CLINICAL TRIAL REPORT

A 4-year non-randomized comparative phase-IV study of early rheumatoid arthritis: integrative anthroposophic medicine for patients with preference against DMARDs versus conventional therapy including DMARDs for patients without preference

Harald J Hamre,[1] Van N Pham,[2] Christian Kern,[3] Rolf Rau,[4] Jörn Klasen,[3] Ute Schendel,[5] Lars Gerlach,[6] Attyla Drabik,[2] Ludger Simon[6,†]

[1]Institute for Applied Epistemology and Medical Methodology at the Witten/Herdecke University, Freiburg, Germany; [2]Institute of Statistics in Medicine, Universitätsklinikum Düsseldorf, Düsseldorf, Germany; [3]Department of Integrative Medicine, Asklepios Westklinikum, Hamburg, Germany; [4]Department of Rheumatology, Evangelisches Fachkrankenhaus Ratingen, Ratingen, Germany; [5]Department of Rheumatology, m&i-Fachklinik Bad Pyrmont, Bad Pyrmont, Germany; [6]Department of Internal Medicine and Gastroenterology, Filderklinik, Filderstadt, Germany

†Dr Ludger Simon passed away on June 10, 2016

Correspondence: Harald J Hamre
Institute for Applied Epistemology and Medical Methodology at the University of Witten-Herdecke, Zechenweg 6, 79111 Freiburg, Germany
Tel +49 761 1560 307
Fax +49 761 6125 6125
Email harald.hamre@ifaemm.de

Background: While disease-modifying antirheumatic drugs (DMARDs) are a mainstay of therapy for rheumatoid arthritis (RA), some patients with early RA refuse DMARDs. In anthroposophic medicine (AM), a treatment strategy for early RA without DMARDs has been developed. Preliminary data suggest that RA symptoms and inflammatory markers can be reduced under AM, without DMARDs.

Patients and methods: Two hundred and fifty-one self-selected patients aged 16–70 years, starting treatment for RA of <3 years duration, without prior DMARD therapy, participated in a prospective, non-randomized, comparative Phase IV study. C-patients were treated in clinics offering conventional therapy including DMARDs, while A-patients had chosen treatment in anthroposophic clinics, without DMARDs. Both groups received corticosteroids and nonsteroidal anti-inflammatory drugs (NSAIDs). Primary outcomes were intensity of RA symptoms measured by self-rating on visual analog scales, C-reactive protein, radiological progression, study withdrawals, serious adverse events (SAE), and adverse drug reactions in months 0–48.

Results: The groups were similar in most baseline characteristics, while A-patients had longer disease duration (mean 15.1 vs 10.8 months, $p<0.0001$), slightly more bone destruction, and a much higher proportion of women (94.6% vs 69.7%, $p<0.0001$). In months 0–12, corticosteroids were used by 45.7% and 81.6% ($p<0.0001$) and NSAIDs by 52.8% and 68.5% ($p=0.0191$) of A- and C-patients, respectively. During follow-up, both groups not only had marked reduction of RA symptoms and C-reactive protein, but also some radiological disease progression. Also, 6.2% of A-patients needed DMARDs. Apart from adverse drug reactions (50.4% and 69.7% of A- and C-patients, respectively, $p=0.0020$), none of the primary outcomes showed any significant between-group difference.

Conclusion: Study results suggest that for most patients preferring anthroposophic treatment, satisfactory results can be achieved without use of DMARDs and with less use of corticosteroids and NSAIDs than in conventional care.

Limitation: Because of the non-randomized study design, with A-patients choosing anthroposophic treatment, one cannot infer how this treatment would have worked for C-patients.

Keywords: clinical trial, Phase IV, disease-modifying antirheumatic drugs, integrative medicine, patient preference, rheumatoid arthritis

Plain language summary

For patients with early rheumatoid arthritis (ERA), a type of drugs called DMARDs are a mainstay of therapy. However, some ERA patients refuse DMARDs, for example, because of side effects. In AM, a treatment strategy for ERA has been developed without DMARDs. Instead, patients have additional physician counseling, use special medications such as herbs, and engage in special artistic or movement exercises. This study was done to see if ERA patients who refuse DMARDs can be treated with AM instead, with acceptable long-term results. ERA patients were treated with AM without DMARDs. They were compared to other ERA patients who were treated with DMARDs. Two other types of drugs for ERA, corticosteroids and so-called "nonsteroidal anti-inflammatory drugs", were allowed in both groups, but were less used among AM patients. After 4 years, rheumatic symptoms, inflammation measured in a blood test, and joint damage measured on X-ray were comparable in the two groups, while the AM patients had less side effects. Eight out of 129 AM patients needed DMARDs. Study results suggest that, for most ERA patients who are willing to engage in AM treatment, satisfactory long-term results can be achieved without DMARDs.

Background

Rheumatoid arthritis (RA) affects 0.5%–1% of the adult population[1] and is three times more frequent in women than men.[2] The symptoms of RA (joint pain, stiffness, tenderness, swelling, and effusion; malaise, fever, weight loss, and fatigue with decreased physical function[1,3]) and associated comorbid conditions (eg, extra-articular autoimmune disease, depression,[4] cardiovascular disease, malignancies and serious infections,[5] with average two comorbidities in each patient[1,6]) and the side effects of standard drug therapy with disease-modifying antirheumatic drugs (DMARDs), nonsteroidal anti-inflammatory drugs (NSAIDs), and possibly corticosteroids (eg, gastrointestinal bleeding, infections, osteoporosis, bone marrow depression, pulmonary fibrosis, liver toxicity[1,2,5]) lead to reduced quality of life, work disability, and considerable societal and personal costs as well as increased risks for hospitalization and mortality.[1]

For patients with early RA, immediate treatment with DMARDs is recommended, in order to retard or stop disease progression.[7] Nonetheless, a proportion of patients with early RA (range 3%–54%[8–14]) is not prescribed DMARDs or discontinues DMARD therapy within 12 months. Reasons for nontreatment with DMARDs, despite active RA, include adverse effects from DMARDs,[1,2,5] contraindications to DMARDs, economic barriers,[15] and patient preferences against DMARD therapy.[16–18] In a German study, DMARD refusal was strongly associated with preference for treatment with complementary and alternative medicine (CAM).[16] CAM use for RA is widespread (current or recent use reported by 34%–82% of patients from different settings around the world[19–26]); reasons for CAM use include dissatisfaction with DMARDs and other conventional medications because of adverse effects.[22,27]

Decision making on RA therapy should be shared between doctor and patient,[7,28] taking account of patient preferences. For patients with definite preference against DMARD therapy, the following questions arise: Can they be treated primarily without DMARDs, while minimizing the risk for rapid disease progression? Does CAM have a role here?

A treatment strategy for RA patients refusing DMARDs should include other standard RA treatment forms (such as NSAIDs, corticosteroids, physiotherapy) as needed, regular patient assessments, and, in case of uncontrollable disease progression, the option to reconsider DMARD therapy with the patient. Such a treatment strategy has been developed in anthroposophic medicine (AM), a physician-provided integrative whole medical system founded by Steiner and Wegman[29] (Box 1).

AM treatment for RA is usually individualized (tailored to individual needs, involving biographical counseling, special AM medications, special nonmedication artistic or movement AM therapies, and nutrition counseling). The aim is to retard disease progression and relieve symptoms, if possible without DMARDs. Corticosteroids are used if needed, with the long-term aim to reduce the dose or even stop the medication.[30–32]

Data from a small single-arm observational study of AM treatment for RA with 2-year follow-up indicate that symptoms and inflammatory markers can be reduced without DMARDs, but neither that study[33] nor studies of other CAM treatments for RA[34–38] included radiological disease progression as an outcome.

This controlled study aimed to evaluate the long-term safety and effectiveness of an AM treatment strategy without DMARDs for early RA – in patients choosing this strategy, compared to conventional treatment with DMARDs – in patients accepting DMARD treatment. Since an inherent feature of the target group (RA patients choosing AM treatment without DMARDs) is therapy preferences, a prospective, non-randomized study design was chosen,[39] allowing patients with preference against DMARDs to be treated without DMARDs. We present here the main study findings.

Box I Anthroposophic medicine

AM is practiced in 24 inpatient hospitals or hospital departments in Europe and in outpatient settings by an estimated 19,000 physicians around the world.[76] Based on the cognitive methods and cognitive results of anthroposophy, AM acknowledges not only physical and chemical forces active in nature, but also formative force systems, with specific forces involved in the formation of mineral substances, plants, animals, and humans. The interactions of these forces are understood to vary between different regions and organs in the human body, resulting in a complex equilibrium.[77] This equilibrium can be distorted in various forms of human disease[78] and is sought to be regulated by AM medications and nonmedication therapies.[79,80]

AM medications are prepared from plants, minerals, animals, and from chemically defined substances in concentrated or potentized form. Potentization implies a successive dilution, each dilution step involving a rhythmic succussion (repeated shaking of liquids) or trituration (grinding of solids into lactose monohydrate). For example, a D6 potency (also called 6X) has been potentized in a 1:10 dilution six times, resulting in a $1:10^{-6}$ dilution.[81] Potencies beyond D23 are unlikely to contain any molecules of the original substance; hence effects cannot readily be explained by molecular mechanisms. Nonetheless, a systematic review of in vitro studies found biologic effects of potencies $\geq$D23 in nearly three-fourths of the studies and in more than two-thirds of the studies with highest quality.[82] All AM medications are manufactured according to Good Manufacturing Practice and national drug regulations;[81] the available evidence suggests that they are generally well tolerated, with infrequent adverse reactions of mostly mild to moderate severity.[80,83]

AM nonmedication therapy includes eurythmy movement exercises,[84] art therapy,[85] and AM-extended physical therapy and nursing (including compresses, oil dispersion baths,[86] and rhythmical embrocation therapy[87]) and is provided by nurses, physiotherapists, and (eurythmy and art therapy) other nonmedical therapists.
Before initiating therapy, AM physicians have prolonged consultations with their patients, in order to take an extended history; address constitutional, psychosocial, and biographic-existential aspect of patients' illness; and select optimal therapy for each patient.[88]

Note: D = decimal, indicating a 1:10 dilution.
Abbreviation: AM, anthroposophic medicine.

Patients and methods
Study design, objectives, and setting
This prospective, non-randomized, comparative Phase IV study was conducted in four outpatient clinics in Germany. The study objectives were to describe the long-term outcomes of early RA and to compare the outcomes in RA patients choosing treatment in outpatient clinics providing either AM treatment without DMARDs (two A-clinics: Filderklinik, Filderstadt; Asklepios Westklinikum, Hamburg) or conventional treatment with DMARDs (two C-clinics: Krankenhaus St Josef, Wuppertal; Johannes Wesling Klinikum Minden, Minden).

All four clinics were part of general hospitals, offering outpatient and inpatient services to study patients, either in internal medicine departments (A-clinics) or in rheumatology departments (C-clinics). A-clinics treated patients referred from various German Federal States, while C-clinics generally treated patients referred from their vicinity.

Participants
Study participants were consecutive, self-selected patients starting treatment for early RA in the anthroposophic (A-patients) or conventional (C-patients) clinics and fulfilling all eligibility criteria. Participants were recruited by the physicians providing outpatient treatment for them in the respective clinics (study physicians).

Inclusion criteria were age 16–70 years and RA according to the 1987 American College of Rheumatology criteria[40] of <3 years duration since onset of initial symptoms (painful swelling of RA-typical joints or tenosynovitis of the hands).

Exclusion criteria were other rheumatic disorders, conditions that could mimic RA (eg, acute infections, osteomalacia, neurologic diseases), severe concomitant disorders (eg, malignant disease, severe hepatic disease, renal insufficiency), mental disorders, alcohol or drug abuse, pregnancy, previous treatment with DMARDs or surgical synovectomy, and treatment with high-dose intravenous corticosteroids, intra-articular (i.a.) corticosteroid injections, or radiological or medical synovectomy in the past 6 weeks.

Scheduled treatment
Treatment of all patients was individualized according to the needs of the patient and, for medications, the manufacturers' recommendations. The available therapy components as well as their selection, dosage, and duration in the patients reflected routine clinical practice for early RA in the participating clinics at the time of the study.

Symptomatic treatment in all patients
All patients in both groups were offered symptomatic treatment with NSAIDs as needed and, if indicated, corticosteroids. Notably, the indication for corticosteroids was stricter in the A-group (high systemic inflammatory activity or visceral organ affection) than in the C-group (moderate systemic inflammatory activity). Corticosteroids were given

at an initial dose according to the inflammatory activity (0.3–1.5 mg/kg prednisolone equivalents daily) and slowly tapered, aiming for a daily dose below 7.5–10 mg prednisolone equivalents. In cases of acute exacerbations of RA or complications from RA, patients were given inpatient treatment in the respective clinic.

Specific treatment in A-patients

All A-patients were offered individualized AM treatment with the aim to relieve pain and other RA symptoms and to retard disease progression. AM treatment comprised the following:

- AM medications, selected from a predefined medication list and applied orally, externally or as subcutaneous (but not i.a.) injections;[30,31]
- at least one AM nonmedication therapy (details in Box 1), typically in weekly sessions of 30–45 min duration;
- nutrition counseling;
- biographical counseling by the AM physician.[41]

Standard nutrition advice was a lacto-vegetarian diet without sugar from sugar beet and without alcohol. Optional for overweight patients was a 7–10-day fast followed by an oligoantigenic diet with gradual food reintroduction according to individual tolerance (adapted from Kjeldsen-Kragh[42]). Further lifestyle advice included warm clothing, preferably from sheep wool.[31]

In A-patients, i.a. corticosteroid injections were to be avoided as far as possible. DMARDs as well as chemical or radiological synoviorthesis and surgical synovectomy were not permitted and, if they occurred, it led to withdrawal from the study.

