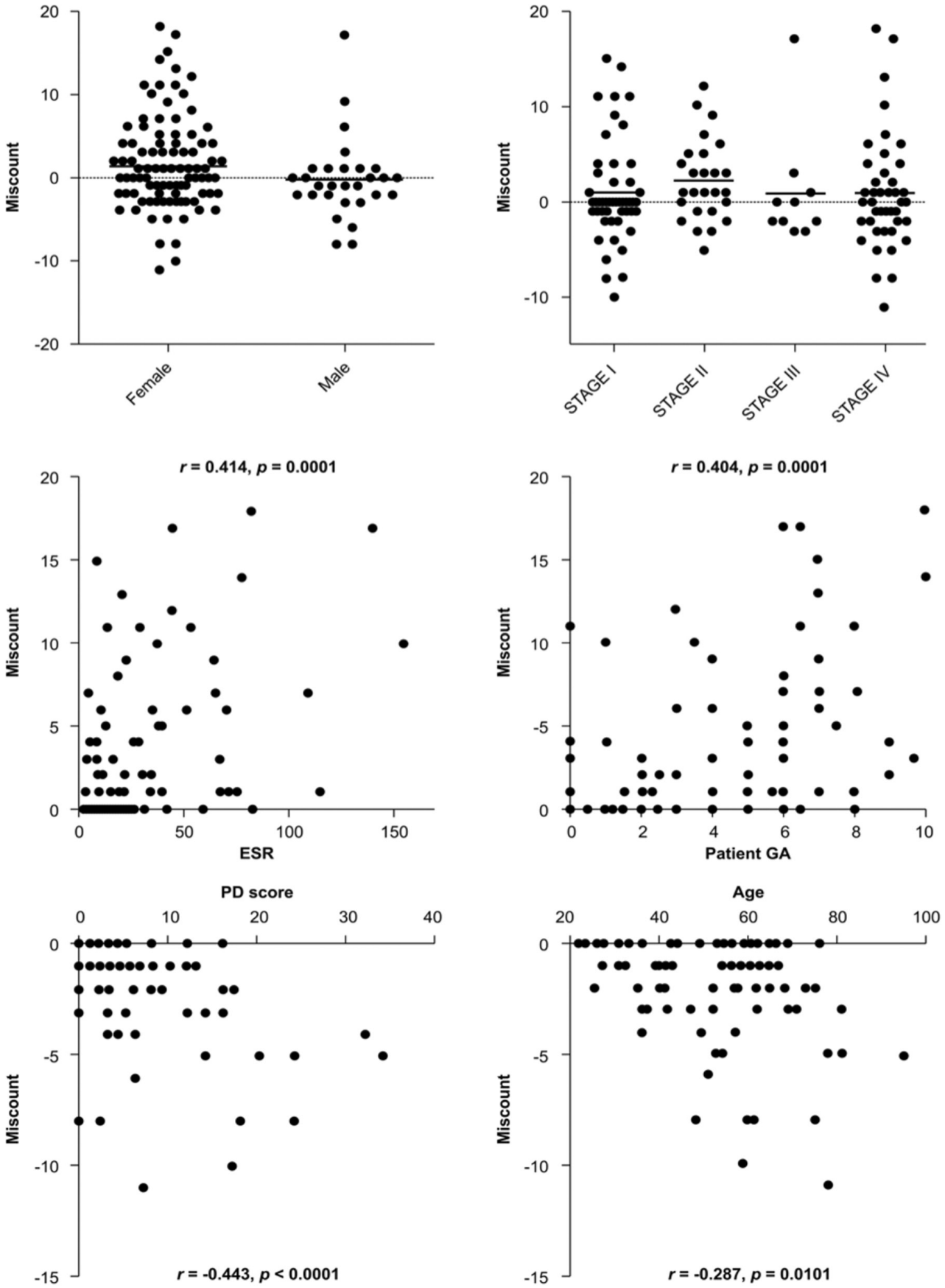

Fig. 4. Analysis of contributing factors. **a**. sex (Mann-Whitney U-test, p=.3735), **b**. Steinbrocker stage (Kruskal-Wallis test, p=.5679), **c**. ESR (r=0.41, p=0.0099), **d**. patient GA (r=0.41, p=0.0048), **e**. PD score (r=-0.44, p=0.0035), **f**. age (r=-0.29, p=0.037).

ommendations of an international task force. *Ann Rheum Dis* 2010; 69: 631-7.
7. ALETAHA D, ALASTI F, SMOLEN JS: Rheumatoid arthritis near remission: clinical rather than laboratory inflammation is associated with radiographic progression. *Ann Rheum Dis* 2011; 70: 1975-80.
