The Effects of Chronic Naltrexone Treatment in Young Autistic Children: A Double-blind Placebo-controlled Crossover Study

Sophie H. N. Willemsen-Swinkels, Jan K. Buitelaar, and Herman van Engeland

In a double-blind placebo-controlled crossover trial 23 autistic children, aged 3–7 years, were treated with a mean daily dosage of 1 mg/kg naltrexone for 4 weeks. Drug effects were monitored with behavior checklists rated by parents and teachers, and ethological playroom observations. On average, parents' checklists and playroom data could not differentiate between naltrexone treatment and placebo treatment; however, teachers significantly favored naltrexone treatment. They reported a decrease in hyperactivity and irritability. No effects of naltrexone on social and stereotypic behavior could be demonstrated.

Key Words: Autism, naltrexone, opiate antagonist, social behavior, locomotor activity, endorphins

Biol Psychiatry 1996;39:1023–1031

Introduction

Panksepp (Panksepp 1979; Panksepp and Sahley 1987) has suggested that phenomena that characterize autistic children, such as social withdrawal and stereotypies, may result from excessive activity of opioid systems in the brain. Evidence for this hypothesis has been sought in two lines of empirical studies: measurement of levels of endorphins and other neuropeptides in the body fluids of autistic subjects, and interventions in the opioid system in autistic subjects.

Research on plasma and cerebrospinal fluid endorphin levels in autistic subjects has yielded inconsistent results (Herman et al 1986; Ross et al 1987; Weizman et al 1988; Gillberg et al 1990; Sandman et al 1991; Leboyer et al 1994). These inconsistencies may relate to discrepancies in assay techniques, failures to account for coexisting abnormalities in other neurotransmitter or neuromodulator systems, and clinical and etiological heterogeneity within the autistic syndrome (Buitelaar 1993).

Intervention studies in the opioid system have focused on the effects of naltrexone, an opiate blocker. Some studies report on the effect of naltrexone on social behavior. Administration of a single dose of naltrexone was found to be effective in increasing verbal production and reducing social withdrawal (Campbell et al 1988, 1989; Leboyer et al 1988), but these studies used small numbers of subjects and open designs. Chronic treatment of autistic subjects with naltrexone yielded contradictory results. Several studies reported positive effects of naltrexone on social behavior, or more generally, the severity of autistic
symptoms, as measured by instruments like the Childhood Autism Rating Scale (CARS) (Borghese et al 1993), the Clinical Global Impressions (CGI) scale (Leboyer et al 1990, 1992; Kolmen et al 1995), or the Clinical Global Consensus Ratings (Campbell et al 1990). Campbell et al (1993), however, using a broad range of rating scales in a double-blind study with 23 children, ages 2–7 years, failed to find an effect on the core items of autism.

On the whole, direct observations of particular aspects of social and communicative behavior, such as the amount of eye contact, social proximity, or verbal utterances, failed to demonstrate beneficial effects of either acute or chronic naltrexone administration (Herman et al 1986, 1991; Borghese et al 1991, 1993). Kolmen et al (1995), in a recent controlled study on 13 children, claimed naltrexone to increase communicative initiations, but this change did not achieve a traditional level of statistical significance.

Instead of confining the study to specific aspects, Buitelaar et al (1990, 1992a and b) evaluated the behavior of autistic children in playroom sessions with broad, detailed, ethologically based observations. They were able to detect small, but significant, changes in social motivation or reciprocity after treatment with ORG 2766, a synthetic ACTH 4–9 analog.

The present study was designed to critically assess the efficacy of a chronic administration of naltrexone under double-blind and placebo-controlled conditions with a variety of instruments including detailed observations of the children’s behavior in a playroom.

Method

Subjects

Autistic children in the age range of 3–7 years and their parents were invited to participate in the study. All subjects were outpatients of the Department of Child Psychiatry of the Utrecht University Hospital or members of the Dutch Autism Association (NVA). Subjects could be included in the study if they met the DSM-III-R criteria of autistic disorder (American Psychiatric Association 1987). There was no further selection on the presence of special symptoms, such as insensitivity to pain or thermal stimuli, reduced crying, or special food preferences (Leboyer et al 1988, 1990; Panksepp et al 1994). The diagnosis was made by two child psychiatrists (JKB, HvE) after extensive diagnostic evaluations that included a review of prior records (developmental history, child psychiatric and psychological observations and tests), a parent interview, and a child psychiatric observation.

