TABLE 1 Clinical and laboratory characteristics of patients at baseline

	ADA	TCZ	*P*-value
Age, mean (range), years	53 (24–78)	56.4 (33–77)	0.516
Sex, female/male, *n*	12/1	18/1	
Duration of symptoms, median (IQR), months	62 (11–147)	142 (72–178)	0.156
ESR, median (IQR), mm/h	48 (34–54)	54 (34–64)	0.389
CRP, median (IQR), mg/dl	0.51 (0.09–0.89)	1.31 (0.24–3.03)	0.089
Swollen joint count, median (IQR)	3 (2–5)	5 (3–7)	0.179
Tender joint count, median (IQR)	5 (1–8)	4 (2–9)	0.984
Patient's global assessment by VAS, median (IQR)	50 (42–65)	67 (40–80)	0.544
Examiner's global assessment by VAS, median (IQR)	40 (40–50)	50 (33–70)	0.56
DAS28-ESR (S.D.)			
Baseline	5.03 (1.16)	5.28 (1.08)	0.575
Week 8	2.96 (0.86)	2.93 (0.81)	0.936
SDAI (S.D.)			
Baseline	21 (10.5)	24.7 (11.3)	0.275
Week 8	7.61 (5.48)	8.84 (4.31)	0.60
TGSS, median (IQR)			
Baseline	99.5 (73–116)	122.75 (98.75–160.75)	0.238
Week 50	108.5 (73–134.5)	125 (99.88–164.88)	0.271

Relationship between positive synovial vascularity and radiographic progression in finger joints

In the ADA group the mean and median of local synovial vascularity at baseline for the MCP and PIP joints were 197 and 0 (range 0–3053) and 218 and 0 (range 0–2414), respectively. In the TCZ group the mean and median of local synovial vascularity at baseline for the MCP and PIP joints were 416 and 0 (range 0–4686) and 167 and 0 (range 0–3195), respectively. Local synovial vascularity in both the ADA and TCZ groups decreased significantly from baseline to week 8 (ADA: MCP $P=0.0001$, PIP $P<0.0001$; TCZ: MCP $P=0.0002$, PIP $P=0.004$). We next categorized finger joints into four groups according to the occurrence of patterns of positive synovial vascularity: joints without synovial vascularity throughout the observational period [the negative (N) group], joints with positive synovial vascularity limited to the period from the baseline to week 8 [the therapeutic response (R) group], joints with intermittent occurrence of positive synovial vascularity in the observational period [the intermittently positive (IP) group] and joints with persistent positive synovial vascularity throughout the observational period [the persistently positive (PP) group]. Each patient had a different pattern of joints with positive synovial vascularity: patients in the N group (ADA $n=2$, TCZ $n=2$), patients in the R group (ADA $n=3$, TCZ $n=3$), patients in the IP or PP groups (ADA $n=3$, TCZ $n=6$) and patients in the mixed R and IP or PP groups (ADA $n=5$, TCZ $n=7$).

The change in the LGSS (ΔLGSS) of the R group showed no progression as compared with the N group or showed improvement of joint damage in the PIP joints of the ADA treatment group (Fig. 1). We next focused on the joints with positive synovial vascularity after week 8, comprising the IP and PP groups. These joints showed an increased ΔLGSS as compared with the N group (Fig. 1). The ΔLGSS between the IP and PP groups showed no significant difference with either ADA or TCZ treatment (Fig. 1).

To analyse the relationship between synovial vascularity and ΔLGSS in more detail in the joints comprising the IP and PP groups, we calculated the sum of synovial vascularity of each finger joint from baseline to week 40 to represent the total exposure to inflammation during the treatment period. The medians of the sum of synovial vascularity with ADA therapy for the MCP and PIP joints were 1456 (range 71–6352) and 1136 (range 71–4757), respectively. The medians of the sum of synovial vascularity with TCZ therapy for the MCP and PIP joints were 2947 (range 71–11289) and 1385 (range 71–5964), respectively. We categorized these joints into two groups: those with a sum of synovial vascularity $\leqslant$ median value [the low-level (L) group], and those with a sum of synovial vascularity $>$ median value [the high-level (H) group]. There were no significant differences in the ΔLGSS between the L group and H group with either ADA or TCZ treatment (Fig. 1).

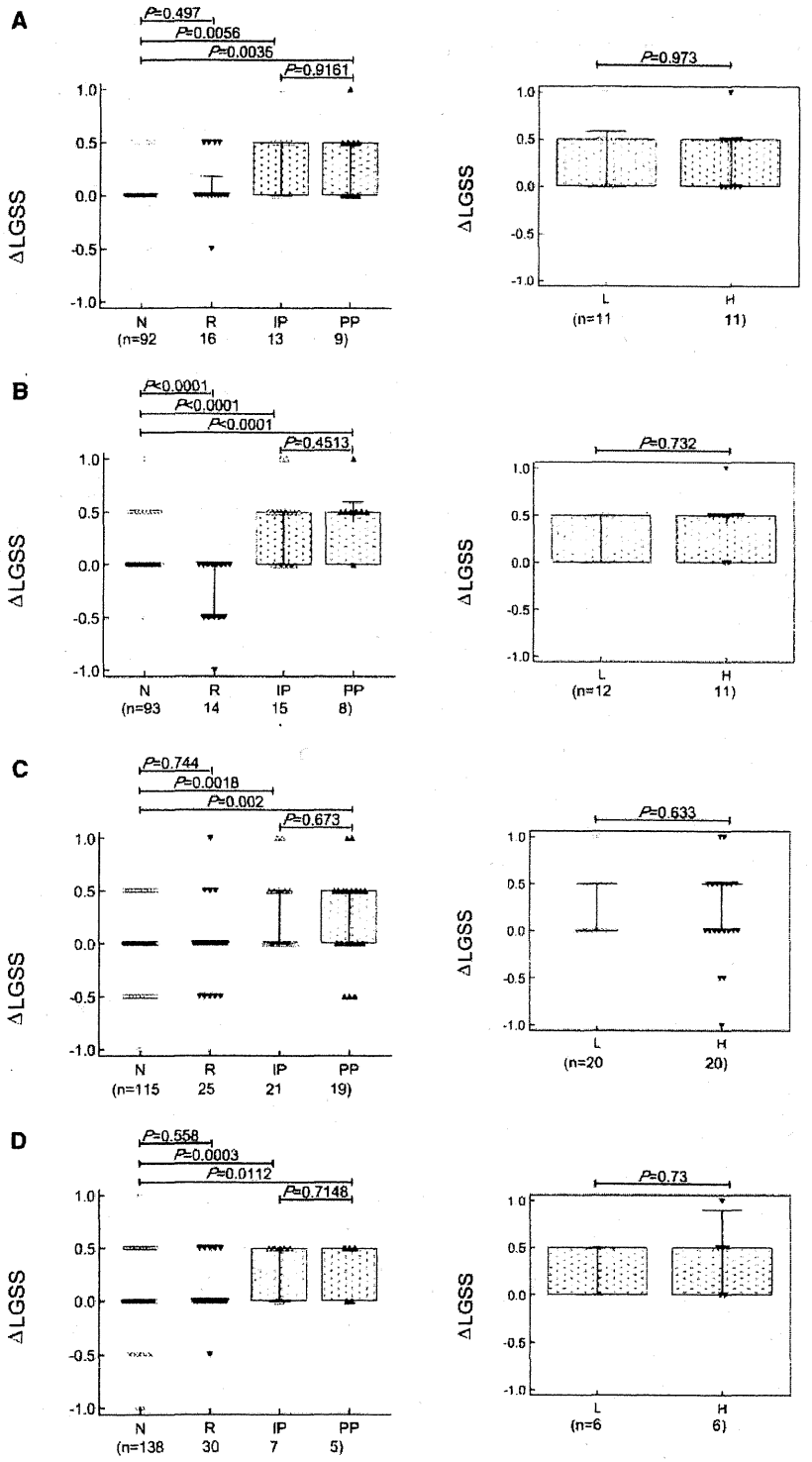
Intra- and interobserver reliability for power Doppler ultrasonography

Representative PDS images for 20 MCP and 20 PIP joints were randomly chosen, and synovial vascularity was measured three times each by the three ultrasonographers (M.H., F.S. and A.N.). The obtained intraobserver ICC values were 0.997–0.999 for MCP joints and 0.998–0.999 for PIP joints. The interobserver ICC values were 0.992–0.996 for MCP joints and 0.991–0.999 for PIP joints.

Discussion

Our study revealed two noteworthy results. First, this study further emphasized a previous report [7] that early improvement and then disappearance of synovial vascularity resulted in reducing joint damage progression.

FIG. 1 Relationship between positive synovial vascularity and LGSS in finger joints.



For ADA treatment, ΔLGSS of MCP (**A**) and PIP joints (**B**) is shown. For TCZ treatment, ΔLGSS of MCP (**C**) and PIP joints (**D**) is shown. Graphs on the left side show ΔLGSS of the N, R, IP and PP groups (Results section), which were categorized according to the occasional occurrence of positive synovial vascularity. For each joint in the IP and PP groups, the sum of synovial vascularity from baseline to week 40 was calculated and then categorized as L and H groups (Results section). Graphs on the right side show ΔLGSS of the L and H groups.

Secondly, a novel result was that persistence of positive synovial vascularity in local finger joints showed joint damage progression despite achieving low disease activity by biologic therapies. Interestingly, the ΔLGSS progressed independently of time-integrated joint inflammation estimated by the sum of synovial vascularity or occasional occurrence of positive synovial vascularity. These joints indicate the presence of low-level local joint inflammation, i.e. smouldering inflammation. The smouldering inflammatory joints could be categorized as a variation of subclinical synovitis described below.

Analysis of RA in the clinical remission phase revealed that there were asymptomatic or symptom-limited joints with poor prognosis. This joint inflammation or so-called subclinical synovitis can only be detected with imaging techniques [11–14]. The growing importance of imaging remission of rheumatoid activity has been confirmed, and imaging techniques such as joint ultrasonography have focused on detailed detection of local joint inflammation [15, 16].

Synovial vascularity detected by PDS is irrefutably linked to the level of joint inflammation [17, 18]. Naredo *et al.* [19] reported the correlation between time-integrated values of joint counts for positive synovial vascularity and total joint damage progression at 1 year. From these results, we speculated that increasing and persistent synovial vascularity might result in advanced joint damage progression; hence an increase in the occasional occurrence of positive synovial vascularity or the sum of synovial vascularity worsens the structural damage in smouldering inflammatory joints. Our data revealed that joints with positive synovial vascularity after week 8 (IP and PP groups) showed joint damage progression; however, their ΔLGSS progression did not relate to the occasional occurrence of positive synovial vascularity or the sum of synovial vascularity (Fig. 1). Accordingly, we concluded that the structural damage in joints with smouldering inflammation progressed independently of the level of the sum of synovial vascularity or the occasional occurrence of positive synovial vascularity. Importantly, the result might indicate that even low levels of positive synovial vascularity that occurred only once during the clinical improvement phase showed a risk for structural damage.

