

Amitriptyline to Relieve Pain in Juvenile Idiopathic Arthritis: A Pilot Study Using Bayesian Metaanalysis of Multiple N-of-1 Clinical Trials

ADAM M. HUBER, GEORGE A. TOMLINSON, GIDEON KOREN, and BRIAN M. FELDMAN

ABSTRACT. *Objective.* Using serial N-of-1 trials and subsequent analysis with Bayesian methods may allow study of therapies using small numbers of subjects. Our research questions were: (1) Can serial N-of-1 trials analyzed with Bayesian statistical techniques be used to estimate the population effect of a therapeutic intervention? (2) Compared to placebo, how likely is it that low-dose amitriptyline therapy in children aged 10–18 years with active polyarticular-course juvenile idiopathic arthritis (JIA) results in a significant improvement in pain?

Methods. Six children (age 10.3–16.3 yrs, 4 girls) were enrolled. There were 3 pairs of randomized, double-blinded treatments (amitriptyline 25 mg or placebo) per participant. Each treatment lasted 2 weeks, with a 1 week washout. The primary outcome was pain, measured by 10 cm visual analog scale. Assessments were at the beginning and end of each treatment. A Bayesian statistical model was used to determine the treatment effect. Values < 0 indicated superiority of amitriptyline.

Results. Bayesian techniques were used successfully to obtain estimates of population effect, despite the small number of participants. The mean treatment effect for pain was 0.67 (SD 0.89, 95% credible interval –0.99, 2.55). The probability that the treatment effect was < 0 was only 16%.

Conclusion. These methods can be used successfully to estimate population effects when sample sizes are small. It is unlikely that amitriptyline reduced pain by a clinically significant amount in these children with polyarticular JIA. These methods may be particularly suited to pilot studies and the study of rare illnesses. (First Release April 15 2007; J Rheumatol 2007;34:1125–32)

Key Indexing Terms:

JUVENILE IDIOPATHIC ARTHRITIS PAIN METAANALYSIS BAYESIAN STATISTICS

Juvenile idiopathic arthritis (JIA) is the most prevalent form of rheumatic illness in children¹. In the past, relatively little attention was paid to the problem of pain in childhood arthritis. It has been suggested that children with arthritis have relatively little pain, and have less pain than adults with arthritis^{2,3}. However, there is an increasing body of literature describing the pain experience of children with arthritis and documenting its effects on children and their families. Pain is now known to be common in children with arthritis⁴⁻⁶, to

interfere with physical, social and school function^{4,7-9}, and to be associated with features of depression and anxiety^{4,8-10}. Interestingly, pain has been shown to have poor correlation with expected predictors, such as disease activity, joint counts, and arthritis type¹¹⁻¹³. This highlights the poor understanding of the causes and influences of arthritis pain in children. It also emphasizes the importance of non-disease-specific factors in the assessment and management of pain in patients with pediatric arthritis.

Current therapies have resulted in excellent disease control and prevention of longterm complications in most children with JIA. Usually, disease control can be equated with pain control. However, there remains a small group of children who continue to experience pain, despite apparent improvements in inflammatory findings. Appropriate management for this group of children, who may have some features of fibromyalgia (FM), is unclear.

It has been suggested that the use of low-dose antidepressants (such as amitriptyline) can reduce pain in a variety of inflammatory and noninflammatory painful conditions. Koh, *et al* studied 100 adult patients with ankylosing spondylitis treated with 30 mg amitriptyline daily at bedtime¹⁴. They described improvements in pain, sleep, fatigue, and disease activity, although only the improvement in disease activity was statistically significant. Rani, *et al* studied a mixed group

From the Division of Pediatric Rheumatology, IWK Health Centre and Dalhousie University, Halifax, Nova Scotia; Division of Rheumatology, The Hospital for Sick Children, Toronto; and Departments of Pediatrics, Health Policy, Management and Evaluation and Public Health Sciences, University of Toronto, Toronto, Ontario, Canada.

Supported by the Canadian Institutes of Health Research. Dr. Feldman holds a Canada Research Chair in Childhood Arthritis.

A.M. Huber, MSc, MD, Division of Pediatric Rheumatology, IWK Health Centre and Dalhousie University; G.A. Tomlinson, MSc, PhD, Department of Health Policy, Management and Evaluation, University of Toronto; G. Koren, MD, The Hospital for Sick Children and University of Toronto; B.M. Feldman, MSc, MD, Division of Rheumatology, The Hospital for Sick Children, Departments of Pediatrics, Health Policy, Management and Evaluation and Public Health Sciences, University of Toronto.

Address reprint requests to Dr. A.M. Huber, Division of Pediatric Rheumatology, IWK Health Centre, 5850 University Avenue, Halifax, Nova Scotia B3K 6R8, Canada. E-mail: adam.huber@iwk.nshealth.ca
Accepted for publication January 18, 2007.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2007. All rights reserved.

of adult patients, including 16 with osteoarthritis and 8 with rheumatoid arthritis, who received amitriptyline 25 mg daily at bedtime¹⁵. Patients in the amitriptyline group had greater reductions in pain intensity and greater perceived pain relief than those in the placebo group. Finally, the use of amitriptyline has been extensively studied in FM^{16,17}. The mechanism of amitriptyline acting as a pain-reducing medication is unknown. Proposed mechanisms include inhibition of reuptake of a variety of neurotransmitters leading to reduced perception of pain, potentiation of the effects of endogenous opioids¹⁸, and possible direct antiinflammatory effects¹⁹. Thus, amitriptyline appears to be a reasonable option for treating pain that has not responded to antiinflammatory therapy in inflammatory disease.