Specific treatment in C-patients

All C-patients were offered individualized conventional treatment according to a predefined therapy guideline, with the aim to retard disease progression. Treatment included one or several DMARDs (chloroquine, hydroxychloroquine, salazosulfapyridine, methotrexate, aurothiomalate, aurothioglucose, auranofin, *Escherichia coli* extract). Biologic DMARDs were only used in selected patients, because at the time of the study, they were not yet established as routine treatment of RA in Germany. Local invasive joint treatment with corticosteroid injections or synovectomy (chemical, radiation, or surgical) was given as needed.

Primary outcome measures
Primary clinical outcome measures

Primary clinical outcome measures were as follows:

- Intensity of RA symptoms and signs, measured by patient self-report on four visual analog scales (VAS) (0= not at

all, 100= extremely severe): joint inflammation in past 3 months, joint swelling today, joint pain today, and joint stiffness today, yielding a total score (Rheumatoid Arthritis Visual Analog Scale [RA-VAS]) with a range from 0 to 400. The four VAS were adapted from items 1–4 of the Rapid Assessment of Disease Activity in Rheumatology questionnaire.[43]

- C-reactive protein (CRP) in serum (normal ≤0.5 mg/dL).
- Radiological disease progression in hands and feet, measured by the Ratingen Score.[44] The Ratingen Score is a modified version of Larsen's scoring method[45] to assess radiological changes in RA: articular surface destruction is measured in 2×19 joints or sites of the hands and feet (five proximal interphalangeal joints, five metacarpophalangeal joints, naviculum, lunatum, radius, ulna, metatarsophalangeal joints 2–5, interphalangeal joint of the great toe). For each joint or site, the amount of destruction is rated on a numerical scale from 0 to 5 points (0= none, 1= destruction of ≤20% of the articular surface, 2=21%–40%, 3=41%–60%, 4=61%–80%, 5=≥81%), yielding a total score range of 0–190. Minimal detectable intraindividual change, as rated by the same radiologist, is 6.6 points.[46–51]

In the predefined primary analyses of RA-VAS and CRP, the area under the curve (AUC) for score values in months 0–3–6–9 and 39–42–45–48, respectively, was compared. AUC was calculated according to the trapezoidal rule, as defined by the formula:

$$AUC_{0-tz} = \frac{\sum_{i=0}^{z}(t_i - t_{i-1})(C_{i-1} + C_i)}{2}$$

where t = time point of follow-up (months 0, 3, 6, and so on), 0= first time point, tz = last time point, i = follow-up number (0 at baseline, 1 after 3 months, 2 after 6 months, and so on), and C = score value (RA-VAS) or concentration (CRP).

The interpretation of AUC values can be illustrated by two examples. A patient with a (low) CRP concentration of 0.5 mg/dL at each of the four measurements in months 0–3–6–9 will have an AUC of 4.5 mg/dL and a patient with a (high) CRP concentration of 10 mg/dL at these four time points will have an AUC of 90 mg/dL.

In the predefined primary analysis of Ratingen Score, score values after 0 and 48 months, respectively, were compared.

Primary safety outcome measures

Primary safety outcome measures were the incidence of withdrawals from the study, serious adverse events (SAE), and adverse drug reactions (ADRs) in months 0–48.

Secondary outcome measures

Secondary clinical outcome measures

Secondary clinical outcome measures presented here are as follows:

- use of corticosteroids and NSAIDs in months 0–48;
- Steinbrocker Radiological Staging System (0–4);[52]
- Steinbrocker Functional Classification (physician rating, I–IV);[52]
- Disease Activity Score[53] (a compound score based on erythrocyte sedimentation rate [ESR], Ritchie-Index of joint tenderness,[54] current health status [VAS 0–100], and number of swollen joints [0–40]). For the latter item, a 40-joint swelling count (DAS40[55]) was used in this study, while the customary DAS44[53] includes four additional joints (two sternoclavicular + two acromioclavicular). Regarding the other three items, DAS40 and DAS44 are identical. Furthermore, the incidence of swelling of the four missing joints in DAS40 among patients with early RA is reportedly very low (2%–3% for sternoclavicular joints,[56,57] 0% for acromioclavicular joints[57]). Therefore, the two DAS versions can be expected to yield very similar score values.

Further secondary clinical outcomes will be described elsewhere.

Secondary safety outcome measures

Secondary safety outcome measures were cause and intensity of SAE and ADR (mild–moderate–severe), ADR associated with stopping of study medication, and ratings of therapy tolerability by patient (poor–average–good) and physician (VAS: 0= no adverse reactions, 100= most severe adverse reactions).

Data collection, diagnostic procedures, financial aspects

All data were documented with questionnaires collected by a study monitor every 6–12 weeks. At baseline, the study physicians documented eligibility criteria and clinical findings; in addition, routine laboratory investigations and pulmonary function tests were performed. RA was diagnosed by the study physician. In addition, for all study patients, the RA diagnosis was independently confirmed after 0 and 48 months by one experienced rheumatologist who was otherwise not involved in the study (Prof H Mathies†, Department of Rheumatology, Asklepios Klinikum Bad Abbach, Bad Abbach, Germany). CRP was measured by turbidimetry (Tina-quant® CRP; Boehringer Ingelheim, Mannheim, Germany) in all patients.

Radiographs of the hands and feet were obtained after 0, 24, 36, and 48 months according to a standardized protocol.

- Ratingen Score was independently documented by two different radiologists who were experienced in the use of this score (author RR, and Prof PL van Riel, Department of Rheumatic Diseases, Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands) and who were blinded toward all patient data as well as the sequence of the radiographs.
- Steinbrocker Radiological Staging was documented by the study physicians, based on inspection of the radiographs as well as the radiographic assessments report from the radiologists in the respective clinic.

Pulmonary function tests were performed every 12 months. Daily use of corticosteroids and NSAIDs was documented by the patients in diaries collected every 3 months. All other data were documented by the study physicians at scheduled visits every 3 months, unless otherwise stated. Since the patients were assigned to A- or C-group according to the clinic they were treated in (see the "Study design, objectives, and setting" section), patients, therapy providers, and (except for Ratingen Score, see above) outcome assessors were not blinded toward the group assignment of patients.

For study purposes, one 75% position as assistant physician was established in each clinic. Otherwise, all out- and inpatient treatment was funded by the ordinary funding schemes in the German health system. The physicians received no extra remuneration beyond their ordinary salary for their participation; likewise, the patients received no financial compensation.

Statistical methods

Data preparation

Data were analyzed using SAS Version 8.2 (SAS Inc., Cary, NC, USA). Missing data for corticosteroid use in patient diaries occurred in n=2 A-patients and n=13 C-patients and were replaced by the corresponding corticosteroid prescription data in the study physicians' documentation. Otherwise, missing data were not replaced and each analysis comprised the respective patients with available data. For each patient, the daily dose of corticosteroids was converted to prednisolone equivalents,[58] while the daily dose of NSAIDs was calculated as a percentage of the maximum daily dose recommended by the manufacturer and subsequently converted to diclofenac equivalents. Ratingen Score had been calculated twice, based on the assessments of radiological findings by two different radiologists; inter-rater correlation was high (r=0.83) with the two different assessments yielding very similar results. In this paper, results of the assessment by

the radiologist with the most experience with the Ratingen Score method (author RR) are presented, unless otherwise specified. Adverse events and ADRs were coded according to Medical Dictionary for Regulatory Activities.

Primary analyses

The data analysis comprised three main assessments with altogether nine primary outcome analyses as follows: (I) pre–post changes of the three primary clinical outcomes (RA-VAS, CRP, and Ratingen Score) in the A- and C-groups, respectively; (II) comparison of pre–post changes of the three primary clinical outcomes between A- and C-groups; and (III) comparison of the three primary safety parameters (study withdrawals, SAE, ADRs) between A- and C-groups.

For the nine primary outcome analyses, alpha was defined as 0.05 for assessments I and II and 0.15 for assessment III. For each of the assessments I + II and III, respectively, *p*-values were adjusted for multiple testing according to the Bonferroni method, resulting in $p<(0.05/3=)0.0167$ for assessments I and II and $p<(0.15/3=)0.05$ for assessment III. For all nine primary analyses, 99.5% CI were calculated. For all other data analyses, statistical significance was defined as $p<0.05$ and, if applicable, 95% CI were calculated.

For bivariate analyses of continuous and rank-ordered variables, two-tailed *t*-tests for paired samples and Wilcoxon test were used for within-group comparisons, and *t*-test for independent samples and Mann–Whitney *U* test were used for between-groups comparisons. For bivariate between-groups analyses of binominal variables, chi-square test and Fisher's exact test were used. All tests were two-tailed. In this paper, we generally present results of nonparametric tests, as they do not rely on assumptions of normal distribution, which were not met in a number of cases. Also, parametric and nonparametric tests showed similar results and the inclusion of parametric tests would not have changed the study interpretation.

Each treatment group (A and C) consisted of patients treated in two different clinics (see the "Study design, objectives, and setting" section). However, treatment protocols were identical in each treatment group. Furthermore, the two clinic subgroups in the A- and C-groups, respectively, had very similar baseline data. Therefore, multilevel analyses to adjust for potential center effects were not deemed necessary.

Sample size estimation

Sample size estimation was conducted separately for clinical outcomes and safety outcomes. Sample size for clinical outcomes was based on previous studies of the Sharp/van der Heijde score for radiological progression[59] and CRP[60] in RA. With a power of 0.90 and adjusted alpha of 0.0167 (see earlier), the highest sample size was estimated for the between-group analysis of Sharp/van der Heijde Score change, indicating a sample size of (2×52=)104 evaluable patients and, accommodating for an estimated dropout rate of 25%, at least (2×70=)140 enrolled patients. Sample size for the three primary safety outcomes was estimated using power 0.80, an estimated outcome rate difference of 20% for each outcome, alpha 0.048, and no dropout (because safety was assessed throughout the study), yielding a sample size of (2×104=)208 enrolled patients. From these estimations, and in order to enable one long-term follow-up analysis as well as earlier follow-up analyses with larger sample sizes within the time frame of the study, it was planned to recruit first a minimum of 140 (2×70) patients to be followed up for 48 months and subsequently another 140 (2×70) patients to be followed up for 24 or 36 months. For RA-VAS, CRP, and Ratingen Score, clinically relevant between-group differences were predefined as ≥20% of the respective mean score value in the C-group.

Further preplanned analyses

The possible influence of demographics, baseline status, and corticosteroid use on the primary clinical outcomes was assessed by bivariate (Kruskal–Wallis test, chi-square test, Pearson's correlation) and multivariate (linear and logistic) regression analyses as well as inspection of scatter plots as follows:

- Age, gender, disease duration, CRP, ESR, treatment group: For the three primary clinical outcomes, we investigated associations between change from baseline to 48 months and age, gender, disease duration, baseline ESR, treatment group, and (Ratingen Score only) baseline CRP. We used generalized linear modeling with estimation of scale parameters by the maximum likelihood method, testing for model effects and between-group differences with Wald chi-square statistics, and estimation of marginal means with Wald type 95% CI.
- Corticosteroid use: For CRP and RA-VAS, we investigated associations between the changes in score values from baseline to each follow-up assessment with the use of corticosteroids (yes/no) in the preceding 3 months. For Ratingen Score, we investigated associations between radiological progression (yes/no) from baseline to 24 and 48 months, respectively, and corticosteroid use in the preceding 12 months. These analyses were restricted to female patients and were performed separately for each of the two independent radiological assessments.

One interim primary clinical outcome analysis was performed after 5½ years.

Relevant amendments to the study protocol

Patient recruitment was slower than expected, which led to two relevant amendments to the study protocol. First, according to the protocol, patient recruitment in A- and C-groups was to be stratified according to Disease Activity Score[53] at baseline, in order to achieve comparable groups regarding baseline morbidity in this non-randomized study; instead, all eligible patients in both groups were recruited. Second, the patient recruitment period was prolonged from original 1½ to 4½ years.

A third relevant protocol amendment concerned the assessment of radiological disease progression, that is, instead of the Sharp/van der Heijde method,[61,62] the Ratingen Score was used because the latter had been found to have better discriminatory properties.[49,50] This amendment was made in March 2001, toward the end of the recruitment period; the predefined sample size estimation (see above) was not amended.

Quality assurance, adherence to regulations, and reporting guidelines

The study was conducted according to the Helsinki Declaration, the International Conference on Harmonisation Good Clinical Practice guidelines, and legal requirements. This paper followed the TREND guidelines for reporting of non-randomized studies.[63] This trial was retrospectively registered with the German Clinical Trials Register on 23 September 2016: TRN DRKS00011083; UTN U1111-1187-5015.

Ethics approval and consent to participate

This study was approved by the Ethics Committee of the Bavarian State Chamber (Ethik-Kommission Nr 95360, date 30 January 1996). Written informed consent was obtained from all patients before study enrollment.