Although a correlation between the progression of systemic joint damage and time-integrated values of joint counts for positive synovial vascularity was reported [19], our study, which focused on synovitis and joint damage in individual finger joints, did not show such correlation. Whereas the previous study [19] showed the effect of non-biologic DMARDs, we studied biologic agents that rapidly improved acute inflammation. The DMARDs have slow therapeutic effect; thus the relationship between exposure to inflammation and joint damage progression may be closer in non-biologic DMARD users. Further, our data showed that some patients were in the mixed R and IP or PP group after starting biologic agents. This might indicate a discrepancy between overall therapeutic response and local joint response. Limitations of our study were its small scale and short observation period. Further larger studies are needed to confirm our observations.

In RA, tight control of joint inflammation is necessary for better outcomes. Treatment strategies should be changed according to the clinical response. Monitoring of synovial vascularity has the potential to provide useful joint information for daily practice and to tailor treatment strategies in RA.

Rheumatology key messages

- Finger joints with positive synovial vascularity under low disease activity showed structural deterioration in RA.
- Monitoring of synovial vascularity has the potential to provide useful information for daily practice in RA.

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

References

1 Jousse-Joulin S, d'Agostino MA, Marhadour T *et al.* Reproducibility of joint swelling assessment by sonography in patients with long-lasting rheumatoid arthritis (SEA-Repro study part II). J Rheumatol 2010;37:938–45.

2 Felson DT, Anderson JJ, Boers M *et al.* Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995;38:727–35.

3 Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum 1995;38:44–8.

4 Koch AE. Review: angiogenesis: implications for rheumatoid arthritis. Arthritis Rheum 1998;41:951–62.

5 Newman JS, Adler RS, Bude RO, Rubin JM. Detection of soft-tissue hyperemia: value of power Doppler sonography. AJR Am J Roentgenol 1994;163:385–9.

6 Naredo E, Moller I, Cruz A, Carmona L, Garrido J. Power Doppler ultrasonographic monitoring of response to anti-tumor necrosis factor therapy in patients with rheumatoid arthritis. Arthritis Rheum 2008;58:2248–56.

7 Fukae J, Isobe M, Kitano A *et al.* Radiographic prognosis of finger joint damage predicted by early alteration in synovial vascularity in patients with rheumatoid arthritis: potential utility of power Doppler sonography in clinical practice. Arthritis Care Res 2011;63:1247–53.

8 Fukae J, Kon Y, Henmi M *et al.* Change of synovial vascularity in a single finger joint assessed by power Doppler sonography correlated with radiographic change in rheumatoid arthritis: comparative study of a novel quantitative score with a semiquantitative score. Arthritis Care Res 2010;62:657–63.

9 Genant HK, Jiang Y, Peterfy C, Lu Y, Redei J, Countryman PJ. Assessment of rheumatoid arthritis using

a modified scoring method on digitized and original radiographs. Arthritis Rheum 1998;41:1583–90.

10 Bruynesteyn K, Boers M, Kostense P, van der Linden S, van der Heijde D. Deciding on progression of joint damage in paired films of individual patients: smallest detectable difference or change. Ann Rheum Dis 2005;64: 179–82.

11 Peluso G, Michelutti A, Bosello S, Gremese E, Tolusso B, Ferraccioli G. Clinical and ultrasonographic remission determines different chances of relapse in early and long standing rheumatoid arthritis. Ann Rheum Dis 2011;70: 172–5.

12 Brown AK, Quinn MA, Karim Z *et al.* Presence of significant synovitis in rheumatoid arthritis patients with disease-modifying antirheumatic drug-induced clinical remission: evidence from an imaging study may explain structural progression. Arthritis Rheum 2006;54: 3761–73.

13 Filippucci E, Iagnocco A, Meenagh G *et al.* Ultrasound imaging for the rheumatologist VII. Ultrasound imaging in rheumatoid arthritis. Clin Exp Rheumatol 2007; 25:5–10.

14 Brown AK, Conaghan PG, Karim Z *et al.* An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. Arthritis Rheum 2008;58:2958–67.

15 Saleem B, Brown AK, Keen H *et al.* Should imaging be a component of rheumatoid arthritis remission criteria? A comparison between traditional and modified composite remission scores and imaging assessments. Ann Rheum Dis 2011;70:792–8.

16 Dougados M, Devauchelle-Pensec V, Ferlet JF *et al.* The ability of synovitis to predict structural damage in rheumatoid arthritis: a comparative study between clinical examination and ultrasound. Ann Rheum Dis 2012. Advance Access published on 7 June 2012, doi:10.1136/annrheumdis-2012-201469.

17 Schirmer M, Duftner C, Schmidt WA, Dejaco C. Ultrasonography in inflammatory rheumatic disease: an overview. Nat Rev Rheumatol 2011;7:479–88.

18 Lainer-Carr D, Brahn E. Angiogenesis inhibition as a therapeutic approach for inflammatory synovitis. Nat Clin Pract Rheumatol 2007;3:434–42.

19 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.

Mod Rheumatol (2013) 23:1179–1185
DOI 10.1007/s10165-012-0816-5

ORIGINAL ARTICLE

Functional disability can deteriorate despite suppression of disease activity in patients with rheumatoid arthritis: a large observational cohort study

Yohei Seto · Eisuke Inoue · Kumi Shidara · Daisuke Hoshi · Naoki Sugimoto · Eri Sato · Eiichi Tanaka · Ayako Nakajima · Atsuo Taniguchi · Shigeki Momohara · Hisashi Yamanaka

Received: 10 October 2012 / Accepted: 10 December 2012 / Published online: 16 January 2013

Abstract
Objective To analyze the relationship between the progression of disability and disease activity in patients with rheumatoid arthritis (RA) in daily practice.
Methods Patients from an observational cohort, IORRA, who completed surveys during 2009–2011 were eligible. Linear regression of disease activity score 28 (DAS28), Japanese version of Health Assessment Questionnaire (J-HAQ), and EQ-5D from baseline were calculated, and the angles of the regression lines were designated DAS28 slope, J-HAQ slope, and EQ-5D slope, respectively, in each patient; averages were compared between treatment groups.
Results A total of 5,038 patients [84.0 % female, mean age 59.4 (SD 13.1) years, disease duration 13.2 (9.6) years, DAS28 3.29 (1.14), and J-HAQ 0.715 (0.760)] were analyzed. The average DAS28 slope indicated improvement in all groups, whereas J-HAQ slopes were negative in patients on methotrexate (MTX), biologics, combination biologics/disease-modifying antirheumatic drugs (DMARDs), and combination biologics/MTX at baseline, but positive in patients on prednisolone >5 mg/day [0.010 (0.153)] and not on MTX at baseline [0.007 (0.122)], representing a worsening of disability.
Conclusion There is some disparity between improvement of disease activity and progression of disability, suggesting that quality of remission must be considered.

Keywords Rheumatoid arthritis · Disease activity · Physical function · Treatment · Cohort study

Introduction

Treatment of rheumatoid arthritis (RA) has improved over the last ten years, following the introduction of new agents and modification of treatment strategies [1, 2]. As a consequence, clinical remission has become a realistic goal [3, 4]. Suppression of disease activity is the major factor that inhibits the progression of disability [1, 5, 6]. Since the treat-to-target (T2T) initiative first proposed the strategy of remission induction in the management of RA in daily practice [7, 8], the proportion of patients in remission has been increasing. In the Institute of Rheumatology, Rheumatoid Arthritis (IORRA) cohort, which we established at the Institute of Rheumatology, Tokyo Women's Medical University in 2000, the improvement in the disease activity in RA patients has been quite apparent, and nearly 40 % of all patients in 2011 achieved clinical remission [9]. This is thought to be the result of improved patient care following the recognition of the importance of remission induction in preventing joint destruction and preserving activities of daily living. However, disability in patients in the IORRA cohort has not improved as much as disease activity: the average disease activity score 28 (DAS28) improved from 4.16 in 2000 to 3.03 in 2011 (−28 %), while the Japanese version of Health Assessment Questionnaire (J-HAQ) score only improved from 0.80 in 2000 to 0.64 in 2011 (−18 %). This triggers the question of whether the

Electronic supplementary material The online version of this article (doi:10.1007/s10165-012-0816-5) contains supplementary material, which is available to authorized users.

Y. Seto (✉) · E. Inoue · K. Shidara · D. Hoshi · N. Sugimoto · E. Sato · E. Tanaka · A. Nakajima · A. Taniguchi · S. Momohara · H. Yamanaka
Institute of Rheumatology, Tokyo Women's Medical University, 10-22 Kawada-cho, Shinjuku-ku, Tokyo 162-0054, Japan
e-mail: seto@ior.twmu.ac.jp

remissions achieved by biologics versus corticosteroids yield comparable outcomes with respect to disability.

We hypothesized that the discrepancy between improvement in disease activity and progression of disability might depend on which agents are administered. Indeed, we have previously demonstrated that patients treated with corticosteroids develop progressive disability even when they are in remission, and suggested that there is a difference in the quality of remission between different agent classes [10, 11]. In the study described in this report, we analyzed the relationship between control of disease activity and progression of disability in our observational cohort, IORRA.

Patients and methods

Patients and the IORRA database

We established a prospective observational cohort of RA patients at the Institute of Rheumatology, Tokyo Women's Medical University in October 2000; this is designated the IORRA cohort. Patients with RA who fulfilled the American College of Rheumatology criteria for RA [2, 12, 13] were registered, and their information and data were collected biannually (in April and October) when patients visited the outpatient unit of our institute for consultation. Informed consent was routinely obtained from each patient at each visit.

The IORRA database consists of three components. The first component is the physician's evaluation, which includes the number of tender joints, number of swollen joints, and a visual analogue scale (VAS) of disease activity rated by the physician. The second component is information collected from patients, which includes VAS for pain, VAS for general health, disability level using J-HAQ score [12], height, body weight, and comorbidities in the previous six months. Information about medication actually taken (not just prescribed) during the period was also reported, including corticosteroid use and daily dose, disease-modifying antirheumatic drug (DMARD) use, methotrexate (MTX) use and weekly dose, and biologics use. Patients were asked by the attending physician to answer these questions by completing questionnaire sheets at home and mailing them back in a pre-stamped envelope within two weeks of their visit. The third component is laboratory data, including C-reactive protein (CRP) level, erythrocyte sedimentation rate (ESR), blood cell count, liver transaminase levels, and urinalysis. All information was integrated into a single database that was used for analysis. DAS28 and EQ-5D scores were calculated according to the original methods [14, 15].

Over 99 % of RA patients in our institute participated in the study during this time period, and over 98 % of patients completed and returned their questionnaires. Patients who participated in the IORRA survey between April 2009 and April 2011 were eligible for this study.

Methods

Institute of Rheumatology, Rheumatoid Arthritis is an observational cohort database in which longitudinal data about disease activity, disability, and treatment have been collected. The linear regressions of DAS28, J-HAQ, and EQ-5D scores from baseline were calculated for each patient, and the angles of the regression lines were designated the DAS28 slope, J-HAQ slope, and EQ-5D slope, respectively. Specifically, for each patient who participated in IORRA at least three times during the study period, a linear regression was conducted with longitudinal DAS28, J-HAQ, and EQ-5D as a response and continuous time, 0, 1, 2, and 3 as an explanatory variable. The DAS28 slope increases as disease activity worsens, the J-HAQ slope increases as disability progresses, and the EQ-5D slope increases as quality of life improves.