Twenty-three subjects entered the study, but three were excluded from data analysis for reasons to be explained later. Of the remaining 20 subjects, 16 were boys and four were girls, ages ranging from 2.8 to 7.4 years (age mean ± SD: 5.5 ± 1.2 years). All subjects selected displayed developmental abnormalities from birth onwards and before their first birthday, including a retardation of language development. The scores on CARS (Schopler et al 1980) ranged from 31.5 to 54.5 (38.8 ± 7.0). Eleven children were classified as severely autistic, and nine children as moderately autistic. Developmental quotients for nonverbal cognitive skills were assessed with a variety of tests [Wechsler Preschool and Primary Scale of Intelligence (WPPSI) in four children, Kaufman-ABC in four children, Stutsman in four children, Snijders-Oomen Nonverbal Intelligence Test in two children, and Psychoeducational Profile in six children]. Four children had a quotient below 40, two between 40 and 55, four between 55 and 70, and in ten subjects the nonverbal intelligence quotient was above 70. Eight children were mute. For the remaining children verbal quotients ranged from 64 to 104. Based on the Wing classification for social behavior (Wing and Gould 1979), six children could be considered “aloof,” three “passive,” and 11 “active but odd.” Five children frequently showed an excess of stereotypic motor behavior. Only two children showed mild self-injurious behavior (one child hand-biting, the other child head-banging). Of five children, there were complaints of severe hyperactive and chaotic behavior.

All subjects underwent a complete medical workup, including chromosomal testing, EEG, metabolic screening, and physical and neurological examination. None of our subjects had chromosomal abnormalities, or known neurological or metabolic diseases that may be considered a biomedical cause of autism; two subjects however had a concomitant abnormality (seizure disorder in one subject and situs inversus in another child). The boy with epilepsy was treated with carbamazepine in a fixed dosage during the trial. None of the subjects was previously treated with psychotrophic drugs.

The study protocol was approved by the Committee for Research on Human Subjects of the Utrecht University Hospital. All parents gave their informed consent prior to the inclusion of their children in the study.

Design

A double-blind and placebo-controlled crossover design was employed. All children but one weighed between 17 and 27 kg. They all received 20 mg naltrexone per day. One boy with a weight of 42 kilograms received an adjusted dosage of 40 mg naltrexone per day. This resulted in dosages ranging from 0.74–1.18 mg/kg/day (mean ± SD 0.98 ± 0.13 mg/kg). Naltrexone was administered in a capsule and a matched capsule was used as placebo. Both
were prepared by the local pharmacy. Medication was administered in the morning, before or during breakfast. The dosage of approximately 1 mg/kg naltrexone per day was chosen in accordance with two double-blind placebo-controlled studies reporting beneficial socializing effects (Walters et al 1990; Campbell et al 1990). The children were randomly assigned to their treatment sequence: naltrexone-placebo and placebo-naltrexone. Each treatment block consisted of a 2-week baseline period, a 4-week period of active treatment (naltrexone or placebo), and a 4-week wash-out period. Prior to baseline measurements, subjects and parents were adapted to all procedures, including administration of capsules and the playroom session.

Assessments

SIDE EFFECTS. Treatment-emergent adverse effects and changes in appetite were assessed systematically after 2 weeks and 4 weeks of treatment by questioning of parents and teachers. Children were weighed before, during, and after each treatment period.