Results

Patient enrollment and availability of follow-up data

Patient enrollment and follow-up

From 1 January 1996 to 30 June 2001, a total of 261 patients were assessed for eligibility; n=251 patients (n=129 A-patients and n=122 C-patients) fulfilled all eligibility criteria and were enrolled in the study. First visit of the first patient was on 20 March 1996, and last visit of the last patient

was on 19 December 2003. Of the 251 enrolled patients, n=221 patients were scheduled for the final follow-up examination after 48 months, while n=30 patients, enrolled toward the end of the recruitment period, had their last follow-up scheduled after 24 months (n=22) or 36 months (n=8). Pooled follow-up rates were 88.4% (n=222/251), 70.1% (n=176/251), 64.2% (n=147/229), and 62.0% (n=137/221) after 12, 24, 36, and 48 months, respectively. On comparing these rates in A- and C-patients at each year (Table S1), no significant differences were found.

Study withdrawals

After 48 months, 46 A-patients and 38 C-patients had been withdrawn from the study. Reasons for study withdrawal (Figure 1) differed significantly between A- and C-patients (p=0.0009, Fisher–Freeman–Halton test). Among patients scheduled for 48-month follow-up, study withdrawal because of nonadherence (noncompliant or lost to follow-up) was lower in A-patients (12.7%, n=15/118) than in C-patients (24.8%, n=25/103; p=0.0120).

Comparing patients with (n=137) and without (n=114) available 48-month follow-up data regarding demographics and baseline morbidity, no significant differences were found, except for age (mean 51.0 vs 48.0 years in patients with and without follow-up data, respectively, p=0.043; Table S2).

Demographics and baseline characteristics

Main data on demographics and baseline characteristics are summarized in Tables 1 and 2. Comparing A- and C-patients, no significant differences were found regarding CRP, ESR, rheumatoid factor, physician and patient ratings of joint swelling, and patient self-ratings of current health and of health satisfaction (Table 1).

Compared to C-patients, A-patients had significantly a significantly higher proportion of women (A-patients 94.6% vs C-patients 69.7%; odds ratio [OR] for female gender in A- vs C-patients 7.59, 95% CI: 3.23–17.82, p<0.0001) and they were younger (mean 47.58 vs 50.65 years, p=0.0205). Furthermore, A-patients had longer disease duration (mean 15.10 vs 10.84 months, median 14.00 vs 7.00 months, p<0.0001) and slightly more joint tenderness as rated by the physician (mean Ritchie Index[54] 13.87 vs 12.56, p=0.0268), but less self-rated joint inflammation (VAS 0–100: mean ± SD 57.83±26.55 vs 77.73±19.78, p<0.0001), joint pain (46.24±25.90 vs 55.10±30.85, p=0.0116), and joint stiffness (42.36±24.98 vs 52.56±29.84, p=0.0067).

Among the routine laboratory investigations, α2-globulin was lower in A-patients than in C-patients (9.21±2.17
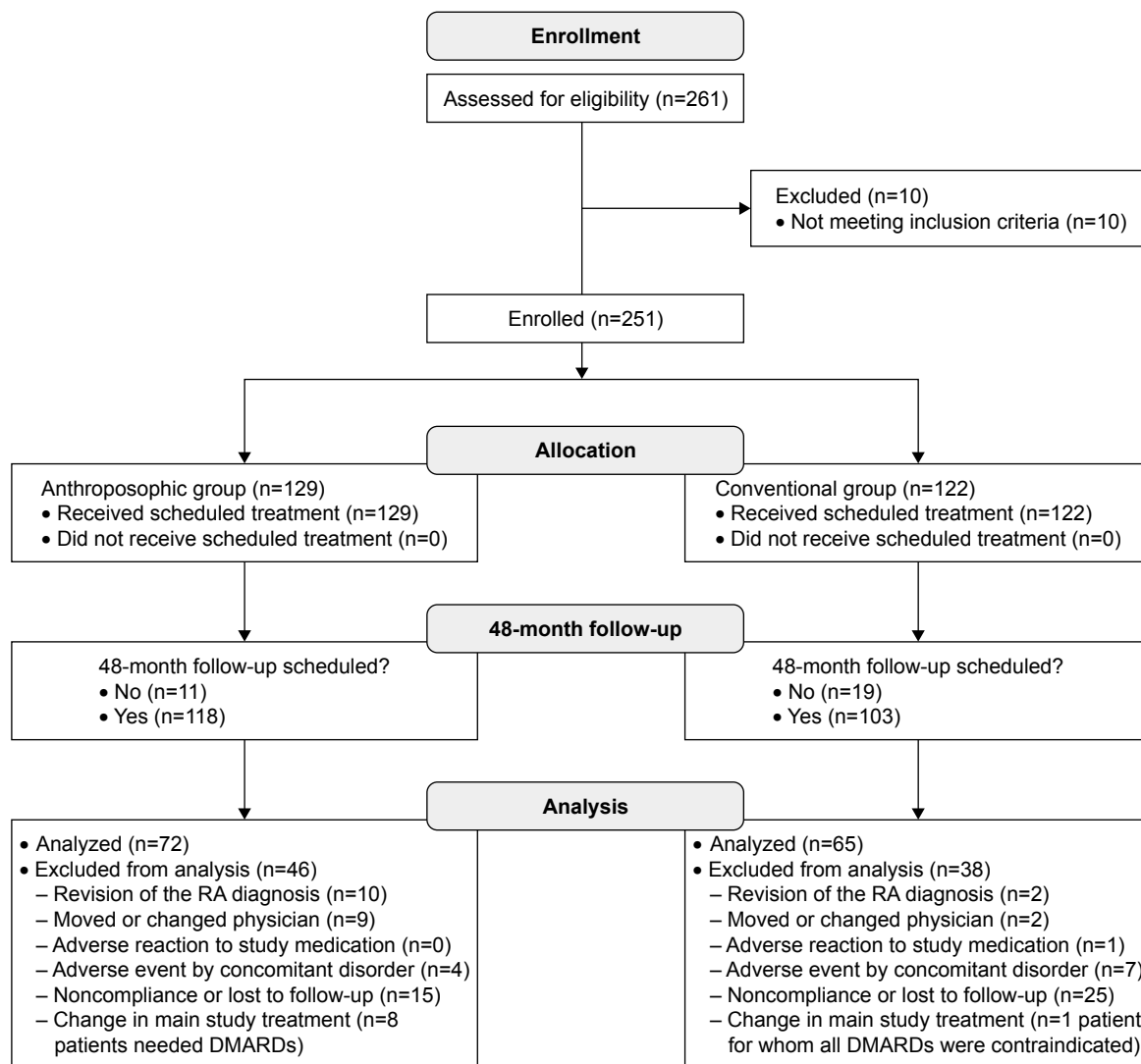
Figure 1 Patient enrollment, follow-up, and analysis.
Abbreviations: DMARD, disease-modifying antirheumatic drug; RA, rheumatoid arthritis.

vs 9.86±2.20, p=0.0207). Otherwise, apart from two differences due to the higher proportion of female in A-patients (hemoglobin 12.66±1.13 vs 13.31±1.28, p<0.0001; ferritin mean 64.11±76.38, median 40.50 vs mean 153.43±149.75, median 102.00, p<0.0001), no significant differences were found. Overall disease activity (DAS, grouped as low–moderate–high) did not differ significantly between the groups (Table 2).

Regarding radiological joint destruction, Ratingen Score[44] values did not differ significantly between the groups (Table 1), while A-patients had more advanced radiological changes according to the Steinbrocker Staging System[52] (p=0.0293; Table 2), with a trend toward erosions occurring more often according to Ratingen Score (score values ≥1 in 54.5% and 42.3% of A- and C-patients, respectively, OR

2.00, 95% CI: 0.93–2.74, p=0.0737), as shown in Table 1. Regarding RA-specific functional limitation, Hanover Functional Ability Questionnaire[64] did not differ significantly between the groups (Table 1), while A-patients had less functional limitation according to the Steinbrocker Functional Classification[52] (p=0.0288), as shown in Table 2. Corticosteroid use at baseline was much less frequent among A-patients than C-patients (25.6% vs 48.4%; OR for corticosteroid use in A- vs C-patients 0.37, 95% CI: 0.21–0.65, p=0.0003). Ratings of tolerability of ongoing therapy at baseline by patients (Table 2; Figure S1) and physicians (Figure S2) did not differ significantly between the groups.

To sum up, A- and C-groups were similar in most aspects, albeit with some relevant differences: The A-group had a higher proportion of women, had more advanced disease

Table 1 Demographics and baseline characteristics

Item	Anthroposophic group	Conventional group	p-value
Female gender, n (%)	122/129 (94.6%)	85/122 (69.7%)	<0.0001
Age (years)	47.58±11.25	50.65±12.15	0.0205
Duration of rheumatoid arthritis (months)	15.10±9.21	10.84±9.17	<0.0001
<12, n (%)	54/125 (43.2%)	78/122 (63.9%)	0.0014
12–23, n (%)	45/125 (36.0%)	27/122 (22.1%)	
24–35, n (%)	26/125 (20.8%)	17/122 (13.9%)	
Rheumatoid factor (IU/mL)	137.37±222.44	194.82±266.22	0.4634
Rheumatoid factor positive (≥20 IU/mL), n (%)	87 (74.4%)	80 (65.6%)	0.1593
C-reactive protein (mg/dL)	2.59±3.49	2.66±4.64	0.2584
Erythrocyte sedimentation rate (mm/hour)	32.72±26.73	35.44±25.42	0.2221
Rheumatoid Arthritis VAS (patient ratings, 0–400)	193.87±87.16	238.30±93.36	<0.0001
Joint inflammation in past 3 months (VAS, 0–100)	57.83±26.55	77.73±19.78	<0.0001
Joint swelling today (VAS, 0–100)	46.58±25.33	52.91±29.72	0.0916
Joint pain today (VAS, 0–100)	46.24±25.90	55.10±30.85	0.0116
Joint stiffness today (VAS, 0–100)	42.36±24.98	52.56±29.84	0.0067
Joint tenderness (Ritchie Index, 0–78)	13.87±7.60	12.56±9.97	0.0268
Number of swollen joints	15.22±8.37	14.00±7.11	0.4242
Hanover Functional Ability Questionnaire	18.02±4.26	16.70±5.10	0.0683
Radiological signs of joint destruction (Ratingen Score)	2.40±3.90	2.17±4.66	0.1059
Patient self-rating of current health status (0–100)	54.72±23.40	56.20±26.15	0.6408
Patient satisfaction with current health status (0–100)	59.23±24.74	54.43±29.18	0.1647
Erosive arthritis with Ratingen Score ≥1, n (%)	67/123 (54.5%)	51/119 (42.3%)	0.0737
Corticosteroid use, n (%)	33/129 (25.6%)	59/122 (48.4%)	0.0003

Note: Data are presented as mean ± SD, unless otherwise stated.
Abbreviation: VAS, visual analog scales.

(duration, radiological changes), and used corticosteroids less frequently than C-patients.

Compared to other treatment studies of DMARD-naïve patients with early RA[65–67] (RA duration <2 vs <3 years in this study), the proportion of women was again much higher in this study, while baseline markers of symptoms, inflammation, and bone destruction were of the same orders of magnitude.

Treatment used

In general, treatment was administered in the respective clinic of the patient. For A-patients with long travel distance from their A-clinic, nonmedication AM therapies were given at local therapy centers (or as part of inpatient treatment in the A-clinic).

Corticosteroids in both groups

The use of corticosteroids in both groups is presented in Table 3. The proportion of patients using corticosteroids was significantly lower among A-patients than C-patients in all four study years, with OR for corticosteroid use (A- vs C-groups) ranging from 0.20 in the first year to 0.55 in the third year (Table 3; Figure S3).

In the subgroup of corticosteroid users (n=148), the daily dose in the first year was significantly higher among A-corticosteroid users (mean 13.56±18.09 prednisolone equivalents, median 8.78 equivalents, n=58) than C-corticosteroid users (mean 7.00±2.85, median 6.16, n=90; p=0.0008, Mann–Whitney U-test). Comparing the first and fourth years, the corticosteroid daily dose among users (n=63 evaluable patients) was significantly reduced in both groups (Wilcoxon test p<0.0001 for each within-group comparison). In the fourth year, the daily dose of corticosteroids did not differ significantly between A-users (mean 4.74±2.84 prednisolone equivalents, n=30) and C-users (mean 5.28±2.32, n=43; p=0.1733, Mann–Whitney U-test).

NSAID use in both groups

The use of NSAIDs in both groups is presented in Table 3. The proportion of patients using NSAIDs was lower among A-patients in all four study years, with OR for NSAID use (A- vs C-groups) ranging from 0.35 in the fourth year to 0.55 in the third year; differences were significant in years 1, 2, and 4 (Table 3; Figure S4).