Average values for the DAS28 slope, J-HAQ slope, and EQ-5D slope were calculated for different patient groups as follows: (a) patients receiving MTX were classified by MTX dose, (b) patients on oral corticosteroids were classified by equivalent prednisolone (PSL) dose, and (c) patients were classified by agents used at baseline: on conventional DMARDs other than MTX, on MTX, on biologics, and on a combination of DMARDs and biologics or a combination of MTX and biologics. At baseline in April 2009, four biologics were available in Japan: infliximab, etanercept, adalimumab, and tocilizumab. The average DAS28 slope, J-HAQ slope, and EQ-5D slope were compared between treatment groups, respectively. Written consent was obtained from each patient who participated in the study, according to the Declaration of Helsinki (most recently revised at the General Assembly in October 2008), and the study was approved by the local ethics committee at Tokyo Women's Medical University.

Statistical analysis

Means and their standard deviations were used to describe data for continuous variables, and proportions were used to describe data for discrete variables. The time coefficients, representing the degree of progression for each outcome in each patient, were analyzed using the mean and its 95 % confidence interval according to the treatment the patient received. The distributions of the DAS28 slope, J-HAQ slope, and EQ-5D slope were visualized by plotting the

cumulative probability [16–18], a method that has been commonly used to present radiographic progression in RA clinical studies, in order to highlight differences in the DAS28 slope, J-HAQ slope, and EQ-5D slope between treatment groups. All of the calculations were done using the statistical software R (http://cran.r-project.org/, version 2.14.0).

Results

Baseline characteristics

A total of 5,038 patients [84.0 % female, mean age 59.4 (SD 13.1) years, disease duration 13.2 (9.6) years at baseline] whose data from consecutive visits were available were recruited from the cohort.

The average DAS28, J-HAQ, and EQ-5D scores at baseline were 3.29 (1.14), 0.715 (0.760), and 0.760 (0.174), respectively. Patients were subclassified by treatment at baseline (April 2009). The baseline characteristics of these treatment groups are shown in Table 1 and in Tables S1, S2, S3, and S4 of the Electronic supplementary material (ESM).

Table 1 Baseline demographic and disease characteristics

Characteristics	Study population ($N = 5,038$)
Age (years)	59.4 (13.1)
Women (%)	84.0
Duration (years)	13.2 (9.6)
DAS28	3.29 (1.14)
CDAI	7.56 (6.48)
SDAI	8.28 (7.04)
EQ-5D	0.76 (0.17)
J-HAQ (0–3 scale)	0.715 (0.760)
Tender joint count (0–45)	1.8 (3.2)
Swollen joint count (0–45)	1.9 (2.8)
Pain VAS (0–100 scale)	28.7 (25.2)
Patient global VAS (0–100 scale)	30.3 (24.6)
Physician global VAS (0–100 scale)	15.6 (15.2)
CRP (mg/dl)	0.73 (1.27)
ESR (mm/h)	31.7 (22.9)
DMARDs (%)	91.0
MTX (%)	68.5
MTX dose (mg/week)	8.0 (3.1)
Prednisolone (%)	46.8
Prednisolone dose (mg/day)	4.2 (2.9)
Biologics (%)	8.7

Values are the mean (SD) unless indicated otherwise

DAS28 disease activity score 28, *CDAI* clinical disease activity index, *SDAI* simplified disease activity index, *J-HAQ* Japanese version of Health Assessment Questionnaire, *VAS* visual analogue scale, *CRP* C-reactive protein, *ESR* erythrocyte sedimentation rate, *DMARDs* disease-modifying anti-rheumatic drugs, *MTX* methotrexate

DAS28 slope

The DAS28 slopes of each treatment group are shown in Table 2. During the observation period, the average DAS28 slope was negative (i.e., it sloped downward) in all subgroups, indicating improvement in disease activity in all subgroups.

J-HAQ slope

The J-HAQ slopes in each treatment group are shown in Table 3. The J-HAQ slope was negative in patients who were receiving MTX, biologics, a combination of biologics and DMARDs, and a combination of biologics and MTX at baseline. In contrast, the J-HAQ slope was positive (i.e., it sloped upward) in patients receiving >5 mg/day of PSL [0.010 (0.153)] and in those without MTX [0.007 (0.122)] at baseline, indicated worsening functional disability.

EQ-5D slope

The EQ-5D slopes in each treatment group are shown in Table 4. The EQ-5D slope was positive in all treatment groups. A significant dose response of EQ-5D was observed in the MTX groups, and the largest improvement in EQ-5D was observed in patients who were on biologics + DMARDs/MTX at baseline. In addition, the increase in EQ-5D slope was larger in the MTX groups than in the PSL groups.

Distribution of the J-HAQ slope

The distribution of the J-HAQ slope was visualized using the probability plot method (Fig. 1). Each plot successfully shows the difference in the J-HAQ slope among groups.

Discussion

In this study, we developed the concept of DAS28 slope, J-HAQ slope, and EQ-5D slope to assess changes over a specified time period. In clinical studies, delta DAS28 and delta HAQ are often used to indicate changes between two time points, i.e., the baseline and endpoint. However, in our longitudinal observational study, DAS28, J-HAQ, and EQ-5D scores were recorded at multiple time points, and these values fluctuated over time; thus, we considered it more logical to evaluate the changes by analyzing linear regression using multiple values at multiple time points. Thus, the DAS28 slope, J-HAQ slope, and EQ-5D slope are

Table 2 DAS28 at baseline and DAS28 slopes

	N	DAS28 (95 % CI)	DAS28 slope (95 % CI)
All	5,038	3.29 (3.26 to 3.33)	−0.071 (−0.080 to −0.061)
Not on MTX	1,586	3.22 (3.16 to 3.28)	−0.041 (−0.058 to −0.025)
MTX 0–4 mg/week	488	3.30 (3.20 to 3.40)	−0.068 (−0.100 to −0.037)
MTX 4–6 mg/week	827	3.28 (3.20 to 3.36)	−0.069 (−0.091 to −0.047)
MTX 6–8 mg/week	966	3.34 (3.26 to 3.41)	−0.083 (−0.104 to −0.062)
MTX 8–10 mg/week	627	3.38 (3.30 to 3.47)	−0.093 (−0.118 to −0.068)
MTX >10 mg/week	521	3.35 (3.26 to 3.44)	−0.106 (−0.131 to −0.081)
Not on PSL	2,682	3.10 (3.06 to 3.15)	−0.056 (−0.068 to −0.044)
PSL 0–3 mg/day	996	3.36 (3.29 to 3.43)	−0.075 (−0.095 to −0.055)
PSL 3–5 mg/day	873	3.56 (3.49 to 3.63)	−0.087 (−0.110 to −0.064)
PSL >5 mg/day	487	3.73 (3.62 to 3.83)	−0.112 (−0.144 to −0.080)
DMARDs	4,587	3.29 (3.26 to 3.33)	−0.071 (−0.080 to −0.061)
MTX	3,429	3.33 (3.29 to 3.37)	−0.083 (−0.094 to −0.072)
Biologics	437	3.28 (3.17 to 3.39)	−0.104 (−0.135 to −0.073)
MTX monotherapy	1,944	3.28 (3.24 to 3.33)	−0.066 (−0.080 to −0.052)
Biologic monotherapy	74	3.28 (3.05 to 3.51)	−0.079 (−0.146 to −0.012)
Biologic with DMARD	363	3.28 (3.15 to 3.41)	−0.109 (−0.144 to −0.074)
Biologic with MTX	345	3.25 (3.12 to 3.38)	−0.107 (−0.143 to −0.071)

CI confidence interval

Table 3 J-HAQ at baseline and J-HAQ slopes

	N	J-HAQ (95 % CI)	J-HAQ slope (95 % CI)
All	5,038	0.715 (0.693 to 0.735)	−0.001 (−0.004 to 0.002)
Not on MTX	1,586	0.690 (0.651 to 0.729)	0.007 (0.000 to 0.013)
MTX 0–4 mg/week	488	0.772 (0.697 to 0.848)	0.001 (−0.010 to 0.011)
MTX 4–6 mg/week	827	0.727 (0.673 to 0.780)	−0.004 (−0.012 to 0.003)
MTX 6–8 mg/week	966	0.733 (0.686 to 0.780)	−0.003 (−0.010 to 0.003)
MTX 8–10 mg/week	627	0.680 (0.628 to 0.732)	−0.005 (−0.014 to 0.004)
MTX >10 mg/week	521	0.703 (0.646 to 0.759)	−0.013 (−0.021 to −0.004)
Not on PSL	2,682	0.542 (0.517 to 0.567)	−0.004 (−0.008 to 0.000)
PSL 0–3 mg/day	996	0.848 (0.797 to 0.897)	0.001 (−0.006 to 0.007)
PSL 3–5 mg/day	873	0.899 (0.847 to 0.952)	0.000 (−0.008 to 0.008)
PSL >5 mg/day	487	1.062 (0.987 to 1.138)	0.010 (−0.004 to 0.024)
DMARDs	4,587	0.705 (0.683 to 0.726)	−0.001 (−0.004 to 0.002)
MTX	3,429	0.723 (0.698 to 0.748)	−0.005 (−0.008 to −0.001)
Biologics	437	0.917 (0.843 to 0.990)	−0.016 (−0.026 to −0.006)
MTX monotherapy	1,944	0.694 (0.661 to 0.727)	−0.002 (−0.007 to 0.003)
Biologic monotherapy	74	1.193 (0.990 to 1.395)	−0.006 (−0.029 to 0.016)
Biologic with DMARD	363	0.861 (0.783 to 0.938)	−0.018 (−0.029 to −0.007)
Biologic with MTX	345	0.853 (0.773 to 0.934)	−0.022 (−0.033 to −0.010)

essentially analogous to delta DAS28, delta J-HAQ, and delta EQ-5D, respectively; however, we consider the former to be more logical indicators of changes over time in the IORRA observational cohort study.