8. ALETAHA D, SMOLEN JS: Joint damage in rheumatoid arthritis progresses in remission according to the Disease Activity Score in 28 joints and is driven by residual swollen joints. *Arthritis Rheum* 2011; 63: 3702-11.
9. WOOD L, PEAT G, WILKIE R, HAY E, THOMAS E, SIM J: A study of the noninstrumented physical examination of the knee found high observer variability. *J Clin Epidemiol* 2006; 59: 512-20.
10. FILER A, DE PABLO P, ALLEN G *et al*.: Utility of ultrasound joint counts in the prediction of rheumatoid arthritis in patients with very early synovitis. *Ann Rheum Dis* 2011; 70: 500-7.
11. NAREDO E, BONILLA G, GAMERO F, USON J, CARMONA L, LAFFON A: Assessment of inflammatory activity in rheumatoid arthritis: a comparative study of clinical evaluation with grey scale and power Doppler ultrasono-graphy. *Ann Rheum Dis* 2005; 64: 375-81.
12. LUUKKAINEN RK, SALTYSHEV M, KOSKI JM, HUHTALA HS: Relationship between clinically detected joint swelling and effusion diagnosed by ultrasonography in metatarsophalangeal and talocrural joints in patients with rheumatoid arthritis. *Clin Exp Rheumatol* 2003; 21: 632-4.
13. FILIPPOU G, CANTARINI L, BERTOLDI I, PICERNO V, FREDIANI B, GALEAZZI M: Ultrasonography vs. clinical examination in children with suspected arthritis. Does it make sense to use poliarticular ultrasonographic screening? *Clin Exp Rheumatol* 2011; 29: 345-50.
14. SCHMIDT WA, SCHICKE B, OSTENDORF B, SCHERER A, KRAUSE A, WALTHER M: Low-field MRI versus ultrasound: which is more sensitive in detecting inflammation and bone damage in MCP and MTP joints in mild or moderate rheumatoid arthritis? *Clin Exp Rheumatol* 2012 Oct 30.
15. BANDINELLI F, PRIGNANO F, BONCIANI D *et al*.: Ultrasound detects occult entheseal involvement in early psoriatic arthritis independently of clinical features and psoriasis severity. *Clin Exp Rheumatol* 2013; 31: 219-24
16. BROLL M, ALBRECHT K, TARNER I, MULLER-LADNER U, STRUNK J: Sensitivity and specificity of ultrasonography and low-field magnetic resonance imaging for diagnosing arthritis. *Clin Exp Rheumatol* 2012; 30: 543-7.
17. NARANJO A, JIMENEZ-NUNEZ FG, MEDINA-LUEZAS J *et al*.: Impact of the use of musculoskeletal ultrasound by rheumatologists in patients with shoulder and hand complaints compared with traditional clinical care. *Clin Exp Rheumatol* 2012; 30: 768-71.
18. ARNETT F, EDWORTHY S, BLOCH D *et al*.: The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988; 31: 315-24.
19. SZKUDLAREK M, COURT-PAYEN M, JACOBSEN S, KLARLUND M, THOMSEN HS, ØSTERGAARD M: Interobserver agreement in ultrasonography of the finger and toe joints in rheumatoid arthritis. *Arthritis Rheum* 2003; 48: 955-62.
20. DOUGADOS M, JOUSSE-JOULIN S, MISTRETTA F *et al*.: Evaluation of several ultrasonography scoring systems for synovitis and comparison to clinical examination: results from a prospective multicentre study of rheumatoid arthritis. *Ann Rheum Dis* 2010; 69: 828-33.
21. NAREDO E, COLLADO P, CRUZ A *et al*.: Longitudinal power Doppler ultrasonographic assessment of joint inflammatory activity in early rheumatoid arthritis: predictive value in disease activity and radiologic progression. *Arthritis Rheum* 2007; 57: 116-24.
22. NAREDO E, RODRÍGUEZ M, CAMPOS C *et al*.: Validity, reproducibility, and responsiveness of a twelve-joint simplified power doppler ultrasonographic assessment of joint inflammation in rheumatoid arthritis. *Arthritis Rheum* 2008; 59: 515-22.
23. OGASAWARA M, MURAYAMA G, YAMADA Y *et al*.: Autofeedback from ultrasound images provides rapid improvement in palpation skills for identifying joint swelling in rheumatoid arthritis. *J Rheumatol* 2012; 39: 1207-14.
24. HARAOUI B, SMOLEN JS, ALETAHA D *et al*.: Treating Rheumatoid Arthritis to Target: multinational recommendations assessment questionnaire. *Ann Rheum Dis* 2011; 70: 1999-2002.
25. SMOLEN JS, ALETAHA D: Monitoring rheumatoid arthritis. *Curr Opin Rheumatol* 2011; 23: 252-8.