In the subgroup of NSAID users, the daily dose was lower among A-users than C-users in each of all four study

Table 2 Secondary categorical outcomes at months 0 and 48

Outcome	Month 0				Month 48			
	A-group		C-group		A-group		C-group	
	n	Percent	n	Percent	n	Percent	n	Percent
Steinbrocker staging system								
0. No radiological abnormalities	14	11.2	40	33.1	15	21.1	16	29.6
1. Periarticular osteoporosis	41	32.8	23	19.0	13	18.3	7	13.0
2. As 1, with initial cartilage or bone erosion	65	52.0	50	41.3	34	47.9	22	40.7
3. Marked cartilage and bone erosion, joint deformation	5	4.0	8	6.6	9	12.7	8	14.8
4. As 3, with joint ankylosis	0	0.0	0	0.0	0	0.0	1	1.9
Total	125	100.0	121	100.0	71	100.0	54	100.0
		$p=0.0289$				$p=0.7779$		
Steinbrocker functional classification								
I. Complete ability to carry out all the usual duties without handicaps	7	5.4	2	1.6	37	51.4	26	40.0
II. Adequate for normal activities despite handicap of discomfort or limited motion of one of the joints	102	79.1	85	69.7	32	44.4	37	56.9
III. Limited to little or none of the duties of usual occupation or self-care	19	14.7	32	26.2	3	4.2	2	3.1
IV. Incapacitated, largely or wholly bed-ridden or confined to a wheelchair with little or no self-care	1	0.8	3	2.5	0	0.0	0	0.0
Total	129	100.0	122	100.0	72	100.0	65	100.0
		$p=0.0043$				$p=0.2536$		
Disease Activity Score								
1. Low (≤2.4)	3	2.3	11	10.1	52	73.2	43	68.3
2. Moderate (2.4 to ≤3.7)	51	39.5	42	38.5	18	25.4	17	27.0
3. High (>3.7)	75	58.1	56	51.4	1	1.4	3	4.8
Total	129	100.0	109	100.0	71	100.0	63	100.0
		$p=0.1431$				$p=0.4848$		
Patient rating of therapy tolerability								
1. Poor	4	4.0	8	9.3	0	0.0	1	1.8
2. Average	27	27.3	18	20.9	2	2.9	14	24.6
3. Good	68	68.7	60	69.8	66	97.1	42	73.7
Total	99	100.0	86	100.0	68	100.0	57	100.0
		$p=0.9616$				$p=0.0001$		

Note: p-values: between-group differences, Mann–Whitney *U* test.
Abbreviations: A-group, Anthroposophic group; C-group, Conventional group.

Table 3 Use of corticosteroids and NSAIDs in years 1–4

Year	Patients using medication				Odds ratio for use (A-group vs C-group)			Daily dose* (mg) in medication users	
	A-group		C-group		OR	95% CI	p-value	A-group	C-group
	n	Percent	n	Percent				Mean ± SD	Mean ± SD
Corticosteroids (oral or intravenous)									
1	58/127	45.7	90/111	81.1	0.20	0.11–0.35	<0.0001	13.56±18.09	7.00±2.85
2	59/118	50.0	70/101	69.3	0.44	0.25–0.77	0.0040	7.46±5.94	5.36±2.40
3	45/99	45.5	55/91	60.4	0.55	0.31–0.97	0.0428	6.18±7.32	5.18±2.77
4	30/84	35.7	43/68	63.2	0.32	0.17–0.63	0.0010	4.74±2.84	5.28±2.32
NSAIDs									
1	67/127	52.8	76/111	68.5	0.51	0.29–0.90	0.0191	75.23±37.98	95.96±31.70
2	33/118	28.0	48/101	47.5	0.43	0.24–0.78	0.0043	57.18±27.59	86.64±40.27
3	28/99	28.3	38/91	41.8	0.55	0.29–1.05	0.0723	60.14±26.99	85.55±30.17
4	20/84	23.8	32/68	47.1	0.35	0.17–0.74	0.0046	46.54±26.44	81.43±39.28

Note: *Daily dose of prednisolone equivalents (corticosteroids) and diclofenac equivalents (NSAID), respectively.
Abbreviations: A-group, Anthroposophic group; C-group, Conventional group; NSAID, nonsteroidal anti-inflammatory drug.

years; differences were tested in the first and fourth years and were significant in both years (p=0.0002 and p=0.0009, respectively, Mann–Whitney U-test).

Comparing the first and fourth years, the NSAID daily dose among users was reduced in both groups; the difference was significant in A-patients (mean reduction of 24.79 mg diclofenac equivalents, 95% CI: 5.48–44.10, p=0.0147, n=20) but not in C-patients (mean reduction 10.25, 95% CI: −9.53 to 30.02, p=0.2961, n=26).

Invasive joint therapy in both groups

The i.a. corticosteroid injections were given to 11.6% (n=15/129) of A-patients and 45.9% (n=56/122) of C-patients (OR for i.a. corticosteroid injection in A- vs C-patients 0.16, 95% CI: 0.08–0.31, p<0.0001).

In the C-group, 26.2% (n=32/122) of patients had chemical synovectomy, 0.8% (n=1) had surgical synovectomy, and 4.9% (n=6) had radiation synovectomy. In the A-group, these treatments did not occur, except for one patient who had radiation synovectomy outside the study setting. (This treatment should, according to protocol, have led to study withdrawal, but that was not detected until after all analyses had been completed. The patient was retained in the analyses.)

Specific treatment in A-patients

Individualized AM medication therapy was administered to all 129 A-patients; most frequently used AM medication types were mistletoe extracts (*Viscum album*, prescribed to ~80% of A-patients), potentized insect extracts (*Apis mellifica*, *Formica rufa*, *Vespa crabro*, also prescribed to ~80%), and winter blooming herbs in potentized forms (*Colchicum autumnale*, *Helleborus niger*, *Hedera helix* to ~40%).

Approximately 80% of A-patients received physical therapy (eg, exercises, heat or cold applications, massage, embrocation, compresses, balneotherapy), about half of A-patients had eurythmy therapy, and about one third had AM art therapy (painting, clay modeling, or speech exercises). In all A-patients, the study physician obtained a detailed biographical anamnesis followed by biographical counseling, usually in conjunction with the follow-up visits. All A-patients had nutrition counseling; for approximately one-third of them, the physician prescribed an individualized nutrition plan.

Approximately half of A-patients had inpatient rehabilitation treatment, usually in the first follow-up year and of several weeks' duration. The inpatient treatment enabled a more extensive use of the above-mentioned AM treatment modalities.

The analysis of predefined treatment modalities not permitted in A-patients (see the "Patients and methods" section

for details) showed that 6.2% (n=8/129) of A-patients needed DMARDs during follow-up because of disease progression that could not be controlled without DMARDs. These eight patients were withdrawn from the study according to the protocol (after 6, 12, 18, 21, 33, 33, 42, and 48 months, respectively). In an extreme scenario sensitivity analysis, the maximum possible number of DMARD users in the A-group was calculated, assuming that all A-patients withdrawn from the study (except withdrawals due to revision of the RA diagnosis) used DMARDs. In this model, the maximum possible number of DMARD users was 7.1% (n=9/127) of A-patients within 12 months, 22.6% (n=28/124) within 24 months, 29.6% (n=34/115) within 36 months, and 33.3% (n=36/108) within 48 months.

Specific treatment in C-patients

In study months 1–3, conventional synthetic DMARDs were used by all 122 C-patients (thereof methotrexate in n=86 patients, sulfasalazine in n=14, parenteral gold salts in n=6, chloroquine in n=2, azathioprine in n=1, and cyclosporine in n=1). Seven patients had an initial combination therapy with a second DMARD (methotrexate in n=3; cyclophosphamide, sulfasalazine, gold salts, and infliximab in n=1 patient, respectively).

In study months 46–48, DMARDs were still being used by n=60 of 65 evaluable C-patients (thereof methotrexate in n=47, leflunamide in n=5, sulfasalazine in n=4, gold salts in n=2, azathioprine in n=1, and cyclosporine in n=1). Of these 60 C-patients, n=10 had a combination therapy with another DMARD (chloroquine in n=4, sulfasalazine in n=2, leflunamide in n=2, methotrexate in n=1, and etanercept in n=1).

Almost all C-patients had physical therapy (eg, exercises, heat and cold applications, massage, balneotherapy).

Primary clinical outcome measures
Bivariate statistical analyses

The results of bivariate statistical analyses have been presented in Table 4. Between months 0–9 and 39–48, RA-VAS and CRP were significantly reduced (indicating improvement) in both groups. Between months 0 and 48, the Ratingen Score increased significantly (indicating radiological disease progression) in both groups. Comparing the amount of change between A- and C-groups for each of the three primary clinical outcomes, no significant differences were found (Table 4).

For RA-VAS and CRP, additional exploratory analyses were performed, comparing mean score values at months 0 and 48 (instead of mean of AUC values in months 0–9 and 39–48, see the "Patients and methods" section for details). Compared to the AUC value analyses, larger samples were

Table 4 Primary clinical outcome analyses

Outcome	Anthroposophic group			Conventional group		
	n	Mean ± SD	Median	n	Mean ± SD	Median
RA-VAS (0–400)						
Months 0–9	117	1,104.7±542.0	1,071.0	94	1,233.6±674.5	1,128.8
Months 39–48	70	773.4±655.7	606.8	50	861.8±661.5	884.3
Change: months 0–9 to 39–48	69	−308.1±637.7	−333.0	45	−381.3±566.9	−315.0
		$p<0.001$			$p<0.001$	
Change in A- vs C-groups*			$p=0.9261$			
C-reactive protein (mg/dL)						
Months 0–9	117	15.07±16.13	7.80	89	9.83±8.14	7.47
Months 39–48	70	8.94±11.03	4.95	54	6.84±5.85	4.66
Change: months 0–9 to 39–48	68	−5.14±13.76	−1.16	45	−3.81±9.20	−2.55
		$p=0.0014$			$p=0.0024$	
Change in A- vs C-groups*			$p=0.8626$			
Ratingen Score (0–190)						
Rating by Prof R Rau						
Month 0	123	2.40±3.90	1.00	118	2.17±4.66	0.00
Month 48	71	5.98±7.68	3.00	61	4.66±7.22	1.00
Change: months 0–48	71	4.16±6.10	1.00	61	2.04±3.98	1.00
		$p<0.001$			$p<0.001$	
Change in A- vs C-groups*			$p=0.1205$			
Rating by Prof PL van Riel						
Month 0	123	2.93±4.40	1.00	118	2.63±5.01	1.00
Month 48	71	7.16±8.91	4.00	61	5.44±8.85	3.00
Change: months 0–48	71	4.28±6.08	2.00	61	2.69±4.87	1.08
		$p<0.001$			$p<0.001$	
Change in A- vs C-groups*			$p=0.1788$			

Notes: RA-VAS and C-reactive protein in months 0–9 and 39–48: area under the curve for score values at months 0, 3, 6, 9 and 39, 42, 45, 48, respectively. *Mann–Whitney U test. All other p-values refer to Wilcoxon test.
Abbreviations: A- vs C-groups, Anthroposophic group versus Conventional group; RA-VAS, Rheumatoid Arthritis Visual Analog Scale.

available for the additional analyses, while the results were very similar (Table S3).

Clinically relevant differences

For the three primary clinical outcomes, a clinically relevant between-group difference had been predefined as ≥20% of the mean value of the respective score in the C-group as follows:

- For RA-VAS, a clinically relevant difference, with lower scores in A-patients, was observed after 18 and 36 months (Figure 2).
- For CRP, a clinically relevant difference, with higher scores in A-patients, was observed at all follow-ups between 3 and 33 months, but not between 36 and 48 months (Figure 3).

- Ratingen Score was higher in A-patients at 36- and 48-month follow-ups, with a difference of 30% and 28%, respectively (Figure 4). According to the above-mentioned criterion, these between-group differences were clinically relevant. On the other hand, the absolute differences (mean 1.25 and 1.32 points after 36 and 48 months, respectively) as well as the 0–48-month differences (mean 4.16 and 2.04 points in A- and C-patients, respectively) were lower than the minimal detectable intraindividual change of 6.6 points.

Associations between primary clinical outcomes and other variables

Age, gender, disease duration, CRP, ESR, treatment group

The 0–48-month changes of CRP and RA-VAS, respectively, showed significant negative associations with baseline ESR.
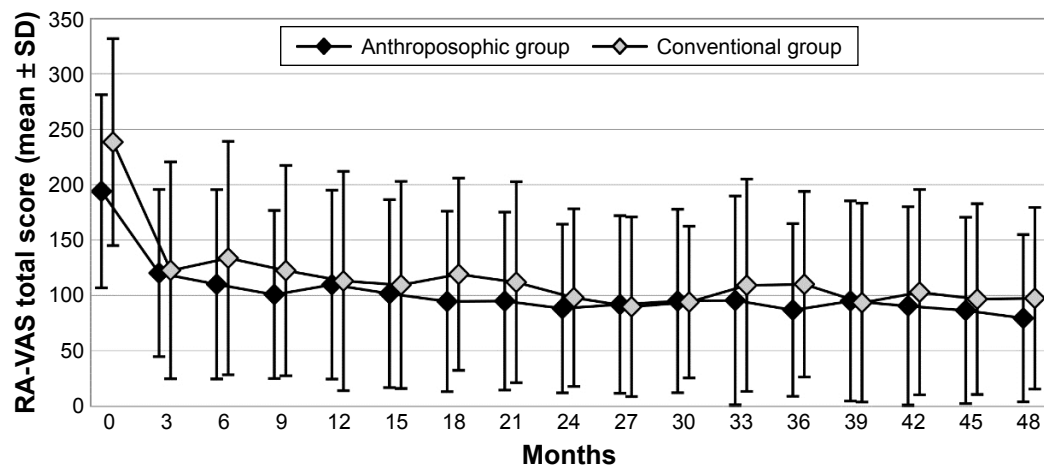
Figure 2 RA-VAS total score.
Note: Total score (0–400) in months 0 (Anthroposophic group: n=127, Conventional group: n=119), 3 (127+110), 6 (123+109), 9 (119+99), 12 (113+89), 15 (107+89), 18 (101+81), 21 (94+84), 24 (94+79), 27 (85+63), 30 (85+59), 33 (83+59), 36 (79+58), 39 (76+61), 42 (74+60), 45 (73+57), and 48 (71+60).
Abbreviation: RA-VAS, Rheumatoid Arthritis Visual Analog Scale.