The baseline characteristics of the treatment groups in this analysis differed, so it is not appropriate to compare the results between treatment groups. This is an inherent limitation of observational studies, in contrast to well-designed clinical trials [19]. It may be possible to compare outcomes by matching patients with comparable disease activity or disability, as done in propensity score matching [20], in order to determine which treatment is superior for

Table 4 EQ-5D at baseline and EQ-5D slopes

	N	EQ-5D (95 % CI)	EQ-5D slope (95 % CI)
All	5,038	0.76 (0.76 to 0.76)	0.0046 (0.0032 to 0.0059)
Not on MTX	1,586	0.77 (0.76 to 0.78)	−0.0008 (−0.0037 to 0.0022)
MTX 0–4 mg/week	488	0.76 (0.74 to 0.78)	0.0046 (0.0005 to 0.0086)
MTX 4–6 mg/week	827	0.76 (0.75 to 0.78)	0.0061 (0.0027 to 0.0095)
MTX 6–8 mg/week	966	0.75 (0.74 to 0.76)	0.0072 (0.0043 to 0.0100)
MTX 8–10 mg/week	627	0.76 (0.74 to 0.77)	0.0075 (0.0042 to 0.0108)
MTX >10 mg/week	521	0.75 (0.73 to 0.76)	0.0082 (0.0049 to 0.0115)
Not on PSL	2,682	0.80 (0.79 to 0.81)	0.0051 (0.0032 to 0.0069)
PSL 0–3 mg/day	996	0.74 (0.73 to 0.75)	0.0047 (0.0018 to 0.0076)
PSL 3–5 mg/day	873	0.72 (0.71 to 0.73)	0.0042 (0.0009 to 0.0074)
PSL >5 mg/day	487	0.67 (0.66 to 0.69)	0.0022 (−0.0034 to 0.0077)
DMARDs	4,587	0.76 (0.76 to 0.77)	0.0049 (0.0035 to 0.0063)
MTX	3,429	0.76 (0.75 to 0.76)	0.0068 (0.0053 to 0.0083)
Biologics	437	0.73 (0.71 to 0.74)	0.0120 (0.0074 to 0.0165)
MTX monotherapy	1,944	0.76 (0.75 to 0.77)	0.0053 (0.0033 to 0.0073)
Biologic monotherapy	74	0.69 (0.65 to 0.73)	0.0079 (−0.0045 to 0.0203)
Biologic with DMARD	363	0.73 (0.71 to 0.75)	0.0128 (0.0079 to 0.0177)
Biologic with MTX	345	0.73 (0.72 to 0.75)	0.0135 (0.0084 to 0.0186)

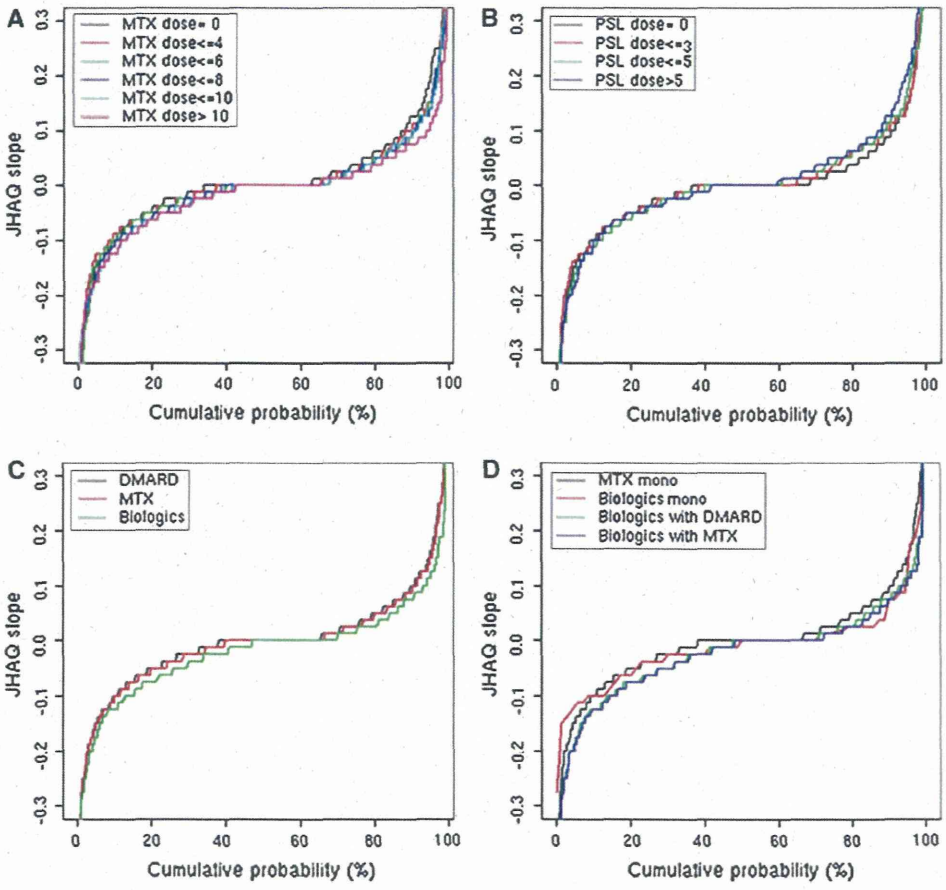


Fig. 1 Probability plots of J-HAQ slopes described for groups subclassified by **a** MTX dose, **b** PSL dose, **c** DMARD, MTX, or biologics users, and **d** monotherapy of MTX/ biologics and combination therapies

the prevention of disability. However, the goal of this study was to examine the differences between control of disease activity as estimated by DAS28 and progression of disability as estimated by J-HAQ among the different treatment groups.

As indicated by the DAS28 slope shown in Table 2, an improvement in DAS28 was observed in every treatment group during the two years from baseline, indicating that the treatment successfully improved the signs and symptoms of RA. Biologics strikingly decreased DAS28 with or without DMARDs, and PSL also decreased DAS28 in a dose-dependent manner. As indicated by the EQ-5D slope, patient QOL also improved during the observation period. However, the J-HAQ slopes differed markedly between treatment groups. Dramatic decreases in the J-HAQ slope were seen in patients on biologics, as well as in patients on MTX (in a dose-dependent manner), whereas increases were observed in patients on PSL dosed at >5 mg/day, indicating worsening functional disability.

These data were also visualized using probability plots. As this method has been frequently used to visualize the distribution of the progression of bone damage as assessed by X-ray scoring methods, we elected to utilize it to show progression of disability. Biologics, sufficiently dosed MTX, and low-dose PSL suppressed this ratio, indicating that patients in these groups are treated properly with respect to preventing the progression of disability. While similar results have been reported from clinical trials [21], the present data reflect patients seen in actual clinical practice.

As proposed in the T2T initiative, the primary target for treatment of RA should be clinical remission [1–4]. As mentioned above, we previously reported that patients treated with corticosteroids experience progression of disability even when they are in remission, and we suggested that there is a difference in the quality of remission between different agent classes [10, 11]. The present study supports this finding and suggests that there should be some discrepancy between suppression of disease activity and maintenance of physical function.

Another mechanism of joint damage prevention in well-controlled RA patients may be the consequence of residual disease activity [22, 23]. We have used DAS28 as a measure of disease activity; however, the progression of disability may be caused by the involvement of joints other than the 28 joints used in this assessment.

A major limitation of this study is the observational nature of the study design; thus, outcomes could not be accurately compared between treatment groups. Indeed, it is also true that the selection of the baseline treatment may be strongly associated with the baseline patients' comorbidities or previous drug histories, etc. Therefore, these background characteristics may influence the outcomes of groups categorized according to the type of agent selected. Our present data do not indicate that corticosteroid should never be given on any occasion in daily practice, or that corticosteroid itself worsens functional disability. In addition, treatment group assignment was based on the treatment at baseline, so the effects of changes in treatment during the observational period could not be assessed.

In conclusion, there is considerable disparity between control of disease activity and progression of disability in RA patients, and this disparity is correlated with the type and dosage of agent used. Remission is a realistic target under current management guidelines; however, differences in the quality of remission must also be considered.

Acknowledgments The authors thank all members involved in the administration of the IORRA cohort database.

Conflict of interest This study was partly supported by a research grant from the Ministry of Health, Welfare and Labor Japan. The IORRA cohort was supported by unrestricted research grants from 36 pharmaceutical companies: Abbott Japan, Asahikasei Kuraray Medical, Asahikasei Pharma, Astellas Pharma, AstraZeneca, MSD, Chugai Pharmaceutical, Daiichi Fine Chemical, Daiichi Sankyo, Dainippon Sumitomo Pharma, Eisai, GlaxoSmithKline, Janssen Pharmaceutical, Japan Tobacco, Kaken Pharmaceutical, Kissei Pharmaceutical, Kowa Pharmaceutical, Mitsubishi Chemical Medience, Mitsubishi Tanabe Pharma, Nippon Chemiphar, Nippon Shinyaku, Novartis Pharma, Otsuka Pharmaceutical, Pfizer, Sanofi-Aventis, Santen Pharmaceutical, Sanwa Kagaku Kenkyusho, Sekisui Medical, Taisho Toyama Pharmaceutical, Takeda Pharmaceutical, Teijin Pharma, Torii Pharmaceutical, Toyama Chemical, UCB Japan, and Zeria Pharmaceutical. YS has received speaking fees from Abbott Japan, Chugai Pharmaceutical, Eisai, Mitsubishi Tanabe Pharma, and Pfizer. ET has received speaking fees and/or consulting fees from Abbott Japan, Astellas Pharma, Bristol–Myers Squibb, Chugai Pharmaceutical, Eisai, Mitsubishi Tanabe Pharma, Pfizer, Santen Pharmaceutical, and Takeda Pharmaceutical. AN has received speaking fees and/or consulting fees from Abbott Japan, Astellas Pharma, Chugai Pharmaceutical, Eisai, Janssen Pharmaceutical, Mitsubishi Tanabe Pharma, Pfizer, Sanwa Kagaku Kenkyusho, and UCB Japan. AT has received speaking fees from Abbott Japan, Bristol–Myers Squibb, Eisai, Mitsubishi Tanabe Pharma, Pfizer, Takeda Pharmaceutical, Teijin Pharma, and Torii Pharmaceutical. SM has received speaking fees from Chugai Pharmaceutical, Eisai, and Mitsubishi Tanabe Pharma. HY has received consultant fees, speaking fees, and/or honoraria from Chugai Pharmaceutical, Daiichi Sankyo, Mitsubishi Tanabe Pharma, Abbott Japan, Eisai, Takeda Pharmaceutical, Janssen Pharmaceutical, Hoffmann-La Roche, and Pfizer. All other authors have declared no conflicts of interest.

References

1. Smolen JS, Aletaha D. Monitoring rheumatoid arthritis. Curr Opin Rheumatol. 2011;23:252–8.
2. Yamanaka H, Inoue E, Singh G, et al. Improvement of disease activity of rheumatoid arthritis patients from 2000 to 2006 in a large observational cohort study IORRA in Japan. Mod Rheumatol. 2007;17:283–9.