BEHAVIOR CHECKLIST RATINGS. Behavior checklists were rated twice before each treatment (2 baselines), halfway and at the end of each treatment (2 weeks and 4 weeks treatment), and 4 weeks after the end of each treatment (wash-out). The parents completed the Aberrant Behavior Checklist (ABC) (Aman et al 1985). In addition, they completed a list of 12 social items (SOC) that were derived from the Matson Evaluation of Social Skills with Youngsters (MESSY) (Matson et al 1985). These items were added because we felt that the ABC items that refer to social behavior did not cover the entire area. Examples of added items are: “Does not look at people when they are speaking.” “Is not friendly to new people he/she meets.” Teachers were asked to rate a condensed version of these questionnaires. Furthermore, the parents and teachers completed a checklist (TAR) that consisted of five target symptoms that were chosen in consultation with the parents and teachers and were specific for each child. The parents were asked to describe specific stereotypic, compulsive, and socially inadequate behaviors of their child, for instance “tapping at windowpanes” and “asks the same question over and over.” These targets had to be rated on a five-point scale. The ABC, SOC, and TAR checklists were rated in the weeks 0 and 2 (baseline ratings) and the weeks 5 and 7 (after 2 weeks and 4 weeks of treatment) of a treatment period. Finally, parents and teachers completed the CGI improvement scale (National Institute of Mental Health 1985) at the end of a 4-week treatment period. The CGI is a seven-point scale ranging from “very much improved” to “very much worse.”

PLAYROOM SESSION. Behavior elements of children and parents were recorded in a 40-min semistructured playroom session that was conducted 1 week before and at the end of each treatment. The playroom session consisted of three parts. In the first part (7 min), the parent was provided with a book and one toy and was asked to await the initiatives of the child. The second part (17 min) included a period of watching television, and performing a playful task. In the third part of the session (16 min), the child was encouraged to play freely. All sessions were videotaped. Behavior elements were recorded by means of “The Observer” (Noldus 1991). Monitored behavior elements included elements describing the activities of the child as well as elements describing the interactions between child and parent.

Thanks to the use of a Video Tape Analysis System, making registration at very slow rates possible, the reliability was good. The percentage agreement between two raters ranged from 85% for the communication elements up to 90% for the elements describing gaze behavior.

During the playroom sessions activity measurements were made by means of an actometer. The actometer was worn on the wrist of the nondominant arm. The characteristics of the actometer (Gachwiler Electronic, CH-8634 Hombrechtikon) used in this study have been described previously (Borbely 1986).

Data Analysis

In order to reduce the number of playroom parameters, several elements were combined when they appeared to belong to the same behavioral system. In order to combine behavior elements, Pearson correlation coefficients were calculated between the change scores of duration or frequency of elements. Those elements, of which changes in duration or frequency were highly correlated, were summed. See the Appendix for a complete and detailed description of the derived variables.

The data were analyzed with repeated measures analysis of variance with a between-subject factor “sequence” (two levels) and within-subject factors “period” (two levels) and “treatment” (two levels) (SPSS-PC 4.0). The factor “treatment” represented the difference between the baseline and endpoint values of a treatment period. In this model the interaction effect “sequence” by “period” by “treatment” corresponds with a naltrexone effect. All testing was two-tailed with $\alpha = 0.05$. To test for the assumptions needed for applying analysis of variance, the Bartlett-Box-$F$ was used to test for homogeneity of variance. The BOX-$M$ was used as a multivariate test for equality of the variance–covariance matrices. If necessary, the data were subjected to a logarithmic transformation to stabilize variances.
Results

Compliance and Side Effects

Three children dropped out of the trial in their second treatment period. The parents of one boy considered him to be a significant clinical responder in the first treatment period. They were so convinced that he had been treated with naltrexone, that they refused to let him participate in the following "placebo" period. It turned out that he had indeed been treated with naltrexone. Two other children refused to take their naltrexone capsules in the second period, complaining of an awful and very bitter taste. As a result 20 children were included in the analysis.

No serious untoward effects were reported by the parents or the teachers. Appetite and weight did not appear to be affected.

Behavior Checklist Ratings

The behavior ratings completed by the parents did not differentiate between naltrexone treatment and placebo treatment. In contrast, naltrexone treatment brought about significant changes in the behavior ratings completed by the teachers. The teacher ABC sum score was significantly decreased after naltrexone treatment \([F(1,18) = 5.62, p = .029]\). Examination of changes in the ABC subscales (Marshburn and Aman 1992) showed that this was primarily due to a significant reduction of the scores on the hyperactivity and irritability subscales \([\text{hyperactivity: } F(1,18) = 12.11, p = .003; \text{ irritability: } F(1,18) = 4.50, p = .048]\).

