Toxicity of anti-rheumatic drugs in a randomzed clinical trial of early rheumatoid arthritis

van Jaarsveld, C.H.M.; Jahangier, Z.N.; Jacobs, J.W.G.; Blaauw, A.A.M.; van Albada-Kuipers, G.A.; Ter Borg, E.J.; Brus, H.J.; Schenk, Y.; van der Veen, M.J.; Bijlsma, J.W.J.

Published in:
Rheumatology

Rheumatology 2000;39:1374–1382

Toxicity of anti-rheumatic drugs in a randomized clinical trial of early rheumatoid arthritis

C. H. M. van Jaarsveld, Z. N. Jahangier, J. W. G. Jacobs, A. A. M. Blaauw, G. A. van Albada-Kuipers[1], E. J. ter Borg[2], H. L. M. Brus[3], Y. Schenk[4], M. J. van der Veen[5] and J. W. J. Bijlsma, on behalf of the Rheumatic Research Foundation, Utrecht, The Netherlands

Department of Rheumatology and Clinical Immunology, University Medical Center, PO Box 85500, 3508 GA Utrecht, [1]Department of Rheumatology, Eemland Hospital, PO Box 1505, 3800 BM Amersfoort, [2]Department of Rheumatology, St Antonius Hospital, Koekoekslaan 1, 3435 CM Nieuwegein, [3]Department of Rheumatology, Hospital Hilversum, van Riebeeckweg 212, 1213 XZ Hilversum, [4]Department of Rheumatology, Diakonessen Hospital, Bosboomstraat 1, 3582 KE Utrecht and [5]Department of Rheumatology, St Jansdal Hospital, PO Box 138, 3840 AC Harderwijk, The Netherlands

Abstract

Objective. To evaluate the toxicity of slow-acting anti-rheumatic drugs (SAARDs) and non-steroidal anti-inflammatory drugs (NSAIDs) in early rheumatoid arthritis.

Methods. Patients were randomized to receive a SAARD—hydroxychloroquine (HCQ; $n=120$), i.m. gold ($n=114$) or methotrexate (MTX; $n=118$)—or a NSAID only ($n=67$). Patients in the three SAARD groups were allowed to take NSAIDs. Follow-up included 545 patient-years (p-yr). Adverse effects were attributed to specific medications using the Naranjo scoring method.

Results. Fifty-five per cent of the patients suffered from adverse effect(s). Adverse effects were most common during i.m. gold therapy (87 per 100 p-yr), which led to permanent discontinuation of this treatment in 31 cases. The incidences of adverse effects that were probably attributable to NSAIDs in patients treated simultaneously with a SAARD were similar for the three SAARD groups. The mean period until the first adverse effect was longer in the MTX group (39 weeks) than in the HCQ group (27 weeks). Baseline clinical and sociodemographic parameters were not predictive of the occurrence of adverse effects.

Conclusion. No adverse effect could be classified as definitely related to either SAARDs or NSAIDs by the Naranjo scoring method. The incidence of possible adverse effects of NSAIDs and SAARDs was 72 per 100 p-yr, and adverse effects led to permanent discontinuation of the therapy in 56 cases (13%) (31 patients receiving i.m. gold, 12 receiving MTX, 10 receiving HCQ and three receiving NSAID only).

KEY WORDS: Toxicity, Anti-rheumatic drugs, Early RA, Naranjo scale.

Rheumatoid arthritis (RA) exhibits a chronic fluctuating course which, if left untreated, results in most cases in progressive joint destruction, disability and premature death. A limited number of anti-rheumatic drugs are available. Aggressive drugs may cause more frequent and serious side-effects than less aggressive drugs. However, patients with a poor prognosis who develop major damage within a short period might benefit from aggressive treatment, while patients with less severe disease might benefit from less aggressive drugs with milder or fewer side-effects. It is obvious that a balance has to be found between joint destruction and disability due to RA and the toxic effects of treatment. Reading the literature on clinical drug trials, one gets the impression that studies comparing the effectiveness of several anti-rheumatic drugs have been performed more often than studies on toxicity. Comparison of the toxicity of

Submitted 7 January 2000; revised version accepted 10 July 2000.

Correspondence to: J. W. G. Jacobs, Department of Rheumatology and Clinical Immunology, University Medical Center, PO Box 85500, 3508 GA Utrecht, The Netherlands.

medication in observational studies may lead to biased conclusions due to confounding by indication, because both the choice of a specific drug and the occurrence of side-effects may be dependent on disease severity. Therefore, randomized trials are to be preferred when toxicity is studied [1, 2].

In the Utrecht Rheumatoid Arthritis Cohort, patients with recent-onset RA were randomized to undergo treatment with non-steroidal anti-inflammatory drugs (NSAIDs) only, or with one of three slow-acting anti-rheumatic drugs (SAARDs)—hydroxychloroquine (HCQ), intramuscular gold (i.m. gold) and methotrexate (MTX). The patients in the last three groups were allowed to take NSAIDs as well. Data from this randomized trial were used to compare the toxicities of the different drugs. Information on medication, adverse effects and clinical disease activity was documented in detail every 3 months, which enabled us to study several aspects of toxicity.

The objective of this study was to make an inventory of adverse effects of NSAIDs, HCQ, i.m. gold and MTX and to compare the toxicities of these drugs. Furthermore, the intervals until the occurrence of adverse effects were studied, and sociodemographic and baseline clinical variables related to the occurrence of adverse effects were identified.