In the two multivariate models, an increase in baseline ESR of 1 mm/hour was associated with the following:

- 0.06 mg/dL less reduction in CRP (Wald 95% CI: 0.03–0.09 mg/dL, $p=0.0003$);
- 1.27-point less reduction in RA-VAS (Wald 95% CI: 0.69–1.85, $p<0.0001$).

All other investigated associations were not significant (for Ratingen Score, the 0–48-month change showed a trend toward more radiological deterioration in A-patients than in C-patients [$p=0.0212$], which did not fulfill the predefined significance criterion of 0.0167 for this comparison, cf the "Patients and methods" section for details).

Corticosteroid use

C-reactive protein: In A-patients, a significant negative association was found between corticosteroid use and CRP change in months 0–3, indicating that A-patients using corticosteroids had more reduction of CRP from baseline than A-patients not using corticosteroids. This association was also found in months 4–6, albeit very weak. For all subsequent 3-month follow-up periods, no such significant associations were found. In C-patients, no significant associations were found between corticosteroid use and CRP change from baseline at any follow-up.

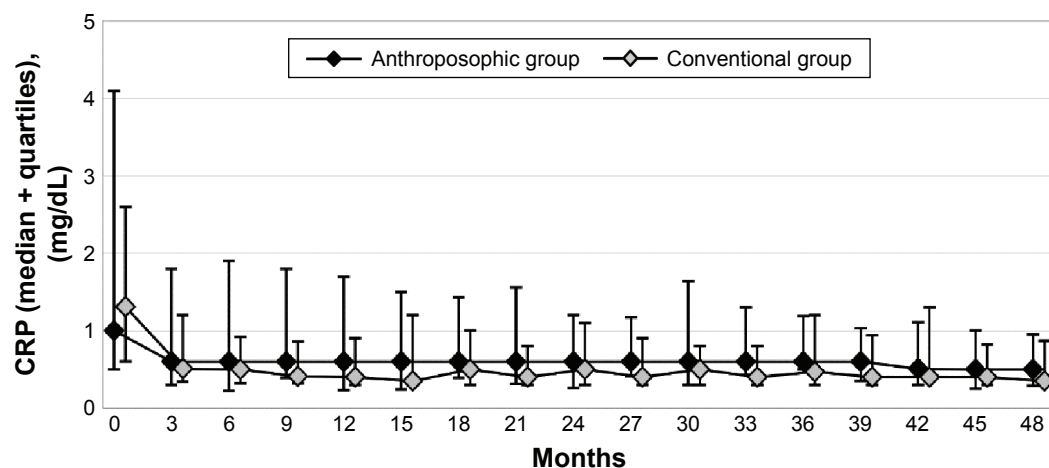Rheumatoid Arthritis Visual Analog Scale: In A-patients, a significant negative association was found between

Figure 3 C-reactive protein.
Note: CRP in months 0 (Anthroposophic group: n=129, Conventional group: n=109), 3 (126+107), 6 (123+106), 9 (118+103), 12 (118+99), 15 (116+96), 18 (112+94), 21 (105+97), 24 (102+91), 27 (93+70), 30 (90+70), 33 (85+70), 36 (86+66), 39 (76+60), 42 (73+60), 45 (73+60), and 48 (72+64).
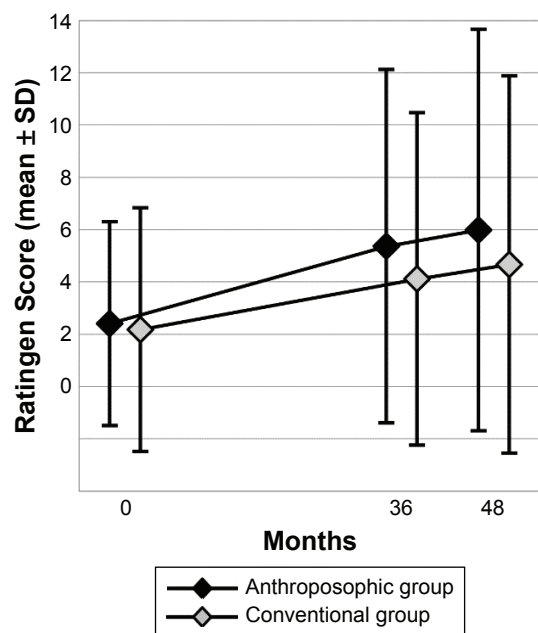**Abbreviation:** CRP, C-reactive protein.

Figure 4 Ratingen Score.
Note: Ratingen Score of radiological joint destruction (0–190) in months 0 (Anthroposophic group: n=123, Conventional group: n=118), 36 (88+63), and 48 (71+61).

corticosteroid use and RA-VAS change in months 0–3, indicating that A-patients using corticosteroids had more improvement in RA-VAS total score, compared to A-patients not using corticosteroids. Corresponding significant associations were also found for the RA-VAS subscales joint pain and joint stiffness. In months 9–12, significant associations were found between corticosteroid use and RA-VAS in A-patients and, to a small degree, in C-patients as well. For all other 3-month follow-up periods, no such significant associations were found either in A- or in C-patients.

Ratingen Score: In A-patients, using the assessment by RR, which was used in all other analyses, a significant but very weak negative association was found between corticosteroid use and radiological progression after 24 months. Using the assessment by the other radiologist, no such associations were found. In C-patients, no significant associations were found between corticosteroid use and radiological progression for any follow-up period, using any of the two radiological assessments.

All these analyses of associations between corticosteroid use and CRP, RA-VAS, and Ratingen Score, respectively, were repeated using "low-dose corticosteroid use" (≤7.5 mg prednisolone equivalents per day) instead of "any corticosteroid use" as the criterion, yielding either similar, albeit less-pronounced, associations (CRP) or no significant association (RA-VAS, Ratingen Score).

Primary safety outcome measures

Among the three primary safety outcome measures, one (ADRs to study medication) showed significantly lower incidence in A-patients (50.4% of patients) than in C-patients (69.7%; OR for ADR in A- vs C-patients 0.44, 99.5% CI: 0.20–0.95, $p=0.0020$), while two measures did not differ significantly between the groups (OR for study withdrawal in A- vs C-patients 1.23, 99.5% CI: 0.56–2.70; OR for SAE 0.97, 99.5% CI: 0.52–1.80; Table 5).

Secondary outcome measures

At 48-month follow-up, no significant differences were found regarding Steinbrocker radiological staging system[52] or Steinbrocker Functional Classification[52] (Table 2). Between months 0 and 48, DAS improved significantly in both groups ($p<0.0001$ for both comparisons; Figure S5); the average improvement did not differ significantly between the groups ($p=0.6261$). At 48-month follow-up, approximately two-thirds of patients in both groups had low disease activity (DAS ≤2.4, Table 2).

At 48-month follow-up, patient rating of therapy tolerability differed significantly between A- and C-patients, indicating better tolerability in A-patients ($p=0.0001$), as shown in Table 2 and Figure S1.

Between months 0 and 48, physician rating of therapy tolerability (VAS: 0= no adverse reactions, 100= most severe adverse reactions) improved significantly in A-patients ($p<0.0001$) and deteriorated significantly in C-patients ($p<0.0001$), as shown in Figure S2. At 48-month follow-up, physician rating of therapy tolerability differed significantly between A- and C-patients, indicating less or less severe ADR in A-patients (mean 0.34±1.56, median 0.00) than in C-patients (mean 6.19±15.28, median 0.00; $p=0.0002$, Mann–Whitney U-test).

The incidence of ADRs associated with stopping of study medication was lower in A-patients (13.2%) than in C-patients (29.5%; OR 0.36, 95% CI: 0.18–0.72, $p=0.0019$). Most frequent causes of ADRs in the A-group were AM medications (n=51 ADRs) and corticosteroids (n=36 ADRs) and in the C-group were DMARDs (n=119 ADRs) and corticosteroids (n=26 ADRs). Intensity of SAE was higher in C-patients than in A-patients ($p=0.0499$).

The remaining secondary safety outcomes did not differ significantly between the two groups (Table 5).

Discussion
Main findings

This was a 4-year prospective, non-randomized, Phase IV study of patients with early RA without prior use of DMARDs,

Table 5 Safety outcomes in months 0 and 48

Primary safety outcomes	Anthroposophic group		Conventional group		A-group vs C-group
	Number of patients	Percent	Number of patients	Percent	p-value
Patients withdrawn from the study	46/129	35.7	38/122	31.1	0.5040
Patients with SAE	31/129	24.0	30/122	24.6	1.0000
Patients with ADR	65/129	50.4	85/122	69.7	0.0020
Secondary safety outcomes					
Patients with ADR, study medication stopped	17/129	13.2	36/122	29.5	0.0019
Patients with ADR of severe intensity	6/129	4.7	13/122	10.7	0.0946
	Number of events	Percent	Number of events	Percent	p-value
Causes of ADR by medication classes					0.0005*
Nonsteroidal anti-inflammatory drugs	18/105	17.1	25/170	14.7	0.6110
Corticosteroids	36/105	34.3	26/170	15.3	0.0003
AM medications (in A-group) vs DMARD (in C-group)	51/105	48.6	119/170	70.0	0.0005
Total	105/105	100.0	170/170	100.0	
Intensity of SAE					0.0567**
Mild	18/41	43.9	12/42	28.6	
Moderate	21/41	51.2	22/42	52.4	
Severe	2/41	4.9	8/42	19.0	
Total	41/41	100.0	42/42	100.0	
Causes of SAE					0.1034*
Concomitant disorder with probable or possible relation to study medication	0/41	0.0	4/42	9.5	
Joint surgery or complications of rheumatoid arthritis	5/41	12.2	7/42	16.7	
Concomitant disorder unrelated to study medication or rheumatoid arthritis	36/41	87.8	31/42	73.8	
Total	41/41	100.0	42/42	100.0	
SAE resolved or improved at last follow-up	37/41	90.2	32/42	76.2	0.1415

Notes: p-values refer to Fisher's exact test, unless otherwise stated. *Fisher–Freeman–Halton test. **Mann–Whitney U test.
Abbreviations: A-group, Anthroposophic group; ADR, adverse drug reaction; AM, anthroposophic medicine; C-group, Conventional group; DMARDs, disease-modifying antirheumatic drugs; SAE, serious adverse events.

starting treatment in four outpatient clinics in Germany. The study comprised two patient groups (A- and C-patients); C-patients, without specific therapy preferences, were treated in two rheumatology departments and received conventional drug therapy with DMARDs, while A-patients, with preference against DMARDs, were treated in internal medicine departments and received AM treatment with special medications, special nonmedication therapies, and additional physician counseling, without DMARDs in almost all (94%) cases. Both groups had physical therapies and used corticosteroids and NSAIDs as needed, with significantly less A-patients using corticosteroids and NSAIDs.

The groups were similar regarding most baseline characteristics, with a longer disease duration, slightly more bone destruction, and a higher proportion of women in the A-group. During follow-up, both groups had substantial reduction of RA symptoms and systemic inflammation and, as expected, a comparatively slight progression of radiological bone destruction; these outcomes did not differ significantly between the groups. The incidence of ADR was significantly lower among A-patients (50% of patients) than C-patients (70%), while the incidence of SAE and study withdrawals did not differ significantly between the groups. Nonadherence to study protocol (noncompliant or lost to follow-up after 48 months) was lower in A-patients (13%) than in C-patients (25%).

Strength and limitations

To our knowledge, this is the first controlled study of a CAM or integrative intervention for RA with a follow-up beyond 12 months, the first with radiological progression as the outcome, with a focus on early RA, and with a CAM intervention implemented without the use of DMARDs. Apart from this study, we know of only four treatment studies of DMARD-naïve patients with early RA in which radiological progression was measured after 48 months or later.[68–71] Other strengths include a rigorous adherence to RA diagnostic criteria with independent assessor technique[72] and blinded rating of radiological findings by two different assessors.

Notably, the study was not designed to test the benefits and risks of different treatment regimens in the same type of patients. The comparisons were between patient groups differing insofar as A-patients had deliberately chosen treatment in AM clinics, without DMARDs, while C-patients were treated in clinics offering conventional therapy including DMARDs. Accordingly, one can neither infer how AM treatment for early RA would have worked for the C-patients nor how conventional therapy would have worked for A-patients.

Apart from this fundamental, intended contrast, there is the issue if other baseline characteristics could have affected the study results. Notably, prognostically relevant baseline variables[73] were either comparable between the two groups (rheumatoid factor, CRP, number of swollen joints, Ratingen Score) or differed, indicating a poorer prognosis in A-patients (gender, disease duration, Steinbrocker staging system).