There are no data concerning the use of amitriptyline in children with arthritis. However, it has been successfully and safely used in other painful conditions such as migraine^{20,21}, chronic abdominal pain²², and pediatric FM²³. Given the clinical impression that some of the children with arthritis and pain that is unresponsive to the usual antiinflammatory therapy appear to have features similar to FM, amitriptyline would appear to be a reasonable potential therapy.

As a first step to investigate this problem, we were interested in using a methodology that would allow us to conduct a pilot study with a very small number of patients, while still obtaining estimates of the effect of the treatment intervention. The availability of such a methodology would be of considerable benefit to the study of this problem, and to the study of rare clinical problems in general. It would also be valuable in pilot work to provide preliminary data prior to embarking on full-scale clinical trials.

Studies of small numbers of patients are common in the psychological and behavioral literature²⁴. Many of these use an N-of-1 experimental design. There is controversy regarding the best way to analyze this type of data. Some have advocated techniques called “visual analysis,” which are based largely on inspection and graphing²⁵⁻²⁷. The preference for these visual techniques was based on a perception of their simplicity, the lack of a need for formal knowledge of statistics, and an opinion that these methods emphasized clinical rather than statistical significance^{28,29}. However, other authors have expressed concerns about the potential for bias, the lack of standardized methods, and poor inter- and intrarater reliability^{24,30}. For these reasons we do not consider visual analysis appropriate for the study of therapy in rheumatology.

In 1991, Jaeschke and colleagues published their experience using N-of-1 trials of amitriptyline in FM³¹. By exposing each patient to multiple periods of amitriptyline or placebo, they were able to evaluate the effectiveness of amitriptyline for each patient. By allowing each patient to act as his/her own control, they eliminated between-subject variability and markedly reduced overall error, thereby increasing power. This work did not attempt to draw any conclusions about the

treatment and its effectiveness in the general population. However, in 1997, Zucker, *et al* reanalyzed a portion of the data from the Jaeschke report using a Bayesian metaanalytic statistical model³². Bayesian statistics are an alternative approach to traditional frequentist statistics, and have several characteristics that make them potentially very powerful for the analysis of small data sets. These include the use of direct probabilities, the ability to obtain results in the context of small data sets (i.e., in circumstances where frequentist statistics would only fail to reject the null hypothesis), insensitivity to repeated analyses, and the ability to explicitly consider previously available information (such as available from previous studies). Zucker and colleagues were able to obtain an estimate of the effect of amitriptyline in this population, despite a sample size of only 23.

For this project, we conducted a pilot study of the effectiveness of amitriptyline for reduction of pain in children aged 10–18 years with polyarticular course JIA. We analyzed the data using a hierarchical Bayesian statistical model, adapted from the work presented by Zucker, *et al*³², and applied it in a novel way for this prospective study. A unique aspect of this study was the intentional use of a very small number of subjects. We were interested in determining if the study goals could be met by such a small data set. The specific research questions we posed were:

1. Is it feasible to use serial N-of-1 trials, and their subsequent analysis with Bayesian statistical techniques, to estimate the population effect of a therapeutic intervention?
2. Compared to placebo, how likely is it that low-dose amitriptyline in children aged 10–18 years with active polyarticular course JIA results in a clinically significant reduction in pain (i.e., likely enough to justify further study)?

MATERIALS AND METHODS

Patients. Although JIA is reasonably prevalent, painful disease in children not otherwise warranting escalation of antiinflammatory therapy is a rare presentation. Six children were enrolled from The Hospital for Sick Children, Toronto, and The IWK Health Centre, Halifax, between March 2000 and March 2003. All children who were approached agreed to participate. Their characteristics are summarized in Table 1. Eligibility criteria were age \geq 10 years, polyarticular course JIA with at least 1 active joint, minimum pain assessment of 1 cm on 10 cm visual analog scale (VAS), and stable doses of usual medications. Children were excluded if there was any known contraindication to amitriptyline, amitriptyline therapy in the preceding month, if they had another indication for amitriptyline therapy, if they were unable to speak English, or if the patient had another illness that would put them at increased risk of adverse outcomes. Research ethics board approval was obtained from both participating centers, and all participants and their parent/legal guardian provided written informed consent.

Measures. The primary outcome was pain, measured by self-report completion of a 10 cm VAS. During each assessment, participants completed 3 VAS per day (morning, afternoon, and before bed) for 2 days. Multiple daily assessments were used to increase reliability, as described by Jensen and McFarland³³.

There were several other secondary outcome measures. Sleep and fatigue were thought to be involved in possible mechanisms of amitriptyline effect, and data for them were collected. Quality of sleep was assessed in 2 ways. An 8 item questionnaire was developed, adapted from Bloom, *et al*³⁴. Scores for

Table 1. Characteristics of study participants. Values are medians (minimum, maximum).