Methods

Patients

The study was designed as a prospective, open-label, randomized, controlled trial. Since 1990 all patients who met the 1987 American College of Rheumatology classification criteria for RA [3] at six rheumatological centres in the Utrecht region of The Netherlands were asked to participate in a randomized trial to compare therapeutic strategies [4]. Disease duration had to have been less than 1 yr; most patients were enrolled shortly after diagnosis. All recently diagnosed RA patients seen by rheumatologists in the Utrecht region were asked to participate, and data were also collected on the 82 patients who refused to be randomized. Therefore, this study was representative of early RA patients referred to hospital (rheumatologists) in The Netherlands (i.e. it was a hospital-based study). The following exclusion criteria were applied: (i) age <17 yr; (ii) comorbid conditions that might interfere with the therapeutic strategies; (iii) previous or current use of SAARDs, glucocorticosteroids, cytotoxic or immunosuppressive drugs; (iv) pregnancy or breast feeding; (v) mental disturbances that would make adherence to the study protocol unlikely. All patients gave informed consent. The study design was approved by the ethics committees of all participating hospitals.

Treatment

Patients entering the study were assigned randomly to one of four treatment groups: (i) NSAID only; (ii) HCQ (400 mg/day); (iii) i.m. gold (aurothioglucose, 50 mg/week); (iv) oral MTX (7.5–15 mg/week). All randomization was done by drawing sealed envelopes from blocks of 100 with equal numbers of patients for each of the four treatments per hospital. In group (i), patients received only NSAIDs, the dose and type of which could be modified at any time. Follow-up of the patients in this group was included in the analysis until a SAARD was administered. Use of NSAIDs by patients of the three SAARD groups was allowed. Follow-up of patients in the SAARD groups was included in the analysis until treatment with a SAARD other than the randomized therapy was started. In this way the toxicities of the original randomly assigned strategies could be compared. The use of pure analgesics was allowed in all groups; the use of oral glucocorticosteroids and intra-articular steroid injections was recorded but was avoided if possible. The assigned medication was continued unless adverse reactions or ineffectiveness necessitated discontinuation. Criteria for the discontinuation or dose adjustment of a SAARD because of adverse reactions were described in the study protocol. Analysis of effectiveness after 1 yr of follow-up showed that patients who had been randomized to receive only NSAIDs exhibited greater disease activity than patients who had been randomized to receive a SAARD [4]. Therefore, in 1994 the NSAID-only arm of the trial was discontinued. As a consequence, there were fewer patients in the NSAID-only group.

The primary aim of the clinical trial in the Utrecht Rheumatoid Arthritis Cohort was to compare the effectivenesses of therapeutic strategies. The present paper reports a comparison of the toxicities of the therapeutic strategies, which was the secondary aim of the trial.

Toxicity

Assessments were performed at the start, every 3 months during the first 2 yr and subsequently every 6 months. Clinical variables, including disease activity variables, medication and the occurrence of adverse effects, were assessed by the same physician or research nurse for each patient on each occasion. Safety investigations included clinical examination and laboratory abnormalities. All adverse experiences were reported, regardless of their relationship to the anti-rheumatic therapy, and are referred to in this report as 'adverse effects'. The term 'side-effect' was avoided since this connotes unconditional attribution to a specific drug, which would be inappropriate. An adverse effect was defined as 'severe' when the SAARD or all NSAIDs were discontinued permanently, 'moderate' when the SAARD or NSAID was temporarily stopped, the SAARD dose was adjusted or the patient switched to another NSAID, and 'mild' when no action regarding medication was taken.

The toxicity profiles of SAARDs and NSAIDs overlap [5]. Since a reported symptom could also be caused by the disease itself, a concomitant illness or other drugs taken at the same time, a causal link between a symptom and a drug is difficult to identify. Therefore, we used the adverse drug reaction scale of Naranjo *et al.* [6] (see Appendix 1) to classify the

probability that an adverse event was related to NSAID or SAARD therapy. If an adverse effect occurred during treatment with both a NSAID and a SAARD, the relationship was assessed for each of the two drugs separately. This scoring method classifies each adverse reaction as definitely, probably, possibly or doubtfully attributable to a drug treatment. The answers to items 6 and 7 of this rating scale were always unknown in the present study; consequently the scores ranged from −3 to 11. Adverse reactions classified as definite (score ≥9) or probable (score 5–8) were studied in more detail.

Statistics

Incidences of adverse effects per 100 patient-years (p-yr) of drug exposure and periods until the first adverse effect were assessed for each treatment strategy. In addition, we studied incidences according to the severity of adverse effects and the probability of a relation to specific drug use. The χ^2-test, Student's *t*-test and the Mann–Whitney *U*-test were used to relate sociodemographic and baseline clinical variables to the occurrence of adverse effects. A *P* value of less than 0.05 was considered statistically significant.

Results

Patient characteristics

A total of 501 RA patients were eligible for the study, 82 of whom refused to be randomized for various reasons. Of the 419 randomized patients, 65% were female. At the start of the study the mean age was 57 yr (S.D. = 14); disease duration was less than 1 yr; 61% of the patients had a positive rheumatoid factor (RF) test and 70% of the radiographs showed no evident damage (Sharp score ≤4) [7]. At baseline, 67 patients (16%) were randomized to receive treatment with NSAIDs only, 120 patients (29%) treatment with HCQ, 118 (28%) treatment with i.m. gold and 114 (27%) treatment with MTX. Baseline clinical variables (Table 1) and sociodemographic variables did not differ significantly between the randomized groups.