3. Jayakumar K, Norton S, Dixey J, et al. Sustained clinical remission in rheumatoid arthritis: prevalence and prognostic factors in an inception cohort of patients treated with conventional DMARDS. Rheumatology (Oxford). 2012;51:169–75.
4. Tanaka Y. Intensive treatment and treatment holiday of TNF-inhibitors in rheumatoid arthritis. Curr Opin Rheumatol. 2012;24: 319–26.
5. Aletaha D, Ward MM, Machold KP, et al. Remission and active disease in rheumatoid arthritis: defining criteria for disease activity states. Arthritis Rheum. 2005;52:2625–36.
6. Tanaka E, Mannalithara A, Inoue E, et al. Efficient management of rheumatoid arthritis significantly reduces long-term functional disability. Ann Rheum Dis. 2008;67:1153–8.
7. Smolen JS, Aletaha D, Bijlsma JW, et al. Treating rheumatoid arthritis to target: recommendations of an international task force. Ann Rheum Dis. 2010;69:631–7.
8. Vermeer M, Kuper HH, Hoekstra M, et al. Implementation of a treat-to-target strategy in very early rheumatoid arthritis: results of the Dutch Rheumatoid Arthritis Monitoring remission induction cohort study. Arthritis Rheum. 2011;63:2865–72.
9. Yamanaka H, Seto Y, Tanaka E, et al. Management of rheumatoid arthritis: the 2012 perspective. Mod Rheumatol. 2012. doi: 10.1007/s10165-012-0702-1.
10. Iikuni N, Inoue E, Tanaka E, et al. Low disease activity state with corticosteroid may not represent 'true' low disease activity state in patients with rheumatoid arthritis. Rheumatology (Oxford). 2008;47:519–21.
11. Tanaka E, Mannalithara A, Inoue E, et al. Effects of long-term corticosteroid usage on functional disability in patients with early rheumatoid arthritis, regardless of controlled disease activity. Rheumatol Int. 2010;32:749–57.
12. Matsuda Y, Singh G, Yamanaka H, et al. Validation of a Japanese version of the Stanford Health Assessment Questionnaire in 3,763 patients with rheumatoid arthritis. Arthritis Rheum. 2003;49: 784–8.
13. Yamanaka H, Inoue E, Tanaka E, et al. Influence of methotrexate dose on its efficacy and safety in rheumatoid arthritis patients: evidence based on the variety of prescribing approaches among practicing Japanese rheumatologists in a single institute-based large observational cohort (IORRA). Mod Rheumatol. 2007;17: 98–105.
14. Prevoo ML, van't Hof MA, Kuper HH, et al. Modified disease activity scores that include twenty-eight-joint counts: development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum. 1995;38: 44–8.
15. Hurst NP, Kind P, Ruta D, et al. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). Br J Rheumatol. 1997;36:551–9.
16. Emery P, Genovese MC, van Vollenhoven R, et al. Less radiographic progression with adalimumab plus methotrexate versus methotrexate monotherapy across the spectrum of clinical response in early rheumatoid arthritis. J Rheumatol. 2009;36: 1429–41.
17. Takeuchi T, Tanaka Y, Kaneko Y, et al. Effectiveness and safety of adalimumab in Japanese patients with rheumatoid arthritis: retrospective analyses of data collected during the first year of adalimumab treatment in routine clinical practice (HARMONY study). Mod Rheumatol. 2011;22:327–38.
18. Kameda H, Kanbe K, Sato E, et al. Continuation of methotrexate resulted in better clinical and radiographic outcomes than discontinuation upon starting etanercept in patients with rheumatoid arthritis: 52-week results from the JESMR study. J Rheumatol. 2011;38:1585–92.
19. Furst DE. Observational cohort studies and well controlled clinical trials—we need them both! J Rheumatol. 2004;31:1476–7.
20. Ozminkowski RJ, Burton WN, Goetzel RZ, et al. The impact of rheumatoid arthritis on medical expenditures, absenteeism, and short-term disability benefits. J Occup Environ Med. 2006;48: 135–48.
21. Smolen JS, Aletaha D, Grisar JC, et al. Estimation of a numerical value for joint damage-related physical disability in rheumatoid arthritis clinical trials. Ann Rheum Dis. 2010;69:1058–64.
22. 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.
23. 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.

Journal of Evaluation in Clinical Practice
International Journal of Public Health Policy and Health Services Research

Journal of Evaluation in Clinical Practice ISSN 1365-2753

Opinion of Japanese rheumatology physicians on methods of assessing the quality of rheumatoid arthritis care

Takahiro Higashi MD PhD,[1] Shunichi Fukuhara MD MSc DMSc[2] and Takeo Nakayama MD MPH PhD[3]

[1]Associate Professor, Department of Public Health/Health Policy, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan
[2]Professor, Department of Epidemiology and Healthcare Research, Kyoto University School of Public Health, Kyoto, Japan
[3]Professor, Department of Health Informatics, Kyoto University School of Public Health, Kyoto, Japan

Keywords
physician opinion, quality of care, rheumatoid arthritis

Correspondence
Dr Takahiro Higashi
Department of Public Health/Health Policy
Graduate School of Medicine
The University of Tokyo
7-3-1 Hongo
Bunkyo-ku
Tokyo 113-0033
Japan
E-mail: higashi@m.u-tokyo.ac.jp

This study was supported in part by a grant-in-aid (health technology assessment) from the Ministry of Health, Labor and Welfare, Japan.

Accepted for publication: 5 August 2010

doi:10.1111/j.1365-2753.2010.01575.x

Abstract

Objective To examine the opinion of rheumatology physicians in Japan regarding desirable quality assessment methods.

Methods We conducted a cross-sectional self-administered mail survey on a random sample of physicians and surgeons registered with the Japan Rheumatism Foundation. In the survey, respondents were asked to rank seven proposed assessment methods for the quality of rheumatoid arthritis care, namely patient satisfaction, risk-adjusted outcomes such as complication incidence and admission rate, guideline compliance, waiting time at clinics, voting by local general practitioners, degree of newspaper and magazine reportage, and volume of patients receiving treatment for rheumatoid arthritis.

Results Among 531 respondents (response rate 48%), the respondents ranked patient satisfaction most favourably (mean rank 1.6), followed by complication/admission rate and number of patients. Guideline adherence was ranked almost the same as voting by local physicians. Waiting time and media reportage were not considered good methods for quality evaluation. Ranking distribution did not differ by working facility or place, volume of patients or years in practice. Multivariate analysis revealed that respondents who care for a large number of rheumatoid arthritis patients (>40 regular patients) were less likely to rank guideline adherence highly (first to third) than those who care for few patients (≤10 regular patients), with an odds ratio of 0.38 ($P < 0.01$) after adjustment for other variables.

Conclusions A majority of Japanese rheumatology physicians consider patient satisfaction the most trustworthy method of assessing the quality of rheumatoid arthritis care. Future research should explore convincing methods of assessing the technical quality of rheumatoid arthritis care.

Introduction

The quality of medical care has gained increasing public attention. Many studies have reported unexplained variations in care across geography [1–4], setting [5] and race [6] in western countries, and gaps between current standards and actual practice [7,8]. This growing concern has fuelled activities to measure and publicly report the quality of medical care for accountability purposes [9,10], and many settings in western countries have gone so far as to adopt payment schemes that reward a high quality of care [11–13]. Concern is growing in Japan also, where professionals have traditionally enjoyed freedom from rigorous quality scrutiny, as exemplified by increases in the number of malpractice litigation cases [14] and in media reporting of quality information, such as the surgical mortality of hospitals based on their own surveys [15,16]. A clear need to systematically measure the quality of care has emerged.

One popular way of measuring quality is to evaluate the process of care in comparison with a set of explicit criteria [7,8,17]. These explicit criteria usually describe standards of care based on clinical evidence and professional consensus, and quality is calculated as the proportion of patients who receive the described care among those eligible for it. These criteria sometimes derive from clinical practice guidelines, which are usually based on clinical evidence and professional consensus. Although the quality of care can also be measured using structural (i.e. staff–patient ratio and presence/absence of high-tech equipment) or outcome (patient survival or re-admission rate) measures, process measures have advantages, such as not requiring the statistical case-mix adjustment necessary in outcome measures and the ability to examine the care provided, for which providers are directly responsible.

Despite the growth of quality measurement in western countries, provider opinions of how quality should be measured are

rarely examined. Providers naturally oppose the idea of 'being measured' and tend to be critical of quality measurement. Simple questioning on whether a certain quality measurement (e.g. process measurement) is appropriate may result in a majority negative response which merely reflects reluctance to be measured. In this regard, a survey of US generalist physicians revealed that 70% of respondents felt that quality is not adequately measured at present, while a majority of the same sample were willing to be paid on quality provided that quality is adequately measured [18]. A qualitative study and an anecdotal story show that the current quality measurement schemes can distort the traditional goodness of the physician–patient encounter [19]. If these critiques shed light on the problems of quality measurement, the need to assure the accountability of health care providers may warrant the consideration of alternative ways of measuring quality. An understanding of physician opinions of how quality should be operationally measured may help identify optimum approaches and facilitate physician cooperation in measuring and improving quality.

Using a survey of attitudes towards the newly revised clinical practice guidelines for rheumatoid arthritis, we investigated current provider opinions of rheumatology physicians defined as physicians whose practice is focused on rheumatic diseases, including rheumatologists, orthopaedic surgeons and some general physicians in Japan regarding which methods are desirable in evaluating the quality of rheumatoid arthritis care. We also analysed the relative degree of acceptance of process-of-care quality measurement among alternative methods of quality assessment.

Methods

Physician survey

We analysed data obtained from a larger survey conducted to evaluate the usefulness of the revised Japanese rheumatoid arthritis clinical practice guidelines [20,21] and rheumatology physicians' general attitudes towards clinical practice guidelines. Details of the survey are reported elsewhere [22]. Briefly, the survey was distributed to a random sample of rheumatology physicians registered with the Japan Rheumatism Foundation. This Foundation is an affiliate of the Japan College of Rheumatology, which plays a central role in supporting research and practice in rheumatology in Japan by funding programmes and disseminating up-to-date information to providers and patients. Eligibility to register with the Foundation is limited to physicians who have been focused on rheumatology practice for at least 5 years and are approved by the review committee based on documentation of cases they have cared for. They are typically but not exclusively rheumatologists and orthopaedic surgeons. The survey was conducted in two waves, the first in December 2002 and the second in March–April 2006. Only the second included questions related to quality of care, and thus the current analysis used this wave only.

Quality of care question item

Among questions about the rheumatoid arthritis guidelines and respondents' practice patterns, the second wave survey included several items that asked about the quality of care in the framework of clinical practice guidelines. The main question asked respondents to rank proposed methods of assessing the quality of institution-provided rheumatoid arthritis care, namely patient satisfaction, risk-adjusted outcomes such as complication incidence and admission rate, guideline compliance, waiting time at clinics, voting by local general practitioners, degree of reportage by newspapers and magazines, and the volume of patients receiving treatment for rheumatoid arthritis. Tied rankings were not explicitly permitted but were treated as such if selected. Because this question of quality assessment was the focus of the present analysis, only respondents who answered this item were entered in the analyses.

Statistical analysis

To obtain a summarized group opinion, we report the modal rank and mean rank for each candidate quality assessment method. After ranking methods by mean rank, we then tested statistical differences in mean ranks between adjacently ranked methods (i.e. first versus second rank, second versus third rank, etc.) using the *t*-test.

Focusing on process measures as represented by guideline adherence, we further examined the relationship of respondent characteristics with the high ranking (i.e. first to third ranking among the proposed measures) of 'guideline adherence' as the quality measure. First, we described the proportion of respondents who ranked guideline adherence highly by stratifying physician characteristics, and then compared proportions using the chi-squared test. Second, we used a multivariable logistic regression to examine the independent association of these factors with the high ranking of guideline adherence. The examined factors included respondent gender, years in practice (<20 years/21–40 years/>41 years), specialty (surgeon/internists), patient volume (~10/11–20/21–30/31–40/≥41 patients for whom the respondents care regularly), type of practice setting (office practice/non-university hospital/university hospital), practice location (eastern/western Japan) and area type (metropolitan/urban/rural). Non-respondents to each item were excluded from the bivariate and multivariate analyses. An alpha level of 0.05 was used to decide statistical significance. The study protocol was approved by the Institutional Review Board of Kyoto University Graduate School of Medicine and Public Health.

Results

Among 1111 physicians surveyed in the second wave, 531 (48%) responded to the question about quality assessment and were entered into the analysis. Respondent characteristics are presented in Table 1; average age was 54 years (range 37–91), with 28 years in practice (12–64), and 5% were female.