Among the six primary outcomes, five did not differ significantly between the groups, while one, the incidence of ADR, was significantly lower in A-patients. In multivariate analyses of the three primary clinical outcomes, significant but weak negative associations were found between baseline ESR and the 0–48-month improvement of CRP and RA-VAS, respectively. Other baseline variables (age, gender, disease duration, treatment group) showed no significant associations with the primary clinical outcomes. Although residual confounding cannot be excluded, it seems unlikely that other baseline characteristics could have affected results in a way that would change the study interpretation.

Data were collected from 1996 to 2003, before newer biologic DMARDs became more widely used in Germany. However, these drugs are still not recommended as first-line therapy for early RA, but only for the minority of patients in whom traditional DMARDs are ineffective or not tolerated.[74] The indication for conventional synthetic DMARDs in mono- or combination therapy in this study is in line with current guidelines.[7,28,75] Also, in most other aspects, the drug therapy for C-patients did not differ substantially from the current standard care for early RA in Germany. Furthermore, the allocation of responsibility for treatment of C-patients to rheumatologists (instead of general internists or general practitioners) and the monitoring of all patients with reassessment of the indication for DMARDs at least every 3 months are in line with current guidelines.[7,28,75] Nonetheless, it cannot be excluded that for a subset of C-patients, switching to biologic DMARDs could have resulted in better clinical outcomes (to a small extent, such switching occurred in this study; eg, at 48-month follow-up, 1 in 60 evaluable C-patients used a biologic DMARD). This issue will be further assessed elsewhere.

The AM treatment used for A-patients in the study corresponds to current recommendations.[76] A limitation in this respect is the lack of exact quantitative data on which AM treatment modalities were used in each A-patient.

Follow-up rates in this study decreased from 88% after 12 months to 62% after 48 months. The latter rate is lower than in other treatment studies of early RA after 48 months (74%–83%[70,77,78]). However, study patients with and without 48-month follow-up had comparable baseline characteristics (except patients without follow-up data being 3 years younger), and follow-up rates did not differ between A- and C-patients at any year. Therefore, dropout bias affecting the between-group differences does not seem likely, although this cannot be excluded.

During follow-up, 8 (6%) of the 129 A-patients were withdrawn from the study because they needed DMARDs. In addition, 29 A-patients were withdrawn because they had moved or changed physician or were lost to follow-up. It cannot be excluded that some of these A-patients also had an unfavorable course of disease and needed DMARDs between their last available follow-up examination and the scheduled end of their study participation; thus, the frequency of DMARD use among A-patients (6% throughout the study) might be underestimated. In an extreme scenario sensitivity analysis, the maximum possible use of DMARDs amounted to only 7% of A-patients within 12 months, increasing to one-third of A-patients within 48 months.

The analyses and discussion of this paper focus on the predefined primary outcomes and statistical analyses, according to the protocol written in 1996. Since then, other RA outcomes and other statistical approaches to this type of studies have become more widespread. Secondary outcomes (eg, response and remission according to the European League against Rheumatism) and supplementary or alternative statistical analyses (eg, more comprehensive outcome predictor analyses including multilevel models, equivalence designs) will be published elsewhere.

Interpretation and implications for practice

While DMARDs are a mainstay of conventional RA therapy, some patients with early RA refuse DMARDs or do not tolerate them well. Furthermore, for some patients, the high drug costs, especially for the newer biologic DMARDs, can represent a barrier to their use.[15]

In AM, a treatment strategy for early RA without DMARDs has been developed, which entails additional physician counseling, special artistic and movement therapies, and special medications. The results of this study suggest that, for

a large majority of patients with early RA who are willing to engage in AM treatment, satisfactory long-term results can be achieved without DMARDs and with less use of corticosteroids and NSAIDs and less ADR than in conventional care.

Admittedly, radiological scores showed more destruction in A-patients after 2 and 4 years, but these differences were small (≤20% of the minimal detectable difference in individual patients) and not statistically significant, although the predefined sample size for that comparison was attained.

While symptom reduction was similar in both groups throughout the study, the reduction of CRP in the first 2¾ years was more outspoken in C-patients than A-patients with a clinically relevant difference, while CRP levels were similar in both groups at 3-year follow-up and beyond. This finding may be related to the more frequent use of corticosteroids as initial treatment in the C-group. Compared to C-patients, A-patients had significantly better safety outcomes (less ADR, less ADR associated with withdrawal of study medication, better therapy tolerability in physician as well as patient ratings) and a trend toward less ADR of severe intensity. These differences are likely a consequence of the lower use of DMARDs, corticosteroids, and NSAIDs in A-patients. Notably, among corticosteroid users in the first year (46% and 81% of A- and C-patients, respectively), the average daily dose was almost twice as high in A-patients. This may be related to the higher threshold for corticosteroid prescription in the A-group, compared to the C-group (see the "Patients and methods" section for details), with consequent need for higher corticosteroid doses to control inflammation and symptoms among A-patients using this medication.

As in conventional therapy, the AM approach requires frequent monitoring of patients (at least every 3 months) with reassessment of therapy response and, if needed, consideration of the option to start treatment with corticosteroids and/or DMARDs. More specifically, the slower initial reduction of CRP in A-patients raises the question if, when initiating AM treatment for early RA, corticosteroids should be offered more liberally to patients with high inflammatory activity.

An overarching principle in current guidelines for RA treatment is shared decision making between the patient and the rheumatologist (or other physician).[7,28,74,75] For physicians offering AM treatment for early RA, the results of this study can be used to inform patients about the possibilities and limitations of the AM approach.

Conclusion

In this non-randomized study of AM or conventional treatment of patients with early RA without prior DMARD use, both treatment regimens were associated with marked reduction in symptoms and CRP and only slight radiological progression after 4 years; these results did not differ significantly between the groups. The incidence of ADRs was significantly lower among A-patients (50% of patients) than C-patients (70%). The patient groups differed according to therapy preference, insofar as A-patients had deliberately chosen treatment in AM clinics, primarily without DMARDs, while C-patients were treated in clinics offering conventional therapy with DMARDs. Accordingly, one cannot infer how AM treatment for early RA would have worked for the C-patients.

Author contributions

LS, JK, and US wrote the study protocol. LS and CK were the study physicians. RR calculated the Ratingen Score. VNP and AD analyzed data, and HJH performed a final audit of the analyses. HJH and LS wrote the first publication draft. LS was the principal investigator, had access to all data, was guarantor until May 2016 (followed by HJH), and approved the penultimate manuscript version before his death in June 2016. All authors contributed toward data analysis, drafting and revising the paper and agree to be accountable for all aspects of the work.

Disclosure

In the past 3 years, the following authors received funding from manufacturers of medications used for RA therapy: HJH

References

1. Gabriel SE, Michaud K. Epidemiological studies in incidence, prevalence, mortality, and comorbidity of the rheumatic diseases. *Arthritis Res Ther*. 2009;11(3):229.
2. Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. *Lancet*. 2010; 376(9746):1094–1108.
3. Lee DM, Weinblatt ME. Rheumatoid arthritis. *Lancet*. 2001;358(9285): 903–911.
4. Matcham F, Rayner L, Steer S, Hotopf M. The prevalence of depression in rheumatoid arthritis: a systematic review and meta-analysis. *Rheumatology (Oxford)*. 2013;52(12):2136–2148.
5. Listing J, Gerhold K, Zink A. The risk of infections associated with rheumatoid arthritis, with its comorbidity and treatment. *Rheumatology (Oxford)*. 2013;52(1):53–61.
6. Michaud K, Wolfe F. Comorbidities in rheumatoid arthritis. *Best Pract Res Clin Rheumatol*. 2007;21(5):885–906.
7. Smolen JS, Landewe R, Bijlsma J, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis*. 2017;76(6):960–977.
8. Bonafede MM, Fox KM, Johnson BH, Watson C, Gandra SR. Factors associated with the initiation of disease-modifying antirheumatic drugs in newly diagnosed rheumatoid arthritis: a retrospective claims database study. *Clin Ther*. 2012;34(2):457–467.
9. Kiely P, Williams R, Walsh D, Young A. Contemporary patterns of care and disease activity outcome in early rheumatoid arthritis: the ERAN cohort. *Rheumatology (Oxford)*. 2009;48(1):57–60.
10. Jamal S, Alibhai SM, Badley EM, Bombardier C. Time to treatment for new patients with rheumatoid arthritis in a major metropolitan city. *J Rheumatol*. 2011;38(7):1282–1288.
11. de Thurah A, Norgaard M, Johansen M, Stengaard-Pedersen K. Time to methotrexate treatment in patients with rheumatoid arthritis referred to hospital. *Scand J Rheumatol*. 2010;39(1):19–25.
12. Rantalaiho V, Kautiainen H, Virta L, Korpela M, Mottonen T, Puolakka K. Trends in treatment strategies and the usage of different disease-modifying anti-rheumatic drugs in early rheumatoid arthritis in Finland. Results from a nationwide register in 2000–2007. *Scand J Rheumatol*. 2011;40(1):16–21.
13. Widdifield J, Bernatsky S, Paterson JM, et al. Quality care in seniors with new-onset rheumatoid arthritis: a Canadian perspective. *Arthritis Care Res*. 2011;63(1):53–57.
14. Navarro-Compan V, Smolen JS, Huizinga TW, et al. Quality indicators in rheumatoid arthritis: results from the METEOR database. *Rheumatology (Oxford)*. 2015;54(9):1630–1639.
15. Hopson S, Saverno K, Liu LZ, et al. Impact of out-of-pocket costs on prescription fills among new initiators of biologic therapies for rheumatoid arthritis. *J Manag Care Spec Pharm*. 2016;22(2):122–130.
16. Westhoff G, Zink A. Basic treatment of early rheumatoid arthritis. Abstaining from rheumatological care and preferring alternative medicine increase the risk of undertreatment. *Z Rheumatol*. 2007;66(2): 121–124, 126–128.
17. Carli C, Ehlin AG, Klareskog L, Lindblad S, Montgomery SM. Trends in disease modifying antirheumatic drug prescription in early rheumatoid arthritis are influenced more by hospital setting than patient or disease characteristics. *Ann Rheum Dis*. 2006;65(8):1102–1105.
18. Brown EM, Garneau KL, Tsao H, Solomon DH. DMARD non-use in low-income, elderly rheumatoid arthritis patients: results of 86 structured interviews. *Arthritis Res Ther*. 2014;16(1):R30.
19. Buchbinder R, Gingold M, Hall S, Cohen M. Non-prescription complementary treatments used by rheumatoid arthritis patients attending a community-based rheumatology practice. *Intern Med J*. 2002; 32(5–6):208–214.
20. Rao JK, Kroenke K, Mihaliak KA, Grambow SC, Weinberger M. Rheumatology patients' use of complementary therapies: results from a one-year longitudinal study. *Arthritis Rheum*. 2003;49(5):619–625.
21. Herman CJ, Allen P, Hunt WC, Prasad A, Brady TJ. Use of complementary therapies among primary care clinic patients with arthritis. *Prev Chronic Dis*. 2004;1(4):A12.
22. Kajiyama H, Akama H, Yamanaka H, et al. One third of Japanese patients with rheumatoid arthritis use complementary and alternative medicine. *Mod Rheumatol*. 2006;16(6):355–359.
23. Zaman T, Agarwal S, Handa R. Complementary and alternative medicine use in rheumatoid arthritis: an audit of patients visiting a tertiary care centre. *Natl Med J India*. 2007;20(5):236–239.
24. Lee MS, Lee MS, Yang CY, et al. Use of complementary and alternative medicine by rheumatoid arthritis patients in Korea. *Clin Rheumatol*. 2008;27(1):29–33.
25. Tokem Y, Parlar KS, Ozer S, Nakas D, Argon G. A multicenter analysis of the use of complementary and alternative medicine in Turkish patients with rheumatoid arthritis: holistic nursing practice review copy. *Holist Nurs Pract*. 2014;28(2):98–105.
26. Mehrpoor G, Owlia MB, Mirjalili MR, Farzan A, Akrami A. Complementary alternative medicine in rheumatologic diseases; data from outpatient clinics in Yazd, Iran. *J Case Rep Practi*. 2014;2(3):92–96.
27. Rose G. Why do patients with rheumatoid arthritis use complementary therapies? *Musculoskeletal Care*. 2006;4(2):101–115.
28. Smolen JS, Landewe R, Breedveld FC, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis*. 2014;73(3):492–509.
29. Steiner R, Wegman I. *Extending Practical Medicine. Fundamental Principles Based on the Science of the Spirit. [GA 27] [First published 1925]*. Bristol: Rudolf Steiner Press; 2000.
30. Simon L. Arzneimittelfindung am Beispiel chronische Polyarthritis (chronisches Gelenkrheuma) [Remedy finding: the example of rheumatoid arthritis]. In: Glöckler M, editor. *Anthroposophische Arzneitherapie für Ärzte und Apotheker [Anthroposophic Medication Therapy for Physicians and Pharmacists]*. Stuttgart: Deutscher Apotheker Verlag; 2005:3-1–3-28.
31. Simon L. Anthroposophische Langzeittherapie entzündlich-rheumatischer Erkrankungen – Konzeption und Perspektiven. In: Streit E, Rist L, editors. *Ethik und Wissenschaft in der Anthroposophischen Medizin. Beiträge zu einer Erneuerung der Medizin*. Bern: Peter Lang Verlag; 2006:215–241.
32. Klasen J. *Autoimmun-Erkrankungen*. Stuttgart: Trias-Verlag; 2011.
33. Simon L, Schietzel T, Gärtner C, Kümmell HC, Schulte M. Ein anthroposophisches Therapiekonzept für entzündlich-rheumatische Erkrankungen. Ergebnisse einer zweijährigen Pilotstudie. [An anthroposophic therapy concept for inflammatory rheumatoid disorders. Results of a two-year pilot study]. *Forsch Komplementärmed*. 1997;4:17–27.
34. Cameron M, Gagnier JJ, Chrubasik S. Herbal therapy for treating rheumatoid arthritis. *Cochrane Database Syst Rev*. 2011;(2):CD002948.
35. Han A, Robinson V, Judd M, Taixiang W, Wells G, Tugwell P. Tai chi for treating rheumatoid arthritis. *Cochrane Database Syst Rev*. 2004;(3):CD004849.
36. Lee MS, Shin BC, Ernst E. Acupuncture for rheumatoid arthritis: a systematic review. *Rheumatology (Oxford)*. 2008;47(12):1747–1753.
37. Macfarlane GJ, El-Metwally A, De SV, Ernst E, Dowds GL, Moots RJ. Evidence for the efficacy of complementary and alternative medicines in the management of rheumatoid arthritis: a systematic review. *Rheumatology (Oxford)*. 2011;50(9):1672–1683.
38. Macfarlane GJ, Paudyal P, Doherty M, et al. A systematic review of evidence for the effectiveness of practitioner-based complementary and alternative therapies in the management of rheumatic diseases: rheumatoid arthritis. *Rheumatology (Oxford)*. 2012;51(9):1707–1713.
39. Leichsenring F. Randomized controlled versus naturalistic studies: a new research agenda. *Bull Menninger Clin*. 2004;68(2):137–151.
40. Arnett FC, Edworthy SM, Bloch D. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum*. 1988;31:315–324.