Age, yrs	12.7 (10.3, 16.3)
Sex	4 female: 2 male
Ethnicity	6 Caucasian
Disease duration, yrs	5 (1.1, 15)
Baseline pain, VAS	3.1 (1.2, 5.9)
Baseline sleep, questionnaire	35.5 (20, 48)
Baseline sleep VAS	7.55 (1.3, 10)
Baseline fatigue, questionnaire	33.5 (15, 39)
Baseline stiffness VAS	3.8 (0.5, 9.5)
Baseline CHAQ	0.45 (0.25, 0.88)
Baseline swollen joints	1 (0, 2)
Baseline active joints	6 (1, 14)
Baseline physician global VAS	1.8 (1.1, 3.1)
Baseline patient global VAS	5.0 (2.4, 9.4)

VAS: visual analog scale, CHAQ: Childhood Health Assessment Questionnaire.

this questionnaire ranged from 8 (very good sleep) to 56 (very poor sleep). There was also a 10 cm VAS response to the question, "How good or bad do you feel your sleep has been over the last week?". Fatigue was assessed using a questionnaire adapted for this study from the Fatigue Severity Scale³⁵. The resulting questionnaire gave a score ranging from 9 (little or no fatigue) to 63 (severe fatigue). Morning stiffness was assessed with a 10 cm VAS response to the question, "How severe has your morning stiffness been on average over the last week?". Physical function was assessed with the Childhood Health Assessment Questionnaire (CHAQ). The CHAQ consists of 30 items in 8 domains, and gives a score from 0 (no or mild physical disability) to 3 (severe physical disability)³⁶.

In order to increase comparability of therapeutic trials in childhood arthritis, a core set of variables has been developed that should be collected in all clinical trials³⁷. Consistent with this recommendation we collected the physician global assessment of overall disease activity (by 10 cm VAS), the parent/patient global assessment of disease severity (10 cm VAS), and the number of active joints. No blood investigations were done as part of this study.

It was determined *a priori* that the clinically significant difference between pain VAS for amitriptyline and placebo periods was 0.8 cm³⁸. A difference of 0.8 cm was also considered clinically significant for other VAS scores. For the CHAQ, a difference of 0.13 was considered clinically significant³⁹.

Procedures. Potential participants were identified through review of patients attending pediatric rheumatology clinics at both hospitals. Inclusion and exclusion criteria were checked, and all participants had an electrocardiogram.

Each participant completed a series of N-of-1 trials. These took the form of 3 paired crossovers (6 treatment periods) per participant. Each treatment pair consisted of one period where the treatment was with amitriptyline, and one period where the treatment was with an identical-appearing placebo. The order within each pair was determined randomly, and the order within each treatment pair was independent of the other treatment periods. These treatment assignments were determined by the research pharmacist using a computer generated random number list, and placed in sealed, numbered envelopes that were only opened at the time of study consent. Each treatment period had a duration of 2 weeks, and was separated from the next treatment period by a 1 week washout when no medication was given (time course based on Jaeschke, *et al*³¹ and clinical experience of the investigators). The dose of amitriptyline was 25 mg by mouth 1–2 hours before bedtime. Both amitriptyline and placebo tablets were crushed and placed within identical gel capsules. Participants, families, and all individuals associated with the study except for the research pharmacist were blinded to treatment allocation. Unblinding was performed at the end of the entire study.

At the initial study visit, a baseline pain assessment and physical exami-

nation were performed. The following baseline measures were also administered: sleep questionnaire, modified Fatigue Severity scale, morning stiffness, CHAQ, swollen joint count, active joint count, screening examination for tender points, physician global assessment of disease activity, and parent/patient global assessment of disease activity.

Study participants were seen at the end of each treatment period. Pain was assessed by completion of three 10 cm VAS daily for the 2 days prior to the study visit. The measures described for the baseline visit were also repeated. In addition, information about adverse events was elicited using a standardized questionnaire and spontaneous report.

At the beginning of each treatment period, participants completed self-administered assessments at home, giving a new baseline for each treatment period. These included pain VAS, sleep questionnaire, modified Fatigue Severity scale, morning stiffness, CHAQ, and parent/patient global assessment of disease activity. For those measures that could not be self-administered (e.g., joint counts), the results from the end of the previous treatment period were carried forward.

Analysis. Data were entered into an Access database (v. 10.0, Microsoft Corp., Seattle, WA, USA) and imported to SAS (v. 8.00, SAS Institute Inc., Cary, NC, USA) for data manipulation. Small data files were then exported to WinBugs (v. 1.4, Imperial College and Medical Research Council, UK, available at <http://www.mrc-bsu.cam.ac.uk/WinBUGS/>) for the primary analysis.

The primary analysis was the difference in pain reduction between amitriptyline and placebo treatment periods. This was the "treatment effect for pain." There were a total of 72 pain scores recorded for each subject, assuming complete data [6 pretreatment scores and 6 posttreatment scores for each of 6 treatment periods = (6 + 6) × 6]. If the medication being taken during the period was effective, it was assumed that there would be a reduction in the pain score. The superior therapy should have acted to reduce the pain score to a greater degree.

For the primary Bayesian analysis, a hierarchical random-effects model was constructed. The choice of a hierarchical model is important because the independence of the data cannot be assumed. Indeed, it is expected that data obtained from the same subject will be more similar than data obtained from different patients. A random-effect model was chosen because it could not be assumed that the effects would be constant across treatment pairs or subjects. The multiply-nested structure of the study design was represented by nested random effects for pair within subject, treatment within pair, and pre/post within treatment. The interaction of pre/post with treatment represents the treatment effect in this model. It is the incremental (i.e., additional) benefit of being in the post phase after receiving amitriptyline versus being in the post phase after receiving placebo (see Appendix). For this analysis, values < 0 cm indicated a superior treatment effect of amitriptyline. Estimates of the overall treatment effect and subject-specific treatment effects were obtained, as well as the probabilities that these values were < 0 cm (probability of any benefit) and < -0.8 cm (*a priori* determined significant reduction in pain).