Figure 1 shows the distribution of assigned rankings for each quality assessment method. Patient satisfaction was most favourably ranked, with a mean rank of 1.6, followed by complication rate (mean rank 2.7) and number of patients (mean rank 3.3). Guideline adherence was ranked mostly in the middle with a mode ranking of 4 (mean 4.0), which was about the same as that for voting by local physicians (mean rank 4.2). Assessment by waiting time and reportage in newspapers and magazines were considered

Table 1 Subject characteristics ($n = 531$) and percentages of respondents ranking 'guideline adherence' highly as a desirable quality measure

	n (%)	High rank for 'guideline adherence'	
Gender			$P = 0.90$
Female	27 (5)	33%	
Male	498 (95)	35%	
Years in practice			$P = 0.41$
<20	207 (39)	32%	
21–40	265 (50)	35%	
≥41	53 (10)	42%	
Specialty			$P = 0.68$
Surgeons	358 (70)	34%	
Internists/others	163 (30)	36%	
Number of rheumatoid arthritis patients (%)*			$P = 0.07$
≤10	104 (20)	41%	
11–20	100 (19)	34%	
21–30	82 (16)	40%	
31–40	54 (10)	37%	
≥41	184 (35)	27%	
Practice settings			$P = 0.34$
University hospital	62 (12)	29%	
Non-university hospital	205 (39)	33%	
Physician office	249 (47)	36%	
Other	11 (2)	55%	
Area type			$P = 0.66$
Metropolitan	149 (28)	34%	
Urban	318 (60)	33%	
Rural	61 (12)	39%	
Practice location			$P = 0.15$
Eastern Japan	218 (41)	37%	
Western Japan	313 (59)	31%	

*Does not add up to 100% because of rounding.
The following non-responding subjects were excluded: 6 subjects for gender and years in practice, 10 subjects for specialty, 4 subjects for practice setting, 7 subjects for number of rheumatoid arthritis patients and 3 subjects for area type.

unfavourable methods of assessment (mean rank of 6 and 6.1, respectively). The differences in mean rank between adjacently ranked methods were significantly different except for that between guideline adherence and voting by local physicians ($P = 0.14$) and between waiting time and reportage in the media ($P = 0.76$). This general ranking trend did not change on stratified analysis by working facility, place of practice, volume of patients or years in practice.

The exploration of factors related to the high ranking of guideline adherence in the unadjusted analysis is presented in the right columns of Table 1. None of the factors examined was associated with the high ranking of guideline adherence as a desirable quality measure. An exception was the number of rheumatoid arthritis patients for whom the respondent cares, which showed the non-significant trend that high-volume respondents with ≥41 regular patients were less likely to rank guideline adherence highly (overall $P = 0.07$). After adjustment for these factors using the logistic regression analysis, this group appeared significantly less likely to rank guideline adherence highly (odds ratio = 0.38 compared to the low-volume group with ≤10 regular patients, $P < 0.01$; Table 2). The Hosmer–Lemeshow test revealed that the model had a reasonable fit, with a *P*-value of 0.88.

Discussion

Our survey revealed a number of interesting points about Japanese rheumatology physicians' opinions on how the quality of care should be assessed. First, patient satisfaction is considered the best method of quality assessment, with this option ranked higher than other methods which target the technical aspects of care. The preference for this interpersonal quality over technical quality may indicate that the assessment of technical care is not considered to capture the 'true' technical quality of care. Alternatively, respondents may be reluctant to subject their practice to the physical and psychological intrusion of technical assessment.

Among the assessment methods targeting the technical aspect of care, the outcome measure of complication/admission rate was preferred over the process measure of guideline adherence. This finding stands in stark contrast to extensive use of process measures in western countries [4,7–9,11]. Because no nationwide quality assessment system for either outcome or process is implemented in Japan, the idea of using guidelines to assess quality may be difficult for the respondents to imagine. Furthermore, process measures used in practice are usually modified from the guideline recommendations themselves so that they can serve a measurement

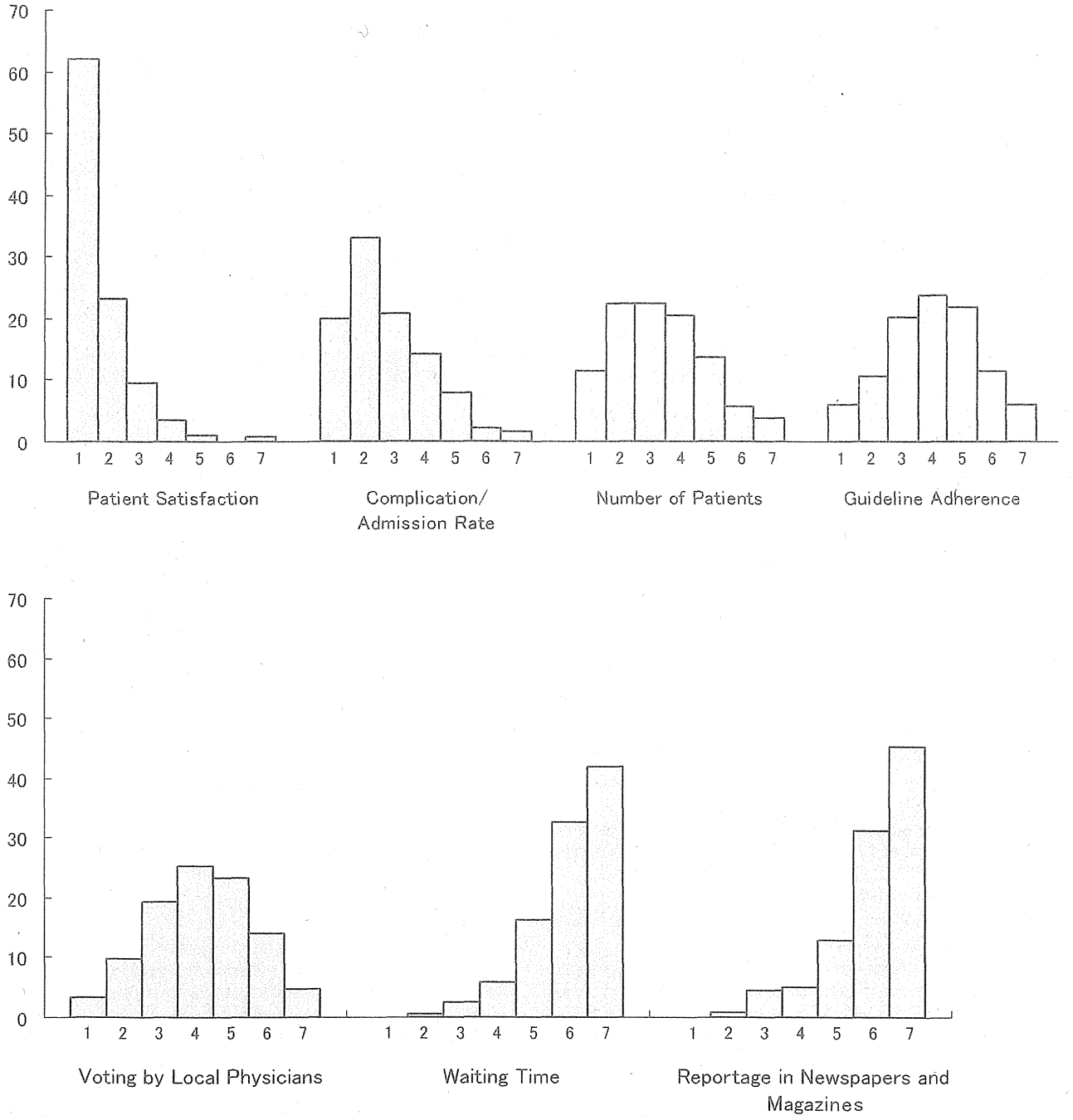


Figure 1 Ranking distribution of quality assessment methods for rheumatoid arthritis care.

purpose. Unfortunately, we suspect that the respondents had little experience or knowledge of process quality-of-care measurement, and concede that use of the term 'guideline adherence' to mean process measures in the questionnaire may have lacked precision. In any case, if guideline recommendations are to be used as quality indicators after pertinent modification, additional effort to convince physicians appears necessary, such as the convening of an expert panel specifically commissioned to examine the validity of each recommendation as a quality indicator.

A second interesting point is that 'number of patients' was considered preferable to guideline adherence as a quality assessment. Although the volume–outcome or volume–quality relationship has been extensively studied in surgical and medical conditions [23], we are unaware of any study showing that

Table 2 Respondent factors in relation to a higher (first to third) ranking of guideline adherence as the quality measure

		Odds ratio (95% CI)	*P*-value*
Sex	Female (vs male)	0.78 (0.33–1.99)	0.60
Years in practice	(vs <20)		0.18
	21–40	1.25 (0.82–1.89)	
	≥41	1.85 (0.95–3.60)	
Specialty	Surgeons (vs others)	0.72 (0.45–1.13)	0.16
Working facility	(vs university hospital)		0.40
	Non-university hospital	0.87 (0.44–1.71)	
	Physician office	0.97 (0.49–1.92)	
	Other	3.25 (0.64–16.45)	
Number of rheumatoid arthritis patients	(vs ≤10)		0.01
	11–20	0.61 (0.34–1.11)	
	21–30	0.80 (0.43–1.47)	
	31–40	0.60 (0.29–1.22)	
	≥41	0.38 (0.22–0.67)	
Area (%)	(vs metropolitan)		0.68
	Urban	0.97 (0.62–1.51)	
	Rural	1.27 (0.64–2.50)	
Practice location	Eastern Japan (vs western)	0.71 (0.47–1.07)	0.10

*Overall *P*-values for categories.

larger-volume providers of rheumatoid arthritis care produce better outcomes. Because the strength of the volume–outcome relationship varies across surgery types [24], future research should test the opinion of rheumatology physicians, as identified in this survey, that volume is a good proxy of quality of care, which leads to better rheumatic care outcomes, or is at least a better proxy than explicit guideline adherence.

Guideline adherence was ranked almost the same as 'voting by local physicians', which is a popular method used by the media. The Best Hospitals report published by the US News is one of the most famous examples [25]. In Japan also, several books have used physician voting to evaluate hospitals [16,26]. In a sense, guideline adherence can be viewed as an evaluation using explicit technical criteria, while voting by local physicians is a form of implicit review of quality, if appropriately performed. However, implicit review is known to be unreliable in the absence of detailed instruction and pertinent training of the reviewers, and bias due to suboptimal methodology is sometimes unclear.

Exploration of factors associated with the high ranking of guideline adherence revealed that high-volume respondents who care for >40 regular rheumatoid arthritis patients were less likely to rank guideline adherence highly. Because guidelines are sometimes criticized as 'too cookbook' [22], high-volume respondents who theoretically have more chance to care for atypical patients may feel less inclined to use guideline recommendations as quality standards. Other factors examined here were not significantly associated with ranking of guideline adherence.

Our results should be interpreted in view of several limitations. First, the survey was conducted among rheumatology physicians engaged in rheumatoid arthritis care, potentially limiting its generalizability to other conditions or types of physicians. Rheumatoid arthritis care is unique in that physicians need to select the most suitable of a wide range of anti-rheumatic medications (i.e. disease-modifying anti-rheumatic drugs) and biological agents and to fine-tune dosages to avoid adverse effects. This process is not only more complex and prolonged than that for most other common diseases, but also highly individualized, limiting the value of guideline recommendations. Second, the ranking of candidate assessment methods reveals relative preference only. The physicians may have thought that patient satisfaction is merely 'less bad' than even worse methods and ranked it highly on this basis alone. We chose ranking to focus on the difference between candidate methods, and expect that future research will examine absolute preference for these potential methods. Finally, potential differences between respondents and non-respondents may have biased the results. The overall survey was about guidelines, and respondents may have had a more favourable attitude to guidelines than non-respondents. Although guideline adherence was ranked about in the middle, non-respondents might have rated this item even lower.

