TEMPLATE RESEARCH PROTOCOL

(September 2018)

- May 2015: adaptation section 11.5: text in accordance to old and new Measure regarding Compulsory Insurance for Clinical Research in Humans
- Sept 2015: adaptation section 9.1, 9.2 and 12.5: text in accordance to WMO amendment on reporting SAE and temporary halt (section 10 of WMO)
- Oct 2015: adaptation section 4.4 – comment [CCMO15], 8.2 and 10.1 with respect to methodology/statistics
- Sept 2018: adaptation section 12.1 and comment [CCMO46] due to applicability GDPR as of May, 2018

PROTOCOL TITLE Pegasus: Equine-assisted Therapy for therapy-resistant adolescents with autism spectrum disorders, a multiple baseline ABA-study

Protocol ID	185 - 20 PEGASUS
Short title	Pegasus
EudraCT number	*n/a*
Version	1.0
Date	24-06-2021
Coordinating investigator/project leader	**Drs. J.C. den Boer,** Kinderpsychiater *T:* 0654310670 *E:* j.denboer@karakter.com Karakter Kinder- en Jeugdpsychiatrie Postbus 68 6710 BB EDE GLD
Principal investigator(s) / multiple sites	**Prof. Dr. N. Lambregts-Rommelse,** Klinisch neuropsycholoog/onderzoeker *T:* 024-3610655 *E:* Nanda.Lambregts-Rommelse@radboudumc.nl Karakter, Kinder- en Jeugdpsychiatrie Postbus 68 6710 BB EDE GLD **Dr. A. Blonk** *T:* 0620538996 *E:* ablonk@gagelhoeve.nl Gagelhoeve Gagelweg 10 5451 NE MILL **E. van Noort** *T:* 0681127700 *E:* esther@horsesenco.nl Horses & Co Lindeweg 12 2995 XK HEERJANSDAM
Sponsor (in Dutch: verrichter/opdrachtgever)	Karakter Kinder- en Jeugdpsychiatrie Postbus 68 6710 BB EDE GLD
Subsidising party	ZonMw 60-63600-98-914
Independent expert (s)	Dr. W. Groen, w.groen@karakter.com

PROTOCOL SIGNATURE SHEET

Name	Signature	Date
Head of Department: Prof. dr. W. Staal Hoogleraar Kinder- en Jeugdpsychiatrie *T:* 024-3512222 *E:* w.staal@karakter.com Karakter Nijmegen Universitaire Kinder- en Jeugdpsychiatrie Reinier Postlaan 12 6525 GC NIJMEGEN		**24-06-2021**
Coordinating investigator/project leader Drs. J.C. den Boer, Kinderpsychiater		
Principal Investigator Prof. Dr. N. Lambregts-Rommelse, Klinisch neuropsycholoog/onderzoeker		

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND RELEVANT DEFINITIONS

ABR	**General Assessment and Registration form (ABR form), the application form that is required for submission to the accredited Ethics Committee; in Dutch: Algemeen Beoordelings- en Registratieformulier (ABR-formulier)**
ASD	**Autism Spectrum Disorder**
CCMO	**Central Committee on Research Involving Human Subjects; in Dutch: Centrale Commissie Mensgebonden Onderzoek**
CV	**Curriculum Vitae**
EAT	**Equine-assisted Therapy**
GCP	**Good Clinical Practice**
GDPR	**General Data Protection Regulation; in Dutch: Algemene Verordening Gegevensbescherming (AVG)**
IC	**Informed Consent**
METC	**Medical research ethics committee (MREC); in Dutch: medisch-ethische toetsingscommissie (METC)**
(S)AE	**(Serious) Adverse Event**
Sponsor	**The sponsor is the party that commissions the organisation or performance of the research, for example a pharmaceutical company, academic hospital, scientific organisation or investigator. A party that provides funding for a study but does not commission it is not regarded as the sponsor, but referred to as a subsidising party.**
SUSAR	**Suspected Unexpected Serious Adverse Reaction**
UAVG	**Dutch Act on Implementation of the General Data Protection Regulation; in Dutch: Uitvoeringswet AVG**
WMO	**Medical Research Involving Human Subjects Act; in Dutch: Wet Medisch-wetenschappelijk Onderzoek met Mensen**

SUMMARY

Rationale: For people with autism spectrum disorder (ASD), daily life is highly stressful and traumatic with many unpredictable events that can evoke emotion dysregulation (ED): a strong difficulty with appropriately regulating negative affect. Heightened levels of ED may aggravate social dysfunctioning in ASD and vice versa. For a part of the patients with ASD, treatment as usual does not have any effect at all on ED. As ASD with severe levels of ED can be considered to have an ultra-high risk profile for developing other disorders (psychosis, anxiety, eating disorders, depression), this treatment-resistant subgroup of patients may end up needing life-long psychiatric treatment. Particularly problematic is that these patients often lack motivation for typically initiated forms of therapy, thereby further limiting their chances for a more favourable outcome. A highly promising method that may prove effective for therapy-resistant individuals with ASD is Equine-Assisted Therapy (EAT). While often met with prejudgment and skepticism, reports from parents and therapists as well as a recent systematic review suggest that EAT may have beneficial effects in youths with ASD. We further argue that an ideal (and perhaps last?) 'window of opportunity' for intervention in treatment-resistant patients with ASD is adolescence, because of the major genetically pre-programmed neurological changes occurring in this period that heighten the sensitivity for environmental input. EAT targeting severe ED offered within this timeframe may improve clinical outcomes both in the short and in the long term in otherwise treatment-resistant adolescents with ASD.