For the Bayesian analysis, noninformative priors were used for all parameters (i.e., assuming that there was no information before the study about whether amitriptyline might be a better pain reliever). A prior distribution is a representation of the state of knowledge before any data are collected in a study. The overall mean treatment effect was assumed to have a normal prior distribution with mean 0 and variance 10,000. This generated a distribution that was essentially flat over the plausible range of the overall mean treatment effect (i.e., all values of the treatment effect were equally likely, over a range of values that might actually be seen). The detailed model is available from the corresponding author upon request.

This analysis was repeated for all of the other measures administered. For those measures where there existed estimates of a clinically significant difference, the probability that the mean of the posterior distribution was less than this value was also calculated^{38,39}.

In order to assess the influence of assumptions about prior distributions on the primary analysis (overall treatment effect for pain), a sensitivity analysis was conducted by repeating the analysis with alternate prior distributions. As noted above, the original analysis used a noninformative prior, which assumed that there was no information available to describe the likely effect-

tiveness of amitriptyline in reducing pain in study participants. Three alternate prior distributions were examined. The first was a pessimistic prior, which assumed that it was very unlikely that amitriptyline was superior to placebo for pain reduction. This distribution was assumed to be normal with mean 0.8 cm and variance 0.17. The mean was chosen to represent a significant difference in pain as defined for this study. The variance was chosen so that the proportion of the distribution < 0 was only 0.025. The second distribution was a skeptical one, which assumed that it was unlikely that there was any difference between amitriptyline and placebo. This distribution was assumed to be normal with mean 0 cm and variance 0.17. The majority of this distribution fell between -0.8 cm and 0.8 cm, with only 0.025 of the distribution < -0.8 and > 0.8 , respectively. The final distribution was an optimistic one, which assumed that it was very likely that amitriptyline was superior to placebo. This distribution was a mirror image of the pessimistic distribution, being normal with a mean of -0.8 cm and variance 0.17. It should be noted that the low variances assigned to these priors assumes considerable certainty, which was not reflective of the information available prior to the conduct of this study. However, this sensitivity analysis provided information about how convincing the results of this study should be to readers with varying levels of belief about the effectiveness of amitriptyline prior to reading the results. For example, someone with a strong belief that amitriptyline was effective in children with arthritis would find this belief reflected in the optimistic prior distribution, and could assess whether their beliefs should be influenced by the results of this study.

An analysis using standard, frequentist methods was then performed on the pain outcome. The SAS procedure PROC MIXED was used to perform a repeated measures analysis. The explanatory variables were patient identification, treatment pair, treatment group, whether the data came from a pre- or posttreatment assessment, and the interaction of treatment group and pre/post assessment. Patient identification and treatment pair were assumed to be random effects, while the remaining variables were considered to be fixed effects. The dependant variable was pain.

RESULTS

Pain. Median pain scores for assessments 1 to 6 were 1.9, 1.5, 1.4, 1.8, 2.5, and 1.2 cm, respectively ($F = 0.57$, $p = 0.72$ by Kruskal-Wallis test), indicating no overall trend in pain scores over the study. Tender point examination was negative for all patients.

Figure 1 shows the posterior distribution that was obtained for the overall treatment effect for pain. The mean of the posterior distribution was 0.67 cm (a positive value suggested that pain reduction was greater during placebo periods) with a standard deviation of 0.89 cm (95% credible interval -0.99 cm, 2.55 cm). The probability that the treatment effect was less than 0 cm was 0.16, and the probability that it was < -0.8 cm was only 0.03 (i.e., it was 97% probable that amitriptyline did not reduce pain to a significant degree in these patients).

The posterior distributions of the treatment effect for pain for individual participants are summarized in Table 2. The individual mean treatment effect ranged from -0.48 cm to 1.84 cm. The probability that the mean treatment effect was < 0 ranged from 0.04 to 0.72, while the probability that it was < -0.8 ranged from 0.01 to 0.33. Thus, it is unlikely that any of the subjects had a clinically significant response to amitriptyline.

When the overall analysis was repeated using a more traditional repeated measures technique, no significant difference in the treatment effect between amitriptyline and placebo periods was observed ($F = 0.80$, $p = 0.37$).