Folic acid (0.5 mg daily) was used at some time during follow-up by 41 (36%) of the 114 patients receiving MTX. Oral prednisone (up to 30 mg daily) was used by 20 of the randomized patients (5%) at some time during follow-up; another three patients received high (200 mg) i.v. doses of dexamethasone for 3 days. At some time during follow-up 18% of the patients were treated with antacids (i.e. antacids in the broad sense; used to protect the stomach or to treat symptoms).

NSAID and SAARD use

The study encompassed 545 p-yr of drug exposure. During the study period, 99% of the patients used NSAIDs at some time. Treatment with the original randomly assigned therapy was continued for a mean period of 67 weeks (1.3 yr, range 0–5.5 yr). Four patients never started the randomized therapy: one patient in the NSAID group was started on a SAARD at baseline and three patients in the i.m. gold group refused injections. Data on medication and adverse effects are listed in Table 1. The NSAID-only strategy was continued for a mean period of 55 weeks, HCQ for 59 weeks, i.m. gold for 62 weeks and MTX for 90 weeks. This resulted in a total drug exposure of 71 p-yr for the NSAID-only group, 136 p-yr for the HCQ group, 141 p-yr for the i.m. gold group and 197 p-yr for the MTX group.

TABLE 1. Disease activity at baseline, medication and adverse effects

	Treatment			
	NSAID only	HCQ (+NSAID)	I.m. gold (+NSAID)	MTX (+NSAID)
No. patients	67	120	118	114
Baseline disease activity (actual range) [mean (S.D.)]				
Disability (0–3)[a]	1.3 (0.7)	1.4 (0.8)	1.3 (0.8)	1.3 (0.8)
Joint score (0–509)[b]	138 (106)	142 (98)	144 (100)	145 (101)
Pain (0–100)[c]	43 (27)	46 (28)	42 (28)	45 (28)
ESR (0–140)[d]	40 (28)	42 (28)	40 (28)	43 (27)
Weeks of treatment [mean (range)]	55 (0–245)	59 (4–262)	62 (0–286)	90 (1–287)[f]
Total exposure (p-yr)	71	136	141	197
Number of adverse effects[e]	34	91	122	144
Patients using NSAIDs (%)	100	99	98	97
Patients with at least one adverse effect (%)	34	49	65	64[g]
Weeks until first adverse effect [mean (range)]	28 (9–56)	27 (9–95)	35 (9–104)	39 (9–234)[h]

[a]Functional disability measured with a validated Dutch version of the Health Assessment Questionnaire [19].
[b]According to Thompson (i.e. assessment of simultaneous presence of joint tenderness and swelling in a selection of joints weighed for joint size) [20, 21].
[c]Measured on a 100-mm visual analogue scale.
[d]Westergren method, mm 1st h. High values indicate more active disease or more disability for all parameters[a–d].
[e]Adjusted frequencies for differences in drug exposure are presented in Table 3.
[f]MTX group differed significantly from each of the other three groups (ANOVA, $P < 0.001$).
[g]NSAID-only and HCQ groups differed significantly from the i.m. gold and MTX groups (χ^2, $P < 0.0001$).
[h]MTX group differed significantly from HCQ group (*t*-test, $P = 0.02$).

Reasons for stopping the assigned medication were adverse effects, ineffectiveness and, in a minority of cases, remission. The observation that MTX was used for a longer period than all the other therapies (ANOVA, $P < 0.001$) indicates a better toxicity profile and/or better effectiveness.

Occurrence of adverse effects

In total, 232 of the 419 patients (55%) suffered 391 adverse effects. The percentage of patients with at least one adverse effect was 65% for the i.m. gold group and 64% for the MTX group, compared with 49% for the HCQ-treated patients and 34% for the patients on NSAIDs only (χ^2-test, $P < 0.0001$). The mean period until the occurrence of the first adverse effect ranged from 27 weeks for the HCQ-treated patients to 39 weeks for the MTX-treated patients (*t*-test, $P = 0.02$) (Table 1). Details of all 391 adverse effects, regardless of their association with the drug, are presented in Table 2 for all four groups. Of these adverse effects, 305 (78%) were reported during clinical visits, 56 (14%) were detected by laboratory tests and 30 (8%) by a combination of the two. Some of the findings are discussed in detail below.

Diarrhoea occurred in 10 patients, five of whom received HCQ. In two of these cases both HCQ and NSAIDs were discontinued permanently, whereas in three cases NSAIDs were continued and the HCQ dose was reduced to 200 mg/day. In the other five cases of diarrhoea, NSAIDs, i.m. gold and MTX were continued.

In total, 77 subjective gastrointestinal (GI) complaints (diarrhoea excluded) were reported, 34 of which occurred in the MTX group (17 per 100 p-yr), 21 in the HCQ group (15 per 100 p-yr), 11 in the i.m. gold group (8 per 100 p-yr) and 11 in the NSAID-only group (15 per 100 p-yr). Although most patients in the three SAARD groups used both a SAARD and a NSAID, incidences did not differ much from that in the NSAID-only group. These subjective GI complaints caused discontinuation of HCQ in two cases and MTX in another two cases, but not of i.m. gold; NSAIDs were stopped permanently in eight cases (four in the HCQ group, one in the MTX group and three in the NSAID-only group).