41. Burkhard G. *Biographical Work: The Anthroposophical Basis*. Edinburgh: Floris Books; 2007.
42. Kjeldsen-Kragh J. Rheumatoid arthritis treated with vegetarian diets. *Am J Clin Nutr*. 1999;70(3 Suppl):594S–600S.
43. Mason JH, Anderson JJ, Meenan RF, Haralson KM, Lewis-Stevens D, Kaine JL. The rapid assessment of disease activity in rheumatology (radar) questionnaire. Validity and sensitivity to change of a patient self-report measure of joint count and clinical status. *Arthritis Rheum*. 1992;35(2):156–162.
44. Mau W, Zeidler H. The internist's differential diagnosis of low back pains (see comments). *Dtsch Med Wochenschr*. 1995;120(46):1589–1594.
45. Larsen A, Thoen J. Hand radiography of 200 patients with rheumatoid arthritis repeated after an interval of one year. *Scand J Rheumatol*. 1987;16(6):395–401.
46. Rau R, Herborn G. A modified version of Larsen's scoring method to assess radiologic changes in rheumatoid arthritis. *J Rheumatol*. 1995;22(10):1976–1982.
47. Rau R, Herborn G. Healing phenomena of erosive changes in rheumatoid arthritis patients undergoing disease-modifying antirheumatic drug therapy. *Arthritis Rheum*. 1996;39(1):162–168.
48. Rau R, Herborn G, Menninger H, Blechschmidt J. Comparison of intramuscular methotrexate and gold sodium thiomalate in the treatment of early erosive rheumatoid arthritis: 12 month data of a double-blind parallel study of 174 patients. *Br J Rheumatol*. 1997;36(3):345–352.
49. Rau R, Wassenberg S, Herborn G, Stucki G, Gebler A. A new method of scoring radiographic change in rheumatoid arthritis. *J Rheumatol*. 1998;25(11):2094–2107.
50. Rau R, Wassenberg S, Herborn G, Perschel WT, Freitag G. Identification of radiologic healing phenomena in patients with rheumatoid arthritis. *J Rheumatol*. 2001;28(12):2608–2615.
51. Wassenberg S, Rau R. Problems in evaluating radiographic findings in rheumatoid arthritis using different methods of radiographic scoring: examples of difficult cases and a study design to develop an improved scoring method. *J Rheumatol*. 1995;22(10):1990–1997.
52. Steinbrocker O, Traeger CH, Battermann RC. Therapeutic criteria in rheumatoid arthritis. *J Am Med Assoc*. 1949;140(8):659–662.
53. van der Heijde DM, van 't Hof MA, van Riel PL, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. *Ann Rheum Dis*. 1990;49(11):916–920.
54. Ritchie DM, Boyle JA, McInnes JM, et al. Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis. *Q J Med*. 1968;37(147):393–406.
55. TORPEDO study. A study on rapid effect of tocilizumab in patients with rheumatoid arthritis with an inadequate response to disease-modifying antirheumatic drugs (DMARDs) or anti-TNF. *ClinicalTrials.gov*. 2014. Available from: https://clinicaltrials.gov/ct2/show/results/NCT00977106. Accessed January 27, 2016.
56. Harris ED. *Rheumatoid Arthritis*. Philadelphia: Saunders; 1997.
57. Koopman WJ, Boulware DW, Heudebert GR. *Clinical Primer of Rheumatology*. Philadelphia: Lippincott Williams & Wilkins; 2003.
58. Kaiser H, Kley HK. *Cortisontherapie. Corticoide in Klinik und Praxis*. 10th ed. Stuttgart: Thieme Verlag; 1997.
59. Sharp JT, Wolfe F, Mitchell DM, Bloch DA. The progression of erosion and joint space narrowing scores in rheumatoid arthritis during the first twenty-five years of disease. *Arthritis Rheum*. 1991;34(6):660–668.
60. Dixon JS, Bird HA, Sitton NG, Pickup ME, Wright V. C-reactive protein in the serial assessment of disease activity in rheumatoid arthritis. *Scand J Rheumatol*. 1984;13(1):39–44.
61. van der Heijde DM, van Riel PL, van Leeuwen MA, van 't Hof MA, van Rijswijk MH, van de Putte LB. Prognostic factors for radiographic damage and physical disability in early rheumatoid arthritis. A prospective follow-up study of 147 patients. *Br J Rheumatol*. 1992;31(8):519–525.
62. van der Heijde D, Dankert T, Nieman F, Rau R, Boers M. Reliability and sensitivity to change of a simplification of the Sharp/van der Heijde radiological assessment in rheumatoid arthritis. *Rheumatology (Oxford)*. 1999;38(10):941–947.
63. Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health*. 2004;94(3):361–366.
64. Raspe HH, Hagedorn U, Kohlmann T, Mattussek S. Der Funktionsfragebogen Hannover (FFbH): Ein Instrument zur Funktionsdiagnostik bei polyartikulären Gelenkerkrankungen. In: Siegrist J, editor. *Wohnortnahe Betreuung Rheumakranker. Ergebnisse sozialwissenschaftlicher Evaluation eines Modellversuches*. Stuttgart: Schattauer Verlag; 1990: 164–1823.
65. Korpela M, Laasonen L, Hannonen P, et al. Retardation of joint damage in patients with early rheumatoid arthritis by initial aggressive treatment with disease-modifying antirheumatic drugs: five-year experience from the FIN-RACo study. *Arthritis Rheum*. 2004;50(7):2072–2081.
66. Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Allaart CF, et al. Clinical and radiographic outcomes of four different treatment strategies in patients with early rheumatoid arthritis (the BeSt study): a randomized, controlled trial. *Arthritis Rheum*. 2005;52(11):3381–3390.
67. Wassenberg S, Rau R, Steinfeld P, Zeidler H. Very low-dose prednisolone in early rheumatoid arthritis retards radiographic progression over two years: a multicenter, double-blind, placebo-controlled trial. *Arthritis Rheum*. 2005;52(11):3371–3380.
68. Rantalaiho V, Korpela M, Laasonen L, et al. Early combination disease-modifying antirheumatic drug therapy and tight disease control improve long-term radiologic outcome in patients with early rheumatoid arthritis: the 11-year results of the Finnish Rheumatoid Arthritis Combination Therapy trial. *Arthritis Res Ther*. 2010;12(3):R122.
69. Rantalaiho V, Kautiainen H, Korpela M, et al. Targeted treatment with a combination of traditional DMARDs produces excellent clinical and radiographic long-term outcomes in early rheumatoid arthritis regardless of initial infliximab. The 5-year follow-up results of a randomised clinical trial, the NEO-RACo trial. *Ann Rheum Dis*. 2014;73(11):1954–1961.
70. van der Kooij SM, Goekoop-Ruiterman YP, de Vries-Bouwstra JK, et al. Drug-free remission, functioning and radiographic damage after 4 years of response-driven treatment in patients with recent-onset rheumatoid arthritis. *Ann Rheum Dis*. 2009;68(6):914–921.
71. Verstappen SM, Jacobs JW, Bijlsma JW, et al. Five-year followup of rheumatoid arthritis patients after early treatment with disease-modifying antirheumatic drugs versus treatment according to the pyramid approach in the first year. *Arthritis Rheum*. 2003;48(7):1797–1807.
72. Smythe HA, Helewa A, Goldsmith CH. "Independent assessor" and "pooled index" as techniques for measuring treatment effects in rheumatoid arthritis. *J Rheumatol*. 1977;4(2):144–152.
73. Markatseli TE, Papagoras C, Drosos AA. Prognostic factors for erosive rheumatoid arthritis. *Clin Exp Rheumatol*. 2010;28(1):114–123.
74. Smolen JS, Landewe R, Breedveld FC, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. *Ann Rheum Dis*. 2010;69(6):964–975.
75. Krüger K, Wollenhaupt J, Albrecht K, et al. S1-Leitlinie der DGRh zur sequenziellen medikamentösen Therapie der rheumatoiden Arthritis 2012. Adaptierte EULAR-Empfehlungen und aktualisierter Therapiealgorithmus [German 2012 guidelines for the sequential medical treatment of rheumatoid arthritis. Adapted EULAR recommendations and updated treatment algorithm]. *Z Rheumatol*. 2012;71(7):592–603.
76. IVAA–International Federation of Anthroposophic Medical Associations; 2012. *Facts and figures on Anthroposophic Medicine worldwide*. Available from: http://www.ivaa.info. Accessed August 15, 2013.
77. Kienle GS, Albonico HU, Baars E, Hamre HJ, Zimmermann P, Kiene H. Anthroposophic medicine: an integrative medical system originating in Europe. *Glob Adv Health Med*. 2013;2(6):20–31.
78. Hamre HJ, Glockmann A, Schwarz R, et al. Antibiotic use in children with acute respiratory or ear infections: prospective observational comparison of anthroposophic and conventional treatment under routine primary care conditions. *Evid Based Complement Alternat Med*. 2014:2014:243801.

79. Simon L. Prescribing anthroposophic medicines: Substances, pharmaceutical processes, principles of effectiveness. In: Arendt A, Debus M, Karutz M, et al, eds. *Vademecum of Anthroposophic Medicines, 3rd English Edition*. Filderstadt: Association of Anthroposophic Physicians in Germany (GAÄD); 2017:579–647.

80. Kienle GS, Kiene H, Albonico HU. *Anthroposophic medicine*: effectiveness, utility, costs, safety. Stuttgart, New York: Schattauer Verlag; 2006.

81. Jones G, Maier J, Pedersen P, Schwarz R, Mennet-von Eiff M, editors. *Anthroposophic Pharmaceutical Codex APC, Fourth Edition*. Dornach: International Association of Anthroposophic Pharmacists; 2017.

82. Witt CM, Bluth M, Albrecht H, Weisshuhn TE, Baumgartner S, Willich SN. The in vitro evidence for an effect of high homeopathic potencies – a systematic review of the literature. *Complement Ther Med*. 2007;15(2):128–138.

83. Hamre HJ, Kiene H, Kienle GS. Clinical research in anthroposophic medicine. *Altern Ther Health Med*. 2009;15(6):52–55.

84. Kirchner-Bockholt M. *Fundations of curative eurythmy*. 2nd ed. Edinburgh: Floris Books; 2004.

85. Hamre HJ, Witt CM, Glockmann A, Ziegler R, Willich SN, Kiene H. Anthroposophic art therapy in chronic disease: a four-year prospective cohort study. *Explore*. 2007;3(4):365–371.

86. Büssing A, Cysarz D, Edelhäuser F, Bornhoft G, Matthiessen PF, Ostermann T. The oil-dispersion bath in anthroposophic medicine – an integrative review. *BMC Complement Altern Med*. 2008;8(61).

87. Layer M. *Handbook for Rhythmical Einreibungen according to Wegman/Hauschka*. Forest Row: Temple Lodge Press; 2006.

88. Hamre HJ, Witt CM, Glockmann A, Ziegler R, Willich SN, Kiene H. Anthroposophic medical therapy in chronic disease: a four-year prospective cohort study. *BMC Complement Altern Med*. 2007;7(10).