Sensitivity analysis. The pessimistic prior distribution yielded

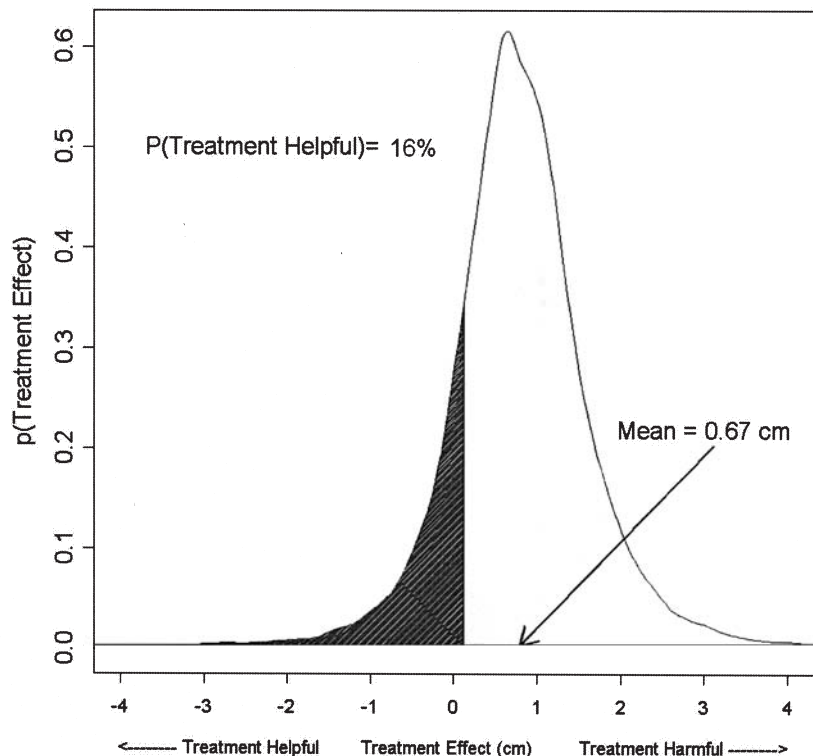


Figure 1. Posterior distribution of the treatment effect for pain (mean 0.67, standard deviation 0.89). Values < 0 represent superiority of amitriptyline for pain reduction.

Table 2. Results obtained from posterior distributions of the subject-specific treatment effects for pain. p (any benefit) is the probability that the treatment effect was less than 0 (i.e., favoring amitriptyline). p (clinically significant benefit) is the probability that the treatment effect was < -0.8.

Subject	Mean of Posterior Distribution	SD	95% Credible Interval	p (any benefit)	p (clinically significant benefit)
1	1.67	0.98	-0.15, 3.62	0.04	< 0.01
2	-0.48	0.81	-2.08, 1.04	0.72	0.33
3	1.84	1.26	-0.30, 4.44	0.05	0.01
4	0.22	0.70	-1.19, 1.64	0.37	0.06
5	0.60	0.70	-0.74, 2.03	0.19	0.02
6	0.17	0.71	-1.26, 1.54	0.39	0.07

a mean for the posterior distribution of 0.75 cm (95% credible interval 0.08 cm, 1.44 cm) with standard deviation 0.34 cm. The skeptical prior distribution yielded a mean for the posterior distribution of 0.18 cm (95% credible interval -0.52 cm, 0.86 cm) with standard deviation 0.35 cm. The optimistic prior distribution yielded a mean for the posterior distribution of -0.47 cm (95% credible interval -1.26 cm, 0.28 cm) with standard deviation 0.39 cm. Even for this optimistic prior, the probability that the treatment effect was clinically important (i.e. < -0.8) was only 0.20.

Other measures. Overall treatment effects were calculated for each of the other measures assessed in this study. These are summarized in Table 3.

Withdrawals. No participant withdrew from the study due to adverse events or side effects. Four participants completed the entire protocol. One participant completed 2 treatment pairs and then withdrew due to a personal matter unrelated to the study. One participant completed one treatment pair and then chose not to continue in the study due to concerns about the time commitment required for followup visits.

Compliance and cointervention. Overall compliance with study medication was 98%, with 452/462 doses of medication (placebo or amitriptyline) taken. For individual patients, compliance ranged from 89% to 100%. In all cases, the number of doses of medication reported to have been taken agreed with the number pills returned at the end of the treatment period.

With regard to overall compliance with the protocol, one patient had a change in medication (increase in dose of methotrexate). This occurred during treatment period 3 in a patient who completed 4 treatment periods. This period was a placebo period. Two participants underwent corticosteroid joint injections. In one participant, 2 joints were injected during period 4 (amitriptyline) in a participant who completed 4 treatment periods. The other participant had injection of 8 joints during period 5 (amitriptyline). This participant completed the entire protocol. When the analysis was repeated, with the treatment periods affected by corticosteroid injections deleted, the mean treatment effect for pain was 0.92 (standard deviation 0.97, 95% credible interval -0.82, 3.03).

Safety. No subject experienced any serious adverse events related to the study medication, and no subject withdrew from the study for medication reasons. Overall, all subjects complained of at least one side effect at some point during the study. Side effects reported were dry mouth (6/6 subjects), sedation (6/6 subjects), nausea (5/6 subjects), blurred vision (2/6 subjects), diarrhea (1/6 subjects), and sore throat (1/6 subjects). There was no difference in reporting of side effects between amitriptyline and placebo periods.

DISCUSSION

This project had 2 main goals. The first was to examine the

Table 3. Overall treatment effects of the other outcomes assessed in this study. P (any benefit) is the probability that the treatment effect was less than 0 (favoring amitriptyline). P (clinically significant benefit) is the probability that the treatment effect was < -0.8.

Subject	Mean of Posterior Distribution	SD	95% Credible Interval	p (any benefit)	p (clinically significant benefit)*
Sleep questionnaire	0.61	3.77	-6.74, 8.24	0.44	NA
Sleep VAS	0.16	1.59	-3.09, 3.24	0.44	0.24
Fatigue questionnaire	-0.75	2.84	-6.4, 2.94	0.61	NA
Morning stiffness VAS	1.04	1.76	-2.54, 4.64	0.26	0.12
CHAQ	-0.09	0.15	-0.34, 0.19	0.78	0.36 [†]
Swollen joints	-0.02	0.39	-0.77, 0.76	0.53	NA
Active joints	-1.28	1.69	-4.53, 2.26	0.80	NA
Physician global VAS	-0.22	0.75	-1.71, 1.25	0.63	0.18
Patient global VAS	-1.53	1.50	-4.46, 1.59	0.87	0.72

* Calculated for VAS scores and the CHAQ score only. [†] Probability that the CHAQ score was less than -0.13. NA: not applicable.

feasibility of using serial N-of-1 trials, analyzed using a Bayesian statistical model, to estimate the population effect of an intervention. The second goal was to assess, in a pilot study, the likelihood of effectiveness of amitriptyline as a pain-reducing intervention in children with polyarticular course JIA.