There were five patients with objective GI symptoms confirmed by endoscopy: colitis ($n = 1$), gastric ulceration ($n = 1$) and gastritis ($n = 3$). Endoscopy was not performed routinely as part of the study. The incidence was 0.9 per 100 p-yr of anti-rheumatic drug exposure. All five patients received NSAIDs in combination with HCQ ($n = 1$), i.m. gold ($n = 1$) or MTX ($n = 3$). In all three cases of gastritis NSAIDs were stopped and antacids were started. The one case of gastric ulceration occurred simultaneously with pneumonitis during MTX and NSAID treatment; MTX was stopped and NSAID therapy was continued, in combination with antacids.

Proteinuria (≥ 0.1 g/l in 24 h) was found in 36 patients, 16 of whom were receiving i.m. gold. Proteinuria (1.1–3.4 g/l) led to permanent discontinuation of i.m. gold in five cases. In one case, i.m. gold was stopped permanently because of a rash and proteinuria (0.3 g/l), and in one case of proteinuria (0.6 g/l) i.m. gold was stopped temporarily. The other nine cases of proteinuria during i.m. gold treatment and the 20 cases during treatments other than i.m. gold were all mild (≤ 0.3 g/l) and therapy was continued.

Elevated serum creatinine (>120 μmol/l) was detected in 15 patients (six persistent and nine single laboratory abnormalities). In three cases, i.m. gold was discontinued temporarily, due partly to elevated serum creatinine and partly to proteinuria and mucocutaneous reactions, which occurred simultaneously. For two MTX-treated patients the dosage was reduced. NSAID treatment was stopped permanently in one case in the i.m. gold group. In three cases, NSAID therapy was stopped temporarily or changed to another NSAID (i.e. one in the NSAID-only, one in the HCQ and one in the MTX group). In six cases (one mild persistent and five single elevations) no action was taken.

The occurrence of multiple rheumatoid nodules was not investigated systematically, but was reported spontaneously by two patients, both of whom were on MTX therapy.

The incidence of adverse effects per 100 p-yr of drug exposure is presented in Table 3. Analysis of all events (irrespective of medication and the severity of the events) showed that the highest incidence was for patients on i.m. gold (87 events per 100 p-yr of i.m. gold exposure) followed by 73 per 100 p-yr for patients on MTX, 67 per 100 p-yr for patients on HCQ and 48 per 100 p-yr for patients on NSAIDs only.

NSAID- or SAARD-related adverse effects

The likelihood that the adverse event was related to NSAIDs or SAARD was classified according to Naranjo *et al.* [6] (Appendix 1). The distribution of categories of association between the adverse effect and SAARD use was as follows: definite, 0 cases; probable, 60 cases; possible, 292 cases; doubtful, 5 cases. Thirty-four adverse effects were unrelated to SAARD because only NSAIDs were taken in these cases. The association with NSAID use was also never classified as definite, but 41 associations were probable, 299 possible and six doubtful; 45 events were unrelated to NSAID use because only a SAARD was taken at the time of the adverse reaction. No adverse effect could be classified as being definitely associated with either SAARD or NSAID use. The 60 events classified as probably related to SAARD use were studied more closely. Table 3 shows that the lowest incidence was found for HCQ therapy (6 per 100 p-yr) compared with 15 per 100 p-yr for MTX and 16 per 100 p-yr for i.m. gold. All adverse effects were also scored according to the method of Naranjo *et al.* to assess their relationship with NSAIDs (Table 3). Twenty-three of the 34 adverse effects which occurred in the NSAID-only group were classified as definitely or probably NSAID-related. In the three SAARD groups, the incidence of NSAID-related adverse effects ranged from 2 to 6 per 100 p-yr. This indicates that NSAID-related toxicity was similar in the three SAARD groups. Only 18 of the adverse effects that occurred in the three

TABLE 2. Clinical symptoms and laboratory abnormalities that occurred in 419 RA patients after a total of 545 p-yr of anti-rheumatic drug use