Despite these limitations, we found that Japanese rheumatology physicians consider that patient satisfaction is the best method for quality assessment, and presently do not fully accept guideline adherence as a standard criterion of quality. Efforts to gain the support of quality monitoring systems focusing on process of care from practising physicians and enable their smooth introduction should focus on ways to construct convincing methods of assessing the technical quality of rheumatoid arthritis care.

References

1. Ess, S., Savidan, A., Frick, H., Rageth, C., Vlastos, G., Lutolf, U. & Thurlimann, B. (2010) Geographic variation in breast cancer care in Switzerland. *Cancer Epidemiology*, 34 (2), 116–121.
2. O'Hare, A. M., Rodriguez, R. A., Hailpern, S. M., Larson, E. B. & Kurella Tamura, M. (2010) Regional variation in health care intensity and treatment practices for end-stage renal disease in older adults. *JAMA*, 304 (2), 180–186.
3. Pilote, L., Califf, R. M., Sapp, S., Miller, D. P., Mark, D. B., Weaver, W. D., Gore, J. M., Armstrong, P. W., Ohman, E. M. & Topol, E. J. (1995) Regional variation across the United States in the management of acute myocardial infarction. GUSTO-1 Investigators. Global Utilization of Streptokinase and Tissue Plasminogen Activator for

Occluded Coronary Arteries. *The New England Journal of Medicine*, 333 (9), 565–572.
4. Jencks, S. F., Cuerdon, T., Burwen, D. R., Fleming, B., Houck, P. M., Kussmaul, A. E., Nilasena, D. S., Ordin, D. L. & Arday, D. R. (2000) Quality of medical care delivered to Medicare beneficiaries: a profile at state and national levels. *JAMA*, 284 (13), 1670–1676.
5. Riley, G. F., Warren, J. L., Potosky, A. L., Klabunde, C. N., Harlan, L. C. & Osswald, M. B. (2008) Comparison of cancer diagnosis and treatment in Medicare fee-for-service and managed care plans. *Medical Care*, 46 (10), 1108–1115.
6. Kressin, N. R. & Petersen, L. A. (2001) Racial differences in the use of invasive cardiovascular procedures: review of the literature and prescription for future research. *Annals of Internal Medicine*, 135 (5), 352–366.
7. McGlynn, E. A., Asch, S. M., Adams, J., Keesey, J., Hicks, J., DeCristofaro, A. & Kerr, E. A. (2003) The quality of health care delivered to adults in the United States. *The New England Journal of Medicine*, 348 (26), 2635–2645.
8. Wenger, N. S., Solomon, D. H., Roth, C. P., *et al.* (2003) The quality of medical care provided to vulnerable community-dwelling older patients. *Annals of Internal Medicine*, 139 (9), 740–747.
9. The US Department of Health and Human Services (2010) *Medicare hospital compare*. Available at: http://www.hospitalcompare.hhs.gov/ (last accessed 28 December 2010).
10. Cancer Care Ontario (2010) *Cancer systems quality index*. Available at: http://csqi.cancercare.on.ca/ (last accessed 28 December 2010).
11. Doran, T., Fullwood, C., Gravelle, H., Reeves, D., Kontopantelis, E., Hiroeh, U. & Roland, M. (2006) Pay-for-performance programs in family practices in the United Kingdom. *The New England Journal of Medicine*, 355 (4), 375–384.
12. Epstein, A. M., Lee, T. H. & Hamel, M. B. (2004) Paying physicians for high-quality care. *The New England Journal of Medicine*, 350 (4), 406–410.
13. Scott, I. A. (2008) Pay for performance programs in Australia: a need for guiding principles. *Australian Health Review*, 32 (4), 740–749.
14. The Supreme Court of Japan General Administrative Office (2006) *Status of medical malpractice cases processes and average litigation length* (in Japanese). Available at: http://www.courts.go.jp/saikosai/about/iinkai/izikankei/toukei_02.html (last accessed 28 December 2010).
15. NIKKEI_NET (2004) *Survey of operative performance in major hospitals* (in Japanese). Available at: http://health.nikkei.co.jp/hResearch/ (last accessed 15 September 2004).
16. Medical, N. (2004) Japanese Best Hospitals (Zenkoku Yuryou Byoin Ranking – Japanese). Tokyo: Nikkei BP.
17. Seddon, M. E., Marshall, M. N., Campbell, S. M. & Roland, M. O. (2001) Systematic review of studies of quality of clinical care in general practice in the UK, Australia and New Zealand. *Quality in Health Care*, 10 (3), 152–158.
18. Casalino, L. P., Alexander, G. C., Jin, L. & Konetzka, R. T. (2007) General internists' views on pay-for-performance and public reporting of quality scores: a national survey. *Health Affairs (Project Hope)*, 26 (2), 492–499.
19. Casalino, L. P. (1999) The unintended consequences of measuring quality on the quality of medical care. *The New England Journal of Medicine*, 341 (15), 1147–1150.
20. Nakayama, T. & Fukuhara, S. T. K. (2003) Contributions of clinical epidemiologists and medical librarians to developing evidence-based clinical practice guidelines in Japan : a case of the treatment of rheumatoid arthritis. *General Medicine*, 4 (1), 21–28.
21. Working Group on Development of Clinical Practice Guidelines for the Treatment of Rheumatoid Arthritis (Principal Investigator: Takahiro Ochi) (2004) A Manual for the Treatment of Rheumatoid Arthritis (in Japanese). Tokyo: Japan Rheumatism Foundation.
22. Higashi, T., Nakayama, T., Fukuhara, S., *et al.* (2010) Opinions of Japanese rheumatology physicians regarding clinical practice guidelines. *International Journal for Quality in Health Care*, 22 (2), 78–85.
23. Institute of Medicine (2000) Interpreting the Volume-Outcome Relationship in the Context of Health Care Quality: Workshop Summary. Washington, D.C.: National Academies Press.
24. Birkmeyer, J. D., Siewers, A. E., Finlayson, E. V., Stukel, T. A., Lucas, F. L., Batista, I., Welch, H. G. & Wennberg, D. E. (2002) Hospital volume and surgical mortality in the United States. *The New England Journal of Medicine*, 346 (15), 1128–1137.
25. US News & World Report (2006) *Best hospitals*. Available at: http://www.usnews.com/usnews/health/best-hospitals/tophosp.htm (last accessed 28 December 2010).
26. Yoshihara, S. (2008) The Best Doctors and Curing Hospitals (Saikouno Meii Plus Naoru Byoin – Japanese). Tokyo: Kodansha.

EXTENDED REPORT

Drug retention rates and relevant risk factors for drug discontinuation due to adverse events in rheumatoid arthritis patients receiving anticytokine therapy with different target molecules

Ryoko Sakai,[1,2] Michi Tanaka,[1,2] Toshihiro Nanki,[1,2] Kaori Watanabe[1,2] Hayato Yamazaki[1,2] Ryuji Koike,[1,2,3] Hayato Nagasawa,[4] Koichi Amano,[4] Kazuyoshi Saito,[5] Yoshiya Tanaka,[5] Satoshi Ito,[6] Takayuki Sumida,[6] Atsushi Ihata,[7] Yoshiaki Ishigatsubo,[7] Tatsuya Atsumi,[8] Takao Koike,[8] Atsuo Nakajima,[9] Naoto Tamura,[10] Takao Fujii,[11] Hiroaki Dobashi,[12] Shigeto Tohma,[13] Takahiko Sugihara,[14] Yukitaka Ueki,[15] Akira Hashiramoto,[16] Atsushi Kawakami,[17] Noboru Hagino,[18] Nobuyuki Miyasaka,[2,19] Masayoshi Harigai[1,2,3] for the REAL Study Group

► Additional supplementary data are published online only. To view these files please visit the journal online (http://ard.bmj.com/content/early/recent)

For numbered affiliations see end of article

Correspondence to
Masayoshi Harigai, Department of Pharmacovigilance, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo, 113-8519, Japan; mharigai.mpha@tmd.ac.jp

Received 29 September 2011
Accepted 25 February 2012
Published Online First
13 April 2011

ABSTRACT

Objective To compare reasons for discontinuation and drug retention rates per reason among anticytokine therapies, infliximab, etanercept and tocilizumab, and the risk of discontinuation of biological agents due to adverse events (AE) in patients with rheumatoid arthritis (RA).

Method This prospective cohort study included Japanese RA patients who started infliximab (n=412, 636.0 patient-years (PY)), etanercept (n=442, 765.3 PY), or tocilizumab (n=168, 206.5 PY) as the first biological therapy after their enrolment in the Registry of Japanese Rheumatoid Arthritis Patients for Long-term Safety (REAL) database. Drug retention rates were calculated using the Kaplan–Meier method. To compare risks of drug discontinuation due to AE for patients treated with these biological agents, the Cox proportional hazard model was applied.

Results The authors found significant differences among the three therapeutic groups in demography, clinical status, comorbidities and usage of concomitant drugs. Development of AE was the most frequent reason for discontinuation of biological agents in the etanercept and tocilizumab groups, and the second most frequent reason in the infliximab group. Discontinuation due to good control was observed most frequently in the infliximab group. Compared with etanercept, the use of infliximab (HR 1.69; 95% CI 1.14 to 2.51) and tocilizumab (HR 1.98; 95% CI 1.04 to 3.76) was significantly associated with a higher risk of discontinuation of biological agents due to AE.

Conclusions Reasons for discontinuation are significantly different among biological agents. The use of infliximab and tocilizumab was significantly associated with treatment discontinuation due to AE compared with etanercept.

Biological disease-modifying antirheumatic drugs (biological agents) are a standard treatment for rheumatoid arthritis (RA).[1 2] A number of clinical trials have demonstrated that biological agents significantly improve signs and symptoms of RA patients with both early and established disease, and that remission of RA can be achieved with biological agents not only in early RA patients, but also in established RA patients who have shown inadequate responses to conventional non-biological disease-modifying antirheumatic drugs (DMARD).

In Japan, six biological agents have been approved for the treatment of RA, infliximab in 2002, etanercept in 2005, tocilizumab and adalimumab in 2008, abatacept in 2010 and golimumab in 2011. These drugs are widely used in clinical practice according to treatment guidelines for biological agents by the Japan College of Rheumatology[3 4] and Japanese drug package inserts. Postmarketing surveillance and some clinical studies have shown short-term effectiveness and safety of these biological agents for Japanese RA patients.[5–8] The European League Against Rheumatism recommendations for the management of RA state that a tumour necrosis factor (TNF) antagonist should be administered as the first biological DMARD for patients who fail to respond to non-biological DMARD, including methotrexate,[9] whereas Japanese guidelines do not clearly specify the precedence of biological agents.

Some RA patients treated with biological agents are compelled to stop the administration of these drugs because of lack of efficacy (LOE), adverse events (AE), or financial reasons. In addition, some RA patients discontinue biological agents in the hope of a biological-free remission or biological-free low disease activity status.[10–12] In general, drugs with high retention rates have a good balance between long-term effectiveness and tolerability, reflecting the satisfaction of patients and doctors with the treatment. Because treatment for RA continues for many years or is life-long in the majority of patients, the examination of long-term drug retention rates using a prospective cohort study is important for the evaluation of biological agents.