With regard to the first goal, this project was an empiric success. For each of the outcomes assessed, we obtained an estimate of the effect of the intervention, and the likelihood that this effect was different than for placebo. Not surprisingly, given the small sample, despite using multiple crossovers for each patient, the frequentist analysis was only able to indicate that no statistically significant difference between the 2 groups could be detected. Given the high likelihood of Type II error, the frequentist analysis does not provide enough information to indicate whether a larger study should be pursued. For this reason, serial N-of-1 trials analyzed with Bayesian metaanalysis is a methodology of considerable potential value in the pilot assessment of new therapeutic interventions and in the study of rare illnesses. By using these methods, very small numbers of study participants can provide important information regarding the likelihood of treatment benefit prior to embarking on the expense and difficulties of constructing large multicenter clinical trials. For example, similar methods have been used successfully and independently to investigate the use of an herbal supplement for the treatment of insomnia⁴⁰.

With regard to the second goal, our analysis indicated that there was a very low probability that amitriptyline acted to reduce pain in this population. Indeed, the mean treatment effect for pain was actually 0.67 cm, meaning that children experienced a greater reduction in pain during periods when they received placebo. The sensitivity analyses indicated that the conclusions of the study were robust — the findings were not sensitive to the choice of prior distribution. The pessimistic and skeptical priors did not change the conclusion that amitriptyline did not act to reduce pain. When a prior distribution that modeled a strong belief in the superiority of amitriptyline was used, the mean treatment effect was < 0 (−0.47 cm). However, there was only a 20% probability that the mean treatment effect was clinically significant. Thus, our pilot data suggest reconsideration of the benefits of amitriptyline even when the prior is optimistic. This work provides evidence that further study of amitriptyline in patients with juvenile arthritis similar to those who participated in this study is not strongly supported.

The protocol used for the conduct of this study required frequent contact with study participants. Subjects were seen in the clinic every 3 weeks during the study, which was difficult for those potential subjects who did not live close to the clinic. Future studies will need to balance the interval of time between assessments to minimize the burden on participants and still complete the study in a reasonable time. However, the potential burdens associated with an intensive protocol

must also be weighed against the burdens, costs, and difficulties associated with the conduct of larger multicenter clinical trials. Further, the advantage of being able to study a small number of subjects to obtain pilot data and perhaps avoid the costs of a larger study should not be underestimated. For these reasons, these methods are potentially an important tool in the study of rare illnesses and for pilot work to determine if further research is warranted.

The intentional enrollment of a very small sample size is a distinctive aspect of this study. We were particularly interested in learning if the study goals could be met, with results that were statistically robust, with a small sample size. Our success suggests that this design could be applied to other pilot work and the study of rare diseases. Although the small sample size is unusual, it is actually a strength of this work, and acts to confirm that surprisingly robust answers can be obtained from small data sets.

There are limitations that may affect the interpretation of our results. First, our results are applicable only to the narrow spectrum of juvenile arthritis studied here. Other groups of children with arthritis might experience improvement with amitriptyline therapy (e.g., those with higher pain scores or more active disease). The relatively low pain scores of participants in this study may have limited the potential for improvement, and obscured treatment effects. However, our results should be valid for children similar to those we have studied. Second, it is possible that amitriptyline used differently (higher doses or longer durations) could have had greater effects on pain relief. However, we chose both treatment parameters based on review of the literature and clinical experience. Third, all study participants received the same dose of amitriptyline. It is possible that for some children this dose was inadequate and for others the dose was more appropriate. Given that the “correct” dose was unknown and that there may have been other factors influencing the effects of amitriptyline (e.g., older children might have been more or less sensitive), the possible effects of this issue are difficult to predict. Finally, a washout period of 1 week was used. This was thought to be appropriate, but if significant carryover effects of amitriptyline periods carried over into placebo periods, this could have obscured benefits of amitriptyline. This is also unlikely, given that effects of amitriptyline typically disappear within a few days when used clinically.

Two participants in this study received corticosteroid joint injections during the study. As noted, both of these occurred during amitriptyline periods. It would be expected that this should have biased the results toward showing benefit for amitriptyline. When the analysis was repeated, deleting those treatment periods affected by the corticosteroid injections, the mean treatment effect for pain was more positive. This indicates an even smaller likelihood that amitriptyline was superior to placebo for pain relief in study participants. Thus, we believe that the use of corticosteroid injections during the trial does not influence our conclusions.

In summary, we have also demonstrated how serial N-of-1 trials could be combined using a Bayesian statistical model to generate estimates of the population effect, even with very small amounts of data. This methodology is a potentially powerful tool in pilot work and the study of rare illnesses, and may facilitate research under circumstances where more conventional methods are inappropriate or prohibitively expensive. However, the methods are time- and labor-intensive, for both researchers and participants, and should be chosen in circumstances when they are the most appropriate. We have also shown that there is little probability that amitriptyline acts as a pain-relieving medication in this population of children with arthritis. Our results do not encourage further research of this intervention in children similar to those studied here.