The target sum score as rated by the teachers was also significantly reduced after naltrexone treatment \([F(1,18) = 5.68, p = .028]\). Inspection of changes of single target behaviors revealed that those targets which were related to hyperactivity showed the greatest decrease after naltrexone treatment.

Figure 1 plots the Clinical Global Impression ratings of the parents after 4 weeks of treatment with naltrexone and with placebo. Subjects were sorted by preference of the parent. The placebo treatment was preferred to the naltrexone treatment in four subjects, whereas in nine subjects the naltrexone treatment was preferred to the placebo treatment. No significant difference between naltrexone and placebo treatment was found on the CGI ratings by the parents (paired t test, \(t = -.90, \text{df} = 19, \text{n.s.}\)).

The CGI ratings of the teachers are displayed in Figure 2. Placebo treatment was preferred to naltrexone treatment in only two subjects, whereas in 13 subjects naltrexone treatment was preferred to treatment with placebo (paired t test, \(t = -3.93, \text{df} = 19, p = .001\)).

Playroom Session

No significant differences between placebo and naltrexone treatment could be found for either the child's interactions with the parent nor the child's activities during the playroom session. Further analysis of sequential patterns
of behavior elements failed to demonstrate beneficial effects of naltrexone on social reciprocity. All analyses were repeated separately for the structured and the unstructured parts of the session. This too revealed no significant naltrexone effects.

Actometer data were complete for only 16 subjects, due to a technical problem in the actometer. Actometer scores averaged over the whole session did not differentiate between naltrexone and placebo treatment.

**Individual Responders**

Individual responders to naltrexone treatment were identified by means of the reliable change (RC) method described by Jacobson and Truax (Jacobson and Truax 1991). The RC method examines the extent to which treatment moved someone outside the range of the dysfunctional population or within the range of the functional population. The clear-cut criterion for improvement is that the difference between the pre- and posttreatment scores, divided by the standard error of the difference score, is greater than 1.96. Assuming a normal distribution, this reflects the chance that the change due to measurement error is lower than .05. For each subject the RC was calculated based on the ABC-sum score as rated by the parents. The correlation coefficient between the first and the second baseline was used as test-retest reliability index of the measure. In seven children, the RC was larger than the p = .05 point (subjects c, d, f, g, j, p, and q of Figures 1 and 2). These seven children were considered to be individual drug responders and have entered an open-label continuation treatment with naltrexone.

The seven responders did not differ from the nonresponding children in age or IQ. There were no differences between responders and nonresponders in behavioral characteristics, except for social behavior. The responders were characterized by significantly higher (i.e., more problematic) baseline scores on the lethargy/social withdrawal subscale of the ABC (see Figure 3; parents: t test; t = 3.58, df = 18, p = .002; teachers: t = -4.52, df = 18, p = .000) and SOC sum score (parents: t test, t = -3.41, df = 18, p = .003; teachers: t = -3.13, df = 18, p = .006).

When the changes on the behavior checklists were examined for the subgroup of responders only, however, not social withdrawal but hyperactivity proved to be consistently reduced by the naltrexone treatment. This is illustrated in Figure 4.

**Discussion**

**Absence of Effects on Social Behavior**

On average, parents could not distinguish between a chronic naltrexone treatment and a chronic placebo treatment. Both treatments appeared to improve the behavior of
the children compared to baseline on rating scales. Accordingly, detailed observations of the child's activities and interactions with the parent during a playroom session as well as actometer scores failed to distinguish between placebo and naltrexone treatment. In contrast, teachers significantly differentiated between naltrexone treatment and placebo treatment. The naltrexone period was preferred significantly more than the placebo period and rating scale scores dropped significantly after naltrexone treatment, whereas placebo values did not differ from baseline values.

Inasmuch as naltrexone was found to benefit the autistic children at school, the changes reported concerned a reduction of hyperactivity and irritability rather than an improvement of their social deficits. Notably our detailed behavior observations in the playroom session, which have
been shown to be sensitive to subtle changes in social behavior in previous studies (Buitelaar et al 1990, 1992a and b), could not substantiate socializing effects of naltrexone. The failure of naltrexone to ameliorate the social behavior of autistic children evidently is at variance with the opioid excess hypothesis of autism. All controlled naltrexone trials in larger samples of autistic children, including the present study, failed to find significant effects on social behavior (Campbell et al 1993; Kolmen et al 1995).