Adverse effect	Treatment			
	NSAID only	HCQ (+NSAID)	I.m. gold (+NSAID)	MTX (+NSAID)
Gastrointestinal (subjective)	12	26	12	37
Nausea, vomiting, dyspepsia, abdominal pain, indigestion	11	21	11	34
Diarrhoea	1	5	1	3
Gastrointestinal (objective)	0	1	1	3
Gastric ulceration	–	–	–	1
Gastritis	–	–	1	2
Colitis	–	1	–	–
Mucocutaneous	2	14	50	15
Rash	2	10	28	3
Stomatitis, mouth ulcers	–	1	11	4
Alopecia	–	1	3	8
Pruritus	–	1	8	–
Photosensitivity	–	1	–	–
Central nervous system	4	2	4	10
Headache, dizziness, tinnitus, mood alterations	2	1	3	8
Confusion	–	–	–	1
Collapse	–	–	–	1
Cerebral haemorrhage (paralytic)	1	–	1	–
Concentration disturbances	1	1	–	–
Renal	4	12	22	20
Proteinuria ($\geq$0.1 g/l in 24 h)	2	8	16	10
Elevated serum creatinine (>120 μmol/l)	2	3	5	5
Oedema (pretibial)	–	1	1	2
Hypertension	–	–	–	2
Haematuria	–	–	–	1
Hepatotoxicity	5	1	3	20
Elevated transaminases or γGT (ALAT > 90 U/l, GT > 100 U/l)[a]	5	1	3	20
Haematological	1	3	12	8
Anaemia (Hb < 6.5 mmol/l)	–	2	–	1
Leukopenia (L < 3.5×10^9/l)	–	–	5	4
Thrombocytopenia (< 150×10^9/l)	–	–	3	–
Eosinophilia (> 0.5×10^9/l)	1	1	2	2
Pancytopenia	–	–	1	1
Leukopenia and thrombocytopenia	–	–	1	–
Respiratory system	0	0	0	3
Severe pulmonary disorder (other than pneumonitis)	–	–	–	1
Pneumonitis	–	–	–	1
Persistent cough	–	–	–	1
Disorders of eye or ear	3	5	1	8
Retinopathy	–	–	–	1
Cataract	–	–	–	1
Disturbed vision (unspecified)	–	2	–	2
Peri-orbital oedema	–	2	–	–
Glaucoma	–	–	1	–
Blurred vision	2	1	–	–
Toxic keratitis	1	–	–	1
Dry eyes/dry mouth	–	–	–	1
Impaired hearing	–	–	–	2
Other	3	27	17	20
Fever, infections	3	19	9	13
Neuropathy	–	1	1	1
Restless legs	–	1	–	–
Lymphadenopathy	–	2	–	–
Heart failure	–	–	1	1
Fatigue	–	2	1	–
Vertebral fracture, osteoporosis	–	1	1	1
Malignancy	–	1	2	2
Rheumatoid nodules	–	–	–	2
Diabetes mellitus	–	–	1	–
Sexual impotence	–	–	1	–
Total	34	91	122	144
Total p-yr of exposure to drug	71	136	141	197

[a]ALAT, alanine aminotransferase.

SAARD groups were classified as probably related to NSAID use, which is low in comparison with the 23 events classified as such in the NSAID-only group. This is a direct consequence of the Naranjo scoring method, as the probability decreases if another drug (the SAARD) could also be responsible for the adverse effect (Appendix 1, item 5). Correspondingly, the probability that an adverse event is related to a SAARD decreases when NSAIDs are taken simultaneously.

Discontinuation of the randomly assigned medication

The definition of the severity of adverse events depended on whether the therapy was stopped or the dose was adjusted. Adverse effects were classified as severe in 17%, moderate in 49% and mild in 34% of the cases. Table 4 shows that, although adverse events occurred frequently during MTX treatment, most of them were classified as mild or moderate. In contrast, 25% of the adverse effects occurring during treatment with i.m. gold resulted in permanent discontinuation of this treatment ($n = 31$). NSAIDs were permanently stopped by 12 patients (3% of 415), three of whom were in the NSAID-only group, five in the HCQ group, two in the i.m. gold group and two in the MTX group. Reasons for stopping NSAIDs included subjective GI complaints ($n = 7$), gastric ulceration ($n = 1$), gastritis ($n = 1$), pretibial oedema ($n = 1$), a combination of anaemia and pancytopenia ($n = 1$), and elevated transaminases ($n = 1$). SAARD therapy was stopped permanently in 53 patients (15% of 352).

Figure 1 shows the frequency of permanent discontinuation of treatment due to adverse effects over a 4-yr period. The total number of patients in follow-up decreased with increasing years of follow-up, as is shown in the table accompanying Fig. 1. Therefore, cumulative absolute numbers are presented in Fig. 1. The percentage of patients in follow-up was similar for the three SAARD groups at 2.5 yr, enabling comparison between the SAARD groups. In the i.m. gold group, 31 patients stopped gold therapy permanently because of adverse effects, compared with 12 patients on MTX and 10 on HCQ therapy. In the i.m. gold group, these adverse effects included mucocutaneous reactions ($n = 21$), proteinuria ($n = 6$) and haematological problems

Table 3. Incidence of adverse events per 100 p-yr

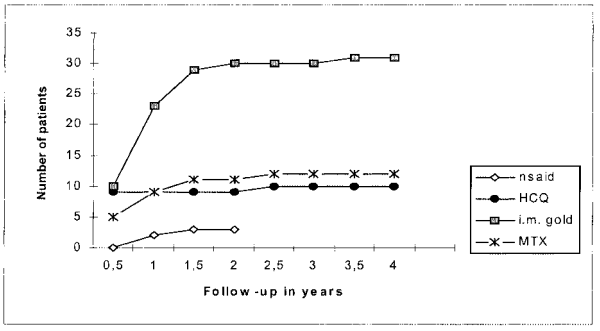
Drug use at time of symptom occurrence[a]	No. of events	Years at risk	Incidence per 100 p-yr
All adverse effects			
NSAID only	34	71	48
HCQ	91	136	67
I.m. gold	122	141	87
MTX	144	197	73
All adverse effects	391	545	72
Adverse effects definitely or probably related to current SAARD use[b]			
NSAID only	–	–	–
HCQ	8	136	6
I.m. gold	22	141	16
MTX	30	197	15
All SAARD-related events	60	545	11
Adverse effects definitely or probably related to current NSAID use[b]			
NSAID only	23	71	32
HCQ	6	136	4
I.m. gold	8	141	6
MTX	4	197	2
All NSAID-related events	41	545	8

[a]98% of patients in the SAARD groups also used NSAIDs at some time during follow-up.
[b]Classification according to Naranjo *et al.* [6] (see Appendix 1); only the adverse events classified as definitely or probably related to current SAARD or NSAID use are presented.