REFERENCES

- Peterson LS, Mason T, Nelson A, O'Fallon WM, Gabriel SE. Juvenile rheumatoid arthritis in Rochester, Minnesota 1960-1993: Is the epidemiology changing? *Arthritis Rheum* 1996;39:1385-90.
- Laaksonen A-L, Laine V. A comparative study of joint pain in adult and juvenile rheumatoid arthritis. *Ann Rheum Dis* 1961;20:386-7.
- Scott PJ, Ansell BM, Huskisson EC. Measurement of pain in juvenile chronic polyarthritis. *Ann Rheum Dis* 1977;36:186-7.
- Schanberg LE, Anthony KK, Gil KM, Maurin EC. Daily pain and symptoms in children with polyarticular arthritis. *Arthritis Rheum* 2003;48:1390-7.
- Lovell DJ, Walco GA. Pain associated with juvenile rheumatoid arthritis. *Pediatr Clin North Am* 1989;36:1015-27.
- Sherry DD, Bohnsack J, Salmonson K, Wallace CA, Mellins E. Painless juvenile rheumatoid arthritis. *J Pediatr* 1990;116:921-3.
- Varni JW, Thompson KL, Hanson V, Brik R. Chronic musculoskeletal pain and functional status in juvenile rheumatoid arthritis: an empirical model. *Pain* 1988;32:1-7.
- Huygen ACJ, Kuis W, Sinnema G. Psychological, behavioral and social adjustment in children and adolescents with juvenile chronic arthritis. *Ann Rheum Dis* 2000;59:276-82.
- Billings AG, Moos RH, Miller JJ, Gottlieb J. Psychosocial adaptation in juvenile rheumatic disease: a controlled evaluation. *Health Psychol* 1987;6:345-59.
- Doherty E, Yanni G, Conroy RM, Bresnihan B. A comparison of child and parent ratings of disability and pain in juvenile chronic arthritis. *J Rheumatol* 1993;21:1563-6.
- Malleson PN, Oen K, Cabral DA, Petty RE, Rosenberg AM, Cheang M. Predictors of pain in children with established juvenile rheumatoid arthritis. *Arthritis Rheum* 2004;51:222-7.
- Ilowite NT, Walco GA, Pochaczewsky R. Assessment of pain in patients with juvenile rheumatoid arthritis: relation between pain intensity and degree of joint inflammation. *Ann Rheum Dis* 1992;51:343-6.
- Hagglund JH, Schopp LM, Alberts KR, Cassidy JT, Frank RG. Predicting pain among children with juvenile rheumatoid arthritis. *Arthritis Care Res* 1995;8:36-42.
- Koh W, Pande I, Samuels A, Jones SD, Calin A. Low dose amitriptyline in ankylosing spondylitis: A short term, double blind, placebo controlled study. *J Rheumatol* 1997;24:2158-61.
- Rani PU, Naidu MUR, Prasad VBN, Rao TR, Shobha JC. An evaluation of antidepressants in rheumatic pain conditions. *Anesth Analg* 1996;83:371-5.
- Goldenberg DI, Mayskiy M, Mossey C, Ruthazer R, Schmid C. A randomized, double-blind crossover trial of fluoxetine and amitriptyline in the treatment of fibromyalgia. *Arthritis Rheum* 1996;39:1852-9.
- Carette S, Bell MJ, Reynolds WJ, et al. Comparison of amitriptyline, cyclobenzaprine and placebo in the treatment of fibromyalgia. *Arthritis Rheum* 1994;37:32-40.
- Godfrey RG. A guide to the understanding and use of tricyclic antidepressants in the overall management of fibromyalgia and other chronic pain syndromes. *Arch Intern Med* 1996;156:1047-52.
- Yaron I, Shirazi I, Judovich R, Levartovsky D, Caspi D, Yaron M. Fluoxetine and amitriptyline inhibit nitric oxide, prostaglandin E2, and hyaluronic acid production in human synovial cells and synovial tissue cultures. *Arthritis Rheum* 1999;42:2561-8.
- Barlow CF. Migraine in the infant and toddler. *J Child Neurol* 1994;9:92-4.
- Elser JM, Woody RC. Migraine headache in the infant and young child. *Headache* 1990;30:366-8.
- Hyams JS, Hyman PE. Recurrent abdominal pain and the biopsychosocial model of medical practice. *J Pediatrics* 1998;133:473-8.
- Romano TJ. Fibromyalgia in children: diagnosis and treatment. *WV Med J* 1991;87:112-4.
- Edgington ES. Statistics and single case analysis. *Prog Behav Modif* 1984;16:83-119.
- Kazdin AE. Single-case research designs: Methods for clinical and applied settings. New York: Oxford University Press; 1982.
- Parsonson BS, Baer DM. The analysis and presentation of single-subject data. In: Kratochwill TR, editor. *Single-subject research: Strategies for evaluating change*. New York: Academic Press; 1978:101-65.
- Barlow DH, Hersen M. *Single case experimental design: Strategies for studying behavioral change*. 2nd ed. New York: Pergamon Press; 1984.
- Bobrovitz CD, Ottenbacher KJ. Comparison of visual inspection and statistical analysis of single-subject data in rehabilitation research. *Am J Phys Med Rehabil* 1998;77:94-102.
- Nourbakhsh MR, Ottenbacher KJ. The statistical analysis of single-subject data: A comparative examination. *Phys Ther* 1994;74:768-76.
- Guyatt G, Sackett D, Adachi J, et al. A clinician's guide to conducting randomized trials in individual patients. *Can Med Assoc J* 1988;139:497-503.
- Jaeschke R, Adachi J, Guyatt G, Keller J, Wong B. Clinical usefulness of amitriptyline in fibromyalgia: The results of 23 N-of-1 randomized controlled trials. *J Rheumatol* 1991;18:447-51.
- Zucker DR, Schmid CH, McIntosh MW, D'Agostino RB, Selker HP, Lau J. Combining single patient (n-of-1) trials to estimate population treatment effects and to evaluate individual patient response to treatment. *J Clin Epidemiol* 1997;50:401-10.
- Jensen MB, McFarland CA. Increasing the reliability and validity of pain intensity measurement in chronic pain patients. *Pain* 1993;55:195-203.
- Bloom BJ, Alario AJ, Nobile C, Owens JA. The association of sleep disturbance with disease activity and parent-reported pain in children with juvenile rheumatoid arthritis [abstract]. *Arthritis Rheum* 1998;41 Suppl:S47.
- Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The Fatigue Severity Scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. *Arch Neurol* 1989;46:1121-3.
- Singh G, Athreya BH, Fries JF, Goldsmith DP. Measurement of health status in children with juvenile rheumatoid arthritis. *Arthritis Rheum* 1994;37:1761-9.
- Giannini EH, Ruperto N, Ravelli A, Lovell DJ, Felson DT, Martini A. A preliminary definition of improvement in juvenile arthritis. *Arthritis Rheum* 1997;40:1202-9.
- Dhanani S, Quenneville J, Perron M, Abdolell M, Feldman BM. Minimal difference in pain associated with change in quality of life