Supplementary materials

Table S1 Patient numbers at each follow-up

Month	Anthroposophic group						Conventional group					
	Scheduled		Follow-up		No follow-up		Scheduled		Follow-up		No follow-up	
	n	Percent	n	Percent	n	Percent	n	Percent	n	Percent	n	Percent
3	129	100.0	127	98.4	2	1.6	122	100.0	115	94.3	7	5.7
6	129	100.0	123	95.3	6	4.7	122	100.0	110	90.2	12	9.8
9	129	100.0	119	92.2	10	7.8	122	100.0	106	86.9	16	13.1
12	129	100.0	118	91.5	11	8.5	122	100.0	104	85.2	18	14.8
15	129	100.0	116	89.9	13	10.1	122	100.0	102	83.6	20	16.4
18	129	100.0	113	87.6	16	12.4	122	100.0	101	82.8	21	17.2
21	129	100.0	106	82.2	23	17.8	122	100.0	100	82.0	22	18.0
24	129	100.0	96	74.4	33	25.6	122	100.0	80	65.6	42	34.4
27	123	100.0	93	75.6	30	24.4	106	100.0	75	70.8	31	29.2
30	123	100.0	90	73.2	33	26.8	106	100.0	74	69.8	32	30.2
33	123	100.0	86	69.9	37	30.1	106	100.0	73	68.9	33	31.1
36	123	100.0	81	65.9	42	34.1	106	100.0	66	62.3	40	37.7
39	118	100.0	77	65.3	41	34.7	103	100.0	66	64.1	37	35.9
42	118	100.0	74	62.7	44	37.3	103	100.0	65	63.1	38	36.9
45	118	100.0	73	61.9	45	38.1	103	100.0	65	63.1	38	36.9
48	118	100.0	72	61.0	46	39.0	103	100.0	65	63.1	38	36.9

Table S2 Dropout analyses

Baseline characteristics	48-month follow-up data available?		
	Yes (n=137)	No (n=114)	*p*-value
Female gender, n (%)	109/137 (79.6%)	99/114 (86.8%)	0.1275
Age (years)	50.96±11.28	48.02±12.01	0.0413
Duration of rheumatoid arthritis (months)	14.44±9.76	12.23±8.93	0.0813
Rheumatoid factor (IU/mL)	148.22±217.62	188.35±276.85	0.6467
Rheumatoid factor positive (≥20 IU/mL), n (%)	112/129 (86.8%)	92/109 (83.6%)	0.4875
C-reactive protein (mg/dL)	2.77±4.74	2.45±3.03	0.6224
Joint tenderness (Ritchie Index, 0–78)	13.14±8.02	13.34±9.77	0.7506
Number of swollen joints	14.11±7.90	15.25±7.64	0.2469
Hanover Functional Ability Questionnaire	17.75±4.44	16.92±5.04	0.2597
Radiological signs of joint destruction (Ratingen Score)	2.12±4.05	2.52±4.58	0.5744
Erosive arthritis with Ratingen Score ≥1, n (%)	45/137 (32.8%)	35/104 (33.7%)	0.8954
Corticosteroid use, n (%)	54/137 (39.4%)	38/114 (33.3%)	0.3194

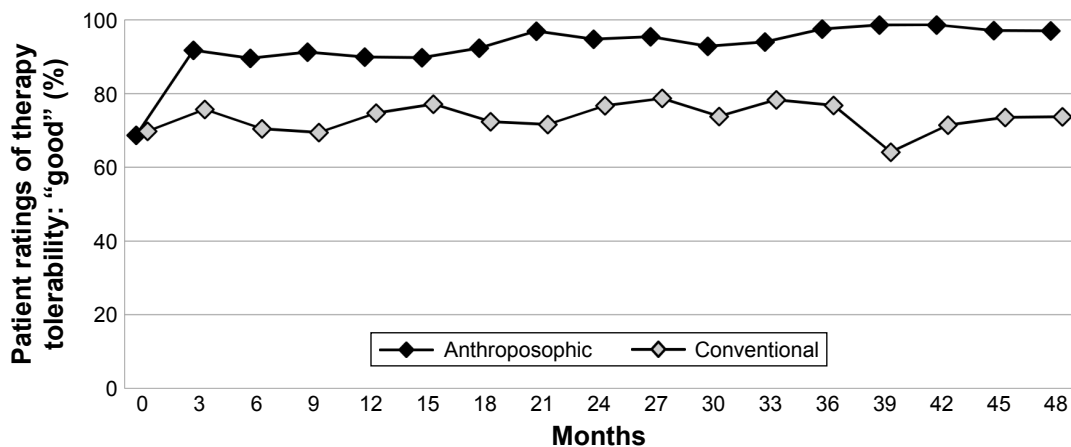**Note:** Data are presented as mean ± SD, unless otherwise stated.

Figure S1 Patient ratings of therapy tolerability ("poor"–"average"–"good").
Note: Percentage of patients with the rating "good tolerability" are as follows: in months 0 (Anthroposophic group: n=99, Conventional group: n=86), 3 (122+107), 6 (115+105), 9 (114+95), 12 (109+95), 15 (107+92), 18 (105+87), 21 (98+88), 24 (95+86), 27 (87+61), 30 (84+61), 33 (84+60), 36 (82+56), 39 (70+53), 42 (73+56), 45 (69+53), and 48 (68+57).
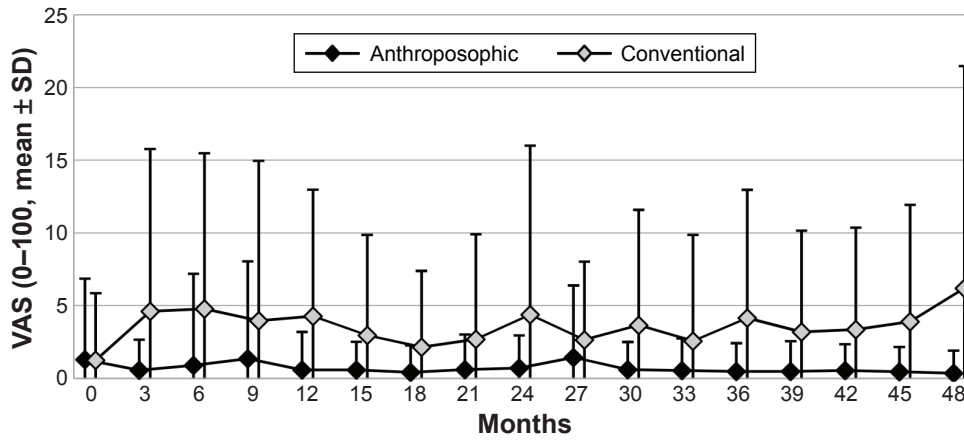
Figure S2 Physician ratings of therapy tolerability.
Note: Physician ratings of therapy tolerability on a VAS (0= no adverse reactions, 100= most severe adverse reactions) in months 0 (Anthroposophic group: n=99, Conventional group: n=86), 3 (126+111), 6 (122+110), 9 (118+107), 12 (117+102), 15 (115+99), 18 (112+94), 21 (103+94), 24 (100+92), 27 (92+69), 30 (89+68), 33 (85+68), 36 (83+65), 39 (75+60), 42 (72+61), 45 (71+59), and 48 (71+64).
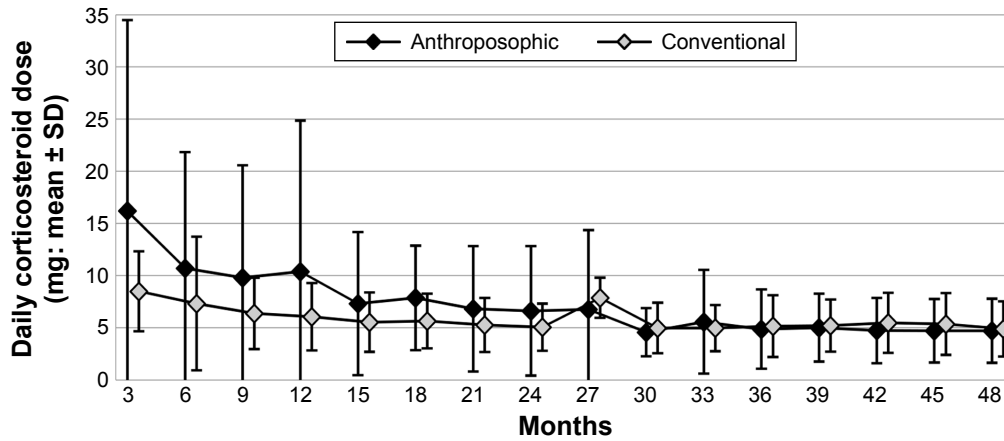**Abbreviation:** VAS, visual analog scale.

Figure S3 Use of corticosteroids in months 0–48.
Note: Use is calculated as average dose in prednisolone equivalents among steroid users in each 3-month period ("3" indicating "months 1–3" and so on).
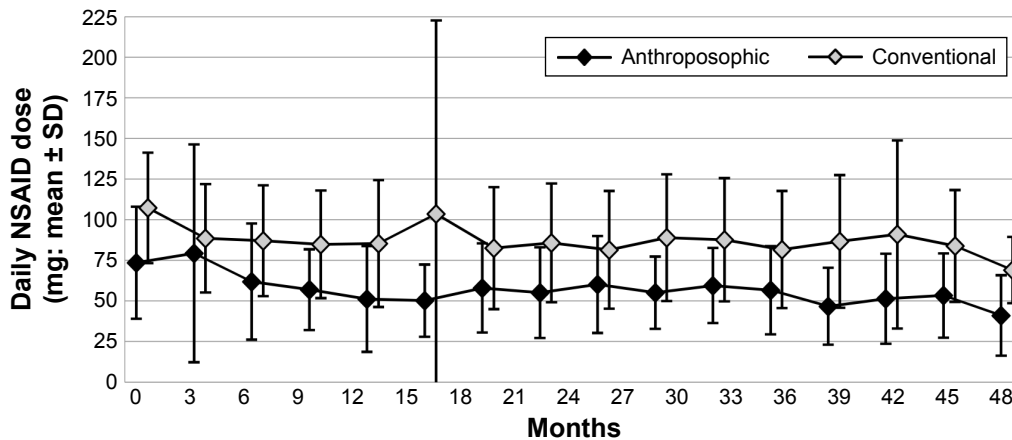
Figure S4 Use of NSAIDs in months 0–48.
Note: Use is calculated as average dose in diclofenac equivalents among NSAID users in each 3-month period ("3" indicating "months 1–3" and so on).
**Abbreviation:** NSAID, nonsteroidal anti-inflammatory drug.

Table S3 RA-VAS and C-reactive protein, alternative analyses

	Anthroposophic group			Conventional group		
	n	Mean ± SD	Median	n	Mean ± SD	Median
RA-VAS (0–400)						
Month 0	127	193.9±87.2	191.0	119	238.3±93.4	240.0
Month 48	71	79.4±75.5	44.0	60	97.4±82.1	86.0
Change: months 0–48	70	−116.6±109.2	−110.5	58	−122.7±99.2	−117.5
Wilcoxon test	$p<0.0001$			$p<0.0001$		
Change in A- vs C-groups*	$p=0.4961$					
C-reactive protein (mg/dL)						
Month 0	129	2.59±3.49	1.00	109	2.66±4.64	1.40
Month 48	72	1.06±1.63	0.50	64	0.92±1.17	0.35
Change: months 0–48	72	−1.64±3.59	−0.21	57	−2.03±6.00	−0.63
Wilcoxon test	$p<0.0001$			$p<0.0001$		
Change in A- vs C-groups*	$p=0.6428$					

Note: *Mann–Whitney U test.
Abbreviatios: A- vs C-groups, Anthroposophic group versus Conventional group; RA-VAS, Rheumatoid Arthritis Visual Analog Scale.
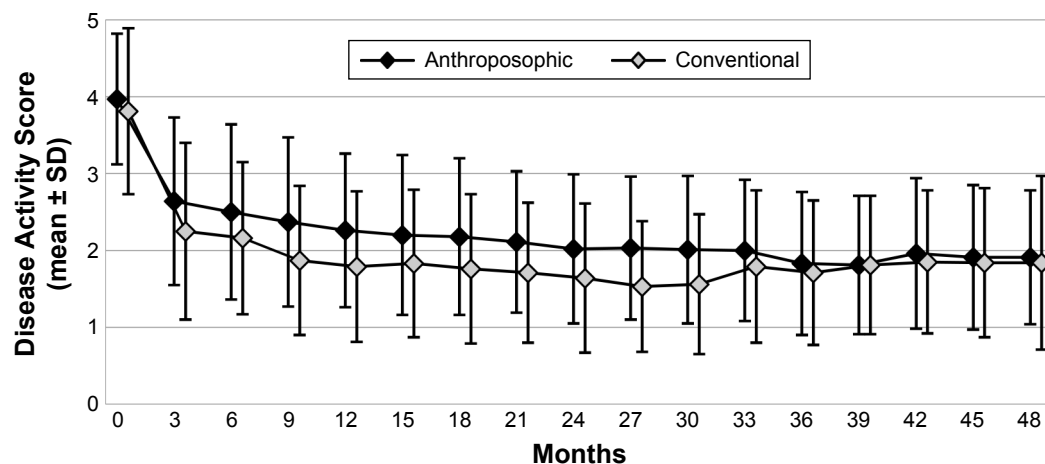
Figure S5 Disease Activity Score (0–10).
Note: Disease Activity Score (0–10) in months 0 (Anthroposophic group: n=129, Conventional group: n=108), 3 (176+104), 6 (123+103), 9 (118+99), 12 (118+99), 15 (115+94), 18 (111+90), 21 (105+90), 24 (101+89), 27 (93+66), 30 (89+64), 33 (85+64), 36 (86+63), 39 (74+56), 42 (73+59), 45 (72+55), and 48 (71+63).