Table 4. Severity of adverse events that occurred in 419 RA patients after a total of 545 p-yr of anti-rheumatic drug use

	Number (% of all adverse events for the drug)			
Severity	NSAID only	HCQ (+NSAID)	I.m. gold (+NSAID)	MTX (+NSAID)
Severe				
SAARD stopped permanently	–	10 (11)	31 (25)	12 (8)
NSAIDs stopped permanently	3 (9)	5 (6)	2 (2)	2 (1)
Moderate[a]	19 (56)	44 (48)	54 (44)	74 (51)
Mild[b]	12 (35)	32 (35)	35 (29)	56 (39)
Total no. of adverse effects	34 (100)	91 (100)	122 (100)	144 (100)
Total p-yr of exposure to drug	71	136	141	197

[a]SAARD or NSAID stopped temporarily, SAARD dose adjusted, or change in NSAID.
[b]No action regarding SAARD and NSAID.



Total number of patients (%) followed in each medication group

	Baseline	1 year	2.5 years	4 years
NSAID	67 (100%)	29 (43%)	4 (6%)	1 (1%)
HCQ	120 (100%)	94 (78%)	61 (51%)	55 (46%)
i.m. gold	118 (100%)	91 (77%)	67 (57%)	60 (51%)
MTX	114 (100%)	96 (84%)	60 (53%)	46 (40%)

FIG. 1. Cumulative number of patients who permanently discontinued the randomized medication due to adverse effects.

($n = 4$). Elevated transaminases or γ-GT ($n = 4$) was the most common adverse effect and resulted in the permanent discontinuation of MTX; less common were GI complaints ($n = 2$), headache ($n = 2$), keratitis ($n = 1$), disturbed vision ($n = 1$), rheumatoid nodules ($n = 1$) and pneumonitis ($n = 1$). HCQ was discontinued because of GI complaints ($n = 5$), mucocutaneous problems ($n = 4$) and disturbed vision ($n = 1$).

Variables related to the occurrence of adverse effects

Sociodemographic and baseline clinical variables of patients with one or more adverse effects ($n = 232$) were compared with those of patients without adverse effects ($n = 187$). No differences in age, gender or RF status at baseline were found. The baseline clinical variables [including erythrocyte sedimentation rate (ESR), C-reactive protein, joint score, pain, functional disability, morning stiffness, radiological damage, haemoglobin, thrombocyte count and general well-being] did not differ between patients with and without adverse effects.

Discussion

In total 391 adverse effects were detected during clinical visits or laboratory tests in a cohort of 419 patients with recent-onset RA and 545 p-yr of anti-rheumatic drug exposure. In contrast with many other studies, these data represent toxicity in recent-onset RA patients who were treated with a SAARD for the first time or who had started only recently with a NSAID. At baseline, patients were randomized to treatment with NSAIDs only, HCQ, i.m. gold or MTX. The mean period until the occurrence of the first adverse effect was longest for MTX-treated patients (39 weeks). MTX was used for a longer mean period (90 weeks) than the other treatments (55–62 weeks), indicating a better toxicity profile and/or better effectiveness. The incidence of adverse effects was low for patients treated with NSAIDs only (48 per 100 p-yr of NSAID use). Higher incidences were found for patients treated with a SAARD in combination with NSAIDs and was highest during i.m. gold treatment (87 per 100 p-yr). Adverse effects caused discontinuation of the SAARD in 26% of i.m. gold-treated patients compared with 11% of MTX-treated patients and 8% of HCQ-treated patients. Meta-analysis has shown similar toxicity-related dropout rates: 30% for i.m. gold, 15% for MTX and 9% for antimalarial agents [8]. Although adverse effects occurred often during MTX in our study, most were mild and MTX could be continued. A difference between i.m. gold and all other therapies is that gold was given as an injection by the rheumatologist or research nurse rather than as pills, which might not be taken. Clearly, a dosage of 50 mg gold per week caused high toxicity rates in our cohort, which resulted in discontinuation in 26% of the cases. Permanent discontinuation might not always be necessary. A small study of 13 RA patients showed that very low dosages (< 20 mg per month) were effective for patients with gold-induced mucocutaneous reactions but a beneficial effect on the normal dosage [9].

Mucocutaneous adverse effects were most common during treatment with i.m. gold, adverse GI effects occurred predominantly during MTX and HCQ treatment, and pulmonary and hepatotoxic effects were most common during MTX therapy. These results are similar to those reported by others [10]. During our study there was one case of MTX-associated pneumonitis (i.e. one case out of 114 patients and 197 p-yr of MTX use). In earlier reports from several centres, the incidence of this type of toxicity ranged from 1 to 3% [11]. In an open trial of 453 MTX-treated RA patients (mean disease duration 13 yr) [12], 59% of patients experienced at least one adverse effect during a mean follow-up of 3 yr compared with 64% in our study. Weinblatt *et al.* [11] reported that 74% of MTX-treated RA patients showed adverse effects. Disease duration was similar to that in our cohort but the maximum MTX dose was higher (20 mg/week) than in our study (15 mg/week).