To establish better treatment strategies for RA, it is important to identify reasons and risk factors causing the discontinuation of a drug, especially for biological agents. Several studies have shown that

Table 1 Characteristics of RA patients treated with infliximab, etanercept or tocilizumab at the start of the observation period

	Infliximab group (n=412)	Etanercept group (n=442)	Tocilizumab group (n=168)	p Value
Age, years	53.6±13.5	58.5±13.0	59.8±13.4	<0.001
Female, %	85.9	78.1	80.4	0.011
Disease duration, years	7.9±7.8	10.3±8.9	10.3±9.6	<0.001
Steinbrocker's class (3 or 4), %	24.8	31.2	27.4	0.108
Steinbrocker's stage (III or IV), %	43.9	57.0	46.4	<0.001
DAS28 (3/CRP)	4.5±1.2 (n=411)	4.5±1.3 (n=440)	5.1±3.4 (n=167)	0.056
Use of ≥3 previous non-biological DMARD, %	41.0	54.5	31.5	<0.001
Biological—naive, %	96.4	83.9	46.4	<0.001
Methotrexate use, %	99.3	44.6	44.0	<0.001
Methotrexate dose, mg/week	8.0±2.1	7.0±2.0	8.2±2.9	<0.001
Use of immunosuppressive drugs, except for methotrexate, %	1.9	5.7	14.9	<0.001
Oral corticosteroid use, %	68.9	73.1	60.1	0.008
Prednisolone-equivalent dose of corticosteroids (mg/day)	5.4±2.6	6.1±3.3	4.9±2.2	<0.001
Chronic pulmonary disease, %	22.6	36.7	40.5	<0.001
Diabetes mellitus, %	8.5	14.9	12.5	0.015

CRP, C-reactive protein; DAS28, disease activity score including 28-joint count; DMARD, disease-modifying antirheumatic drug; RA, rheumatoid arthritis

a frequent reason for the discontinuation of biological agents is the development of AE.[5–7 13–16] Mid to long-term tolerability of TNF inhibitors[6 13 14 16–24] and tocilizumab[7 15 25] has been reported, and some studies have directly compared drug retention rates among TNF inhibitors or between TNF inhibitors and other biological agents.[14 16 17 25–27] To summarise, infliximab had the lowest overall retention rate among infliximab, etanercept and adalimumab[14 16 17] and among infliximab, etanercept and anakinra.[26] A recent report from the CORRONA registry demonstrated the highest retention rate of infliximab compared with etanercept and adalimumab.[27] However, drug retention rates have not been compared between TNF inhibitors and the interleukin-6 receptor inhibitor, tocilizumab, in the real world. In addition, the risk factors causing drug discontinuation due to AE for patients given these biological agents have not been thoroughly evaluated.

The purpose of this study was to compare drug retention rates and reasons for discontinuation of infliximab, etanercept and tocilizumab among Japanese RA patients, and to investigate the association of the use of these biological agents and other clinical characteristics with drug discontinuation due to AE.

PATIENTS AND METHODS

Database

The Registry of Japanese Rheumatoid Arthritis Patients for Long-term Safety (REAL) is an ongoing prospective cohort established to investigate the long-term safety of biological agents in RA patients. Twenty-seven institutions participate, including 16 university hospitals and 11 referring hospitals. Details of REAL have previously been described.[28 29] Briefly, the criteria for enrolment in REAL include patients meeting the 1987 American College of Rheumatology criteria for RA, written informed consent, and starting or switching treatment with biological agents or starting, adding or switching non-biological DMARD at the time of enrolment in the study. Enrolment in the REAL database was started in June 2005 and closed in January 2012. To facilitate enrolment to the REAL registry, participating physicians were asked to enrol their patients already registered in postmarketing surveillance programmes previously implemented by pharmaceutical companies for biological agents.[5 8] In addition, our investigators were also encouraged to enrol as many patients as possible who fulfilled the inclusion criteria.[29]

Data were retrieved from the REAL database on 4 April 2011 for this study. The REAL study was approved by the ethics committees of the Tokyo Medical and Dental University Hospital and other participating institutions.

Data collection

Each patient's recorded baseline data included demography, disease activity, physical disability, comorbidities, treatments and laboratory data at the beginning of the observation period. A follow-up form was submitted by the site investigators every 6 months to the REAL data centre at the Department of Pharmacovigilance of Tokyo Medical and Dental University to report the occurrence of serious AE, current RA disease activity, treatments and clinical laboratory data.[28 29] We collected the Steinbrocker class[30] as the baseline measurement for each patient's physical disability, instead of the health assessment questionnaire disability index.[31] The investigators in each hospital confirmed the accuracy of their data submitted to the REAL data centre. The centre examined all the data sent by site investigators and sent queries if necessary to verify the accuracy of the data.

Patients

By April 2011, 2067 RA patients were registered in REAL, of these 1044 patients started treatment with infliximab, etanercept or tocilizumab at the time of enrolment or after enrolment in REAL. Four patients were excluded from this study because the reason for discontinuation of the initial biological agents was not identified. Eighteen patients who were enrolled in another clinical study requiring the discontinuation of infliximab were also excluded. We did not include patients who used adalimumab, abatacept or golimumab as the first biological agent in REAL because we did not have sufficient numbers of patients on adalimumab in the database (n=98) compared with infliximab and etanercept and had no patients given abatacept or golimumab in the database at the time our data were compiled. Our analysis included 412 patients who started infliximab, 442 patients who started etanercept and 168 patients who started tocilizumab.

Follow-up

For patients who initiated biological agents (infliximab, etanercept, or tocilizumab) at enrolment in REAL, the start date

Table 2 Reasons for drug discontinuation in RA patients treated with infliximab, etanercept or tocilizumab*

Reason for discontinuation	Infliximab (n=157)†	Etanercept (n=130)†	Tocilizumab (n=51)†
Adverse events	57 Cases (36.3%)	57 Cases (43.8%)	23 Cases (45.1%)
Infection	20 Cases (12.7%)	22 Cases (16.9%)	8 Cases (15.7%)
Pulmonary diseases except infection‡	7 Cases (4.5%)	7 Cases (4%)	3 Cases (5.9%)
Infusion reaction	6 Cases (3.8%)	NA	0 Case (0%)
Allergy except infusion reaction	7 Cases (4.5%)	12 Cases (9.2%)	6 Cases (11.8%)
Malignancy	6 Cases (3.8%)	3 Cases (2.3%)	1 Case (2%)
Cardiovascular system disease	2 Cases (1.3%)	2 Cases (1.5%)	2 Cases (3.9%)
Others	9 Cases (5.7%)	11 Cases (8.5%)	3 Cases (5.9%)
Lack of efficacy	68 Cases (43.3%)	47 Cases (36.2%)	23 Cases (45.1%)
Good control	21 Cases (13.4%)	7 Cases (5.4%)	2 Cases (3.9%)
Miscellaneous§	11 Cases (7.0%)	19 Cases (14.6%) §	3 Cases (5.9%)

The χ^2 test was applied to assess differences in the proportion of causes for discontinuation (ie, adverse event, lack of efficacy, good control and miscellaneous), and the adjusted residuals were calculated. A significant difference among the three groups (p=0.026) was observed. The adjusted residuals indicated that significantly higher percentages of patients in the infliximab group stopped the treatment due to good disease control compared with the other two groups (p<0.05).
*Values are the number (percentage) of patients who discontinued use because of each reason.
†Number of patients who discontinued their first biological DMARD for any reason.
‡Pulmonary diseases except for infection included interstitial pneumonia (three cases for infliximab, five for etanercept, two for tocilizumab) and other pulmonary diseases (four for infliximab, two for etanercept, one for tocilizumab).
§Miscellaneous reasons for drug discontinuation include patients' preference, financial reasons, and pregnancy.
DMARD, disease-modifying antirheumatic drugs; RA, rheumatoid arthritis.
NA, not applicable

of the observation period was the date these agents were first administered. For patients who started non-biological DMARD at the time of entry in REAL and who later started treatment with biological agents, the start of the observation period was the date of the first administration of biological agents in REAL. Observation was stopped either at 2.5 years after the start of the observation period, on the date of death of a patient, loss to follow up, enrolment in clinical trials, or when therapy was stopped with the first biological agent in REAL for more than 90 days, or on 4 April 2011, whichever came first. The period following switching to a second biological agent was excluded from this study. We defined termination of treatment with biological agents as stopping treatment with the agent for more than 90 days. The date of the last administration of each biological DMARD was retrieved from medical records and reported by the site investigators. Reasons for drug discontinuation were obtained from case report forms of REAL supplemented by medical records, if necessary, and classified into AE, good control, LOE or miscellaneous. We did not discriminate between a primary and secondary LOE. Note that we collected only serious AE in REAL, but also collected AE in this study if it was the main reason for the discontinuation of a biological agent. When a patient had two or more reasons for drug discontinuation, site investigators assigned precedence and we used the primary reason contributing to drug discontinuation for that patient.

Statistical analysis

The primary outcome of this study was the investigation of the association of the use of infliximab, etanercept and tocilizumab with drug discontinuation due to AE. We also sought to identify other risk factors for drug discontinuation due to AE. Drug retention rates were calculated by the Kaplan–Meier method and compared using the log-rank test among groups. For univariate analysis, the χ^2 test was used for comparison of categorical variables and the Kruskal–Wallis test was used for continuous variables among the three agents. For multivariate analysis, the Cox regression hazard model with the forced entry method was employed to compare risks for drug discontinuation due to AE. The validity of the proportional hazards assumption was confirmed by the log-minus-log survival function. We followed the STROBE statement[32] for clear reporting except for 'the number and reasons for non-participation' in this study.

These statistical analyses were conducted using SPSS (version 16.0Illinois,). All p values were two-tailed and p<0.05 was considered statistically significant.

RESULTS

Baseline characteristics of the patients

This analysis included 412 patients in the infliximab group (636.0 patient-years (PY)), 442 in the etanercept group (765.3 PY) and 168 in the tocilizumab group (206.5 PY). Table 1 shows the baseline characteristics of the groups. There were significant differences in age, gender, disease duration and clinical status of the patients. The etanercept and tocilizumab groups had longer disease duration (p<0.001) and higher percentages of comorbidities than the infliximab group (p<0.001 for chronic pulmonary disease, p=0.011 for diabetes mellitus). The rates of biological-naive patients (96.4% for the infliximab group, 83.9% for the etanercept group and 46.4% for the tocilizumab group) (p<0.001) and of the use of three or more non-biological DMARD (p<0.001) in the tocilizumab group were the lowest among the three groups. The rate of the use (p=0.007) and dose (p<0.001) of oral corticosteroids of the etanercept group were higher than those for the other two groups. Disease activity did not differ significantly among the groups.

Occurrence of treatment termination

The median IQR of the observation period for each group was 1.50 (0.74–2.50) years for the infliximab group, 2.1 (0.98–2.50) years for the etanercept group and 1.0 (0.5–2.0) years for the tocilizumab group. The number of patients who discontinued biological agents for any reason during the observation period was 157 (38.1%) for the infliximab group, 130 (29.4%) for the etanercept group and 51 (30.4%) for the tocilizumab group (p=0.019 by χ^2). Table 2 shows the reasons for drug discontinuation for each group. A significant difference among the three groups (p=0.026 by χ^2) was seen in the proportions of reasons for discontinuation, and the adjusted residuals indicated that significantly higher percentages of patients in the infliximab group stopped treatment due to good disease control compared with the other two groups (p<0.05). The most frequently reported