Effect on Hyperactivity
The effects of naltrexone on hyperactivity seem to be a remarkably consistent finding (Campbell et al 1988, 1989, 1990, 1993, Leboyer et al 1988; Asleson et al 1991; Herman et al 1991, 1993; Kolmen et al 1995; Willemsen-Swinkels et al in submission). The opioid system has been found important in mediating activity and exploratory behavior in animal and human studies (Katz 1988) and several explanations may be considered for this effect of naltrexone. First of all, the explanation might be simply that naltrexone exerted a generalized depressant effect upon activity. This however is unlikely, since the scores on the social withdrawal/lethargy subscale of the ABC completed by parents and teachers manifested even a tendency to reduction (i.e., improvement) following naltrexone. Furthermore, the playroom data did not reveal nonspecific sedative effects on behavior.

Second, there is some evidence that the opiate system is involved in a person’s ability to allocate attention to relevant sources (Strahlendorf et al 1980; Berridge et al 1993). Finally, it is possible that the naltrexone effects on activity are secondary to a blockade of opioids released by stress. Support for this possibility may be found in the situation-dependency of the naltrexone effect. In animal studies naloxone, a short-acting opiate antagonist, has been found to reduce locomotor activity of rats in a social context, whereas locomotor activity of rats tested individually proved not to be affected (Dokla 1992). School, as a social context and presumably high arousal condition, may serve to promote endogenous opioid release; subsequent opioid blockade may then contribute to reduced activity.

Individual Response
Based on changes of the ABC checklist at home, seven children were identified as individual responders to naltrexone treatment. Further examination of the ABC subscales completed by the parents revealed that the responders had improved in hyperactivity and irritability, but not in social behavior. Analysis of the characteristics of these seven children suggested that the hyperactivity/irritability decreasing effect reported by the parents was related to baseline severity of social problems, not to the severity of hyperactivity/irritability. This questions the clinical usefulness of naltrexone for treating hyperactive behavior in children who are less severely disturbed in their social behavior.

In conclusion, the present data indicate that a 1.0 mg/kg dose of naltrexone brought about a modest and situation-dependent reduction in hyperactivity and irritability, rather than improvements in social behavior. The hypothesis of opioid involvement in the core symptoms of autism is therefore not supported. The hyperactivity-reducing effect of naltrexone may have clinical relevance in some autistic children, if the effects are shown to be maintained or even enlarged during continuation of treatment. So far, however, naltrexone cannot be recommended for clinical use in autism.

This research is supported by the Janusz Korczak Foundation. We wish to thank Florence Weijnen, Paul Stemerding, and Ineke den Hartog for their assistance in data collection. Du Pont Pharma (Dr. Klein) was generous in supplying part of the drug required.

Appendix
The originally 44 behavior elements of the playroom observations were transformed into the following variables:

SOCIAL REACTIVITY: percentage of initiations of the parent that are followed by a reaction of the child.
LOOK FACE: percentage of time spent looking at the face of the other during the task part of the session.
VERBAL INITIATIVE: number of verbal remarks, questions, and requests.
NONVERBAL INITIATIVE: number of communicative gestures and touches.
VOCALIZATION: number of vocalizations including situation-irrelevant echolalia.
BEHAVIOR SHIFT: Total number of times that the child starts one of the following behaviors: self-injurious behavior, stereotypic behavior, distress, (auto)manipulation, change toy, functional play, walk.
STEREOTYPY: percentage of time that the child shows self-injurious behavior, repetitive movements, or licks or sniffs objects.
FUNCTIONAL ACTIVITY: percentage of time that the child is involved in functional play or automanipulation. (The behavior elements functional play and
automanipulation had negative Pearson correlation coefficients with the other activity elements and positive coefficients with each other.)

**ACTIVE OTHER**: percentage of time that the child manipulates an object, changes toys, walks, or shows clear signs of distress (screaming, yelling, crying).

**LOOK TASK**: percentage of time spent in looking at the given task during the task part of the session.

### References