Although most patients in the three SAARD groups also used NSAIDs, the incidence of GI complaints did not differ much from that in the NSAID-only group. This result might be biased because the type and dosage of NSAIDs were not similar for the four groups. Both efficacy and toxicity differ among NSAIDs [13]. Patients in the NSAID-only group probably used higher doses and/or more toxic NSAIDs. Objective GI events occurred in 1% (five of 419), which is low in comparison with the 10–20% reported in the literature [14].

Endoscopy was performed only for clinical reasons. The occurrence of subjective GI complaints was also low in our study: 21% (87 of 419) compared with reported estimates of 30–40% of patients on long-term NSAID treatment [14]. These observations can be explained in part by population differences. At baseline, patients with active ulcer disease or corticosteroid therapy were excluded from our study. Reported risk factors for toxicity [13, 14], such as advanced age, prior ulcer disease, high-dose NSAIDs, use of multiple NSAIDs and corticosteroid use, were probably less common in our population of early RA patients than in other study populations. Moreover, 18% of patients used antacids for some period during follow-up. In the NSAID group, antacids were used by 25% of the patients compared with 18% of the HCQ and i.m. gold groups and 15% of the MTX group.

No clear relationship between sociodemographic or clinical variables at baseline and the occurrence of adverse effects was found. In other studies, some predictors were found: a low joint score, a lower polymorphonuclear cell count and the absence of RF were predictive of adverse effects in MTX-treated patients [12]. Renal impairment correlated with higher MTX toxicity rates [15], while comorbidity, marked disability, advanced age, long duration of disease and the start of a new drug *vs* continued use were weakly but statistically significantly correlated with toxicity in patients who had been treated with several SAARDs [16]. It is not clear whether predictors reported in the literature of established RA are applicable to a cohort of early RA patients, such as that in the present study.

The scoring method of Naranjo *et al.* [6] was used to specify causal links between symptoms and drugs. Sixty adverse effects were classified as being at least probably related to the use of a specific SAARD (eight to HCQ, 22 to i.m. gold, 30 to MTX) and 41 to NSAID use. In most cases the relationship with the specific medication remained unclear. Even for the 65 patients who stopped treatment because of adverse effects, the Naranjo score was below 8, indicating no definite association. In theory, the score in our study could exceed 8. In our opinion, the difficulty of relating adverse effects to a specific drug is not a consequence of the classification method used but is more likely to be attributable to the use of other medication, the disease itself or comorbidity.

The systematic assessment of toxicity has received little attention compared with the assessment of the effectiveness of treatments. A drug toxicity index, taking into account both frequency and severity, has been proposed by Fries *et al.* [17]. In a study using this index, HCQ was the least toxic, followed by i.m. gold; MTX was the most toxic [16]. One can also focus on the rates of severe adverse effects only or on the proportion of dropouts due to toxicity. Antimalarial drugs (such as HCQ) had the lowest dropout rates due to toxicity, with MTX close behind, while i.m. gold was the most toxic according to a meta-analysis [18].

Comparison of toxicity in non-randomized studies is not straightforward because both toxicity and the choice of treatment depend on the disease stage, disease activity and co-medication. We compared toxicity using data from a randomized controlled trial of patients with early RA, which minimizes the confounding effects of the disease itself. In general, our results are consistent with the literature. The relevance of this study is that four different treatment strategies were compared in a truly prospective way, and that the study included a large group of recently diagnosed RA patients.

Some limitations of our study will now be discussed. A possible bias might be introduced by the open character of our study. Both patients and rheumatologists knew which anti-rheumatic drug was used. Known adverse effects might be over-reported during clinical visits, while less well-known adverse effects might be under-reported (information bias). Adverse effects might be over-reported for i.m. gold and MTX, because these are thought to be more toxic than HCQ. This bias is unavoidable in an open study, but is probably less pronounced for objective and severe adverse effects. Therefore, a major bias in the number of patients who discontinued therapy is not likely to have been responsible for the high dropout rate found for i.m. gold.

Assessment of toxicity was a secondary aim of the clinical trial. The attending rheumatologist recorded information on a standard form that included an open question on whether any possible adverse effect had occurred since the last visit and, if so, of what kind. This recording of adverse effects was not fully structured, but it did not differ systematically between treatment groups. The reporting of adverse effects during clinical visits should cover the whole period since the previous visit; however, minor symptoms might be missed. Laboratory test results were recorded every 3 months during the first 2 yr of the study and every 6 months in the subsequent years. Temporary laboratory abnormalities occurring between recording dates could have been missed.

In conclusion, adverse effects were most common during i.m. gold therapy (87 per 100 p-yr) and led to discontinuation of this therapy within 2.5 yr in 26% of cases. The HCQ, MTX and NSAID-only treatments were less toxic. Among the adverse effects in this study that were probably related to SAARDs, none was irreversible or lethal. Although adverse effects were less common during NSAID-only treatment than during treatment with a SAARD, one might favour the latter because of its better effectiveness and because no irreversible or lethal adverse effects were seen.

Acknowledgement

This work was supported by a grant from The Dutch League against Rheumatism (Het Nationaal Reumafonds).