- in children with rheumatic disease. *Arthritis Rheum* 2002;47:501-5.
39. Dempster H, Porepa M, Young N, Feldman BM. The clinical meaning of functional outcome scores in children with juvenile arthritis. *Arthritis Rheum* 2001;44:1768-74.
40. Coxeter P, Schluter P, Eastwood H, Nikles C, Glasziou P. Valerian does not appear to reduce symptoms for patients with chronic insomnia in general practice using a series of randomized n-of-1 trials. *Complement Ther Med* 2003;11:215-22.

APPENDIX. The Bayesian hierarchical random-effects model.

Appendix 1. Explanation of Bayesian, hierarchical, random-effects model.

Pair 1				Pair 2				Pair 3			
Treatment		Placebo		Treatment		Placebo		Treatment		Placebo	
Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
$\alpha +$	$\alpha +$	$\alpha +$	$\alpha +$	$\alpha +$	$\alpha +$	$\alpha +$	$\alpha +$	$\alpha +$	$\alpha +$	$\alpha +$	$\alpha +$
Pair ₁	Pair ₁ +	Pair ₁	Pair ₁ +	Pair ₂ +	Pair ₂ +	Pair ₂	Pair ₂ +	Pair ₃ +	Pair ₃ +	Pair ₃	Pair ₃ +
+ trt ₁	trt ₁ +		post ₁	trt ₂	trt ₂ +		post ₂	trt ₃	trt ₃ +		post ₃
	post ₁ +				post ₂ +				post ₃ +		
	(trt*post) ₁				(trt*post) ₂				(trt*post) ₃		

The table shows the mean values for the 6 exchangeable VAS measurements for a single subject in the trial. This structure would be repeated independently for each of the subjects in the study. The 6 VAS scores in each study period (box) each come from a normal distribution with the corresponding mean and a common error variance; e.g., the VAS scores in the pre period for treatment in Pair 1 have mean $\alpha + \text{Pair}_1 + \text{trt}_1$

Definition of the parameters in the table and corresponding parameters in the model

There is an overall mean α for the subject:
 $\alpha = \text{subject Mean}$ [vector of length 6]

There are random effects $\text{Pair}_j, (j = 1 \text{ to } 3)$ that allow the means of the 3 pairs to be different:
 $\text{Pair}_j = \text{pair Deviation}$ [array with 6 rows, 3 columns]

There are effects of being post_{j, (j = 1 to 3)} that allow the pre and post means to differ. These are assumed the same for each pre-post couple within a pair, but different in the 3 pairs
 $\text{trt}_j = \text{treat.beta}$ [array with 6 rows, 3 columns]

There is the random effect for treatment trt that allows the overall mean in the treatment period to be higher than the mean in the placebo period:
 $\text{Post}_j = \text{prepost.beta}$ [array with 6 rows, 3 columns]

Finally, there is the effect of interest, the additional effect of treatment on the pre-post difference. This is modeled by the interaction between post and treatment and it is given a random effect that has a common mean for each subject. These are parameters here.

$$(\text{trt*post})_j = \text{treat.ixn.beta}$$
 [array with 6 rows, 3 columns]

What is the model doing? It would be possible to calculate the arithmetic mean of the pain scores for each treatment period. However, we are working here with parameters so there is an inherent weighting according to sample size and variability of the 6 VAS scores that contribute to the means in each cell of the table. Consider Pair 1:

The post-pre period mean for the treatment period is

$$\alpha + \text{Pair}_1 + \text{trt}_1 + \text{post}_1 + (\text{trt*post})_1 - [\alpha + \text{Pair}_1 + \text{trt}_1] = \text{post}_1 + (\text{trt*post})_1$$

The post-pre period mean for the placebo period is

$$\alpha + \text{Pair}_1 + \text{post}_1 - [\alpha + \text{Pair}_1] = \text{post}_1$$

The difference between these 2 differences is

$$\text{post}_1 + (\text{trt*post})_1 - \text{post}_1 = (\text{trt*post})_1$$