References

1. Psaty BM, Koepsell TD, Lin D *et al*. Assessment and control for confounding by indication in observational studies. J Am Geriatr Soc 1999;47:749–54.
2. Mitchell AA, Lesko SM. When a randomised controlled trial is needed to assess drug safety. The case of paediatric ibuprofen. Drug Safety 1995;13:15–24.
3. Arnett FC, Edworthy SM, Bloch DA *et al*. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.
4. van der Heide A, Jacobs JWG, Bijlsma JWJ, Heurkens AHM, van Booma-Frankfort C, van der Veen MJ. The effectiveness of early treatment with 'second-line' antirheumatic drugs. A randomized, controlled trial. Ann Intern Med 1996;124:699–707.
5. Akil M, Amos RS. Rheumatoid arthritis. II: treatment. Br Med J 1995;310:652–5.
6. Naranjo CA, Busto U, Sellers EM *et al*. A method for estimating the probability of adverse drug reactions. Clin Pharmacol Ther 1981;30:239–45.
7. Sharp JT, Young DY, Bluhm GB *et al*. How many joints in the hands and wrists should be included in a score of radiologic abnormalities used to assess rheumatoid arthritis? Arthritis Rheum 1985;28:1326–35.
8. Felson DT, Anderson JJ, Meenan RF. The comparative efficacy and toxicity of second-line drugs in rheumatoid arthritis—results of two meta-analyses. Arthritis Rheum 1990;33:1449–61.
9. Klinkhoff AV, Teufel A. How low can you go? Use of very low dosage of gold in patients with mucocutaneous reactions. J Rheumatol 1995;22:1657–9.
10. Furst DE. Rational disease-modifying anti-rheumatic drug (DMARD) therapy: is it possible? Br J Rheumatol 1996; 35:707–10.
11. Weinblatt ME, Kaplan H, Germain BF *et al*. Methotrexate in rheumatoid arthritis—a five-year prospective multicenter study. Arthritis Rheum 1994;37:1492–8.
12. Bologna C, Viu P, Picot M-C, Jorgensen C, Sany J. Long-term follow-up of 453 rheumatoid arthritis patients treated with methotrexate: an open, retrospective, observational study. Br J Rheumatol 1997;36:535–40.
13. Furst DE. Are there differences among nonsteroidal antiinflammatory drugs? Comparing acetylated salicylates, nonacetylated salicylates and nonacetylated nonsteroidal antiinflammatory drugs. Arthritis Rheum 1994;37:1–9.
14. Lichtenstein DR, Syngal S, Wolfe MM. Nonsteroidal antiinflammatory drugs and the gastrointestinal tract—the double-edged sword. Arthritis Rheum 1995;38:5–18.
15. Rheumatoid arthritis clinical trial archive group. The effect of age and renal function on the efficacy and toxicity of methotrexate in rheumatoid arthritis. J Rheumatol 1995;22:218–23.
16. Fries JF, Williams CA, Ramey D, Bloch DA. The relative toxicity of disease modifying antirheumatic drugs. Arthritis Rheum 1993;36:297–306.
17. Fries JF, Spitz PW, Williams CA, Bloch DA, Singh G, Hubert HB. A toxicity index for comparison of side effects among different drugs. Arthritis Rheum 1990;33:121–30.
18. Felson DT, Anderson JJ, Meenan RF. Use of short-term efficacy/toxicity tradeoffs to select second-line drugs in rheumatoid arthritis—A meta-analysis of published clinical trials. Arthritis Rheum 1992;35:1117–25.
19. Bijlsma JWJ, Oude Heuvel CHB, Zaalberg A. Development and validation of the Dutch questionnaire capacities of daily life (VDF) for patients with rheumatoid arthritis. Rehab Sci 1990;3:71–4.
20. Thompson PW, Silman AJ, Kirwan JR, Currey HLF. Articular indices for joint inflammation with rheumatoid arthritis. Correlation with the acute-phase response. Arthritis Rheum 1987;30:618–23.
21. van der Brink HR, van der Heide A, Jacobs JWG, van der Veen MJ, Bijlsma JWJ. Evaluation of the Thompson articular index. J Rheumatol 1993;20:28–32.

Appendix 1. Adverse drug reaction scale according to Naranjo *et al*. [6]

To assess the adverse drug reaction, the following questionnaire is completed by entering in the last column a score for each question and totalling the individual scores. The total score is then used to categorize the probability that the adverse reaction is attributable to the drug: ≥ 9, definitely; 5–8, probably; 1–4, possibly; ≤ 0, doubtfully.

		Yes	No	Do not know	Score
1	Are there previous conclusive reports on this reaction?[a]	+1	0	0	
2	Did the adverse event appear after the suspected drug was administered?	+2	−1	0	
3	Did the adverse reaction improve when the drug was discontinued or a specific antagonist was administered?	+1	0	0	
4	Did the adverse reaction reappear when the drug was readministered?	+2	−1	0	
5	Are there alternative causes (other than the drug) that could on their own have caused the reaction?	−1	+2	0	
6	Did the reaction reappear when a placebo was given?[b]	−1	+1	0	
7	Was the drug detected in the blood (or other fluids) in concentrations known to be toxic?[b]	+1	0	0	
8	Was the reaction more severe when the dose was increased or less severe when the dose was decreased?	+1	0	0	
9	Did the patient have a similar reaction to the same or similar drugs in any previous exposure?	+1	0	0	
10	Was the adverse event confirmed by any objective evidence?	+1	0	0	
Total score (−4 to 13)					

[a]Refers to any report listed in the National Pharmacotherapeutic Catalogue of 1997 as an adverse reaction to the specific SAARD or NSAID.
[b]Answers to questions 6 and 7 were unknown in the present study. The minimum and maximum possible scores were therefore −3 and 11 respectively (instead of −4 and 13).