Objective: To quantify the short-term (15 weeks) and long-term (1 year) (cost-)effectiveness of Equine-Assisted Therapy (EAT) in adolescents with therapy-resistant ASD (aged 11-18) and, when proven (cost-) effective, implement EAT in clinical practice.

Hypothesis: In a therapy-resistant group (N=35) of youths with ASD, a 15-week treatment with EAT significantly improves:

1. the primary outcome measure emotion dysregulation in the short term
2. secondary outcome measures of emotion dysregulation in the long term (1 year), ASD symptom severity, quality of life, self-esteem, global and family functioning and goal attainment as well as qualitative outcomes in the short term (15 weeks) and the long term (1 year).
3. is more cost-effective on the long term (1 year) when compared to continued CAU.

Study design: Mixed-methods strategy consisting of three elements: a randomized, multiple-baseline single-case design (n=35), a qualitative study (n=8-10) and a cost-effectiveness study (n=6).

Study population: Treatment-resistant adolescents with ASD (11-18 years), with clinical levels of emotion dysregulation despite receiving care as usual (CAU) for at least 18 months

Intervention: 15 weekly sessions of 60 minutes EAT using a standardized protocol by certified EAT therapists.

Main study parameters/endpoints: the main study parameter is the change on the Emotion Dysregulation Index

Nature and extent of the burden and risks associated with participation, benefit and group relatedness: Participants receive regular therapy and participate in interviews/ questionnaires. The risks are negligible and the extra burden is minimal.

1. INTRODUCTION AND RATIONALE

Autism spectrum disorder (ASD) is a severe neurodevelopmental disorder with a steadily rising world-wide prevalence of about 1-3% (Fombonne, 2018). Autism is marked by two phenomena. Firstly, patients with autism experience difficulties in the integration of information. This may result in difficulties in understanding others (theory of mind), a more detail-focused style of processing information (executive functioning), or difficulties in imagining future events. Secondly, patients with autism experience difficulties in processing sensory information, which may result in either overwhelming or decreased experiences of sensory input. As a consequence, daily life is highly stressful and traumatic with many unpredictable events that evoke emotion dysregulation (ED), a strong difficulty with regulating negative affect (Aldao, Nolen-Hoeksema, & Schweizer, 2010), resulting in low self-esteem. Individuals with ASD rely on more maladaptive emotion regulation strategies (Samson, Wells, Phillips, Hardan, & Gross, 2015), which is associated with a wide range of negative outcomes, such as poorer social functioning (Nader-Grosbois & Mazzone, 2014) and more depression and anxiety symptoms(Mazefsky, Borue, Day, & Minshew, 2014; Swain, Scarpa, White, & Laugeson, 2015). Moreover, heightened levels of negative emotions may influence social functioning in ASD and vice versa. For example, a reduced ability for cognitive-reappraisal (i.e. the ability to change cognitions about a situation) is associated with higher levels of negative emotion, which in turn is related to more maladaptive behavior in ASD (Mazefsky, 2015). Vice versa, irritability may promote misinterpretation of social cues or inappropriate behavior in social interactions (Mazefsky, 2015). As such, ASD with severe levels of ED can be considered to have an ultra-high risk profile for developing other disorders, such as psychosis, anxiety, eating disorders, and depression. Treatment targeting this maladaptive cycle may improve outcomes for individuals with ASD.

'Therapy-resistant' ASD

In a part of the patients suffering from ASD, care as usual (psycho-education, medication, cognitive therapy, environmental adaptations) does not diminish ED at all. There are currently no guidelines or treatment options for this subgroup of severely impaired, therapy-resistant individuals with ASD. The ‘dooming’ perspective of a lifelong disease may significantly damage patients and parents by creating a self-fulfilling prophecy and a cascade towards comorbid disorders, such as psychosis, depression, trauma, personality disorders, eating disorders and school dropout(Gillespie-Lynch, Kapp, Brooks, Pickens, & Schwartzman, 2017; Mogensen & Mason, 2015).

A promising novel treatment method targeting emotion dysregulation: equine-assisted therapy (eat)

A highly promising method that may prove effective for therapy-resistant individuals with ASD is Equine-Assisted Therapy (EAT) (McDaniel Peters & Wood, 2017). While often met with prejudgment and skepticism, reports from parents and therapists suggest that EAT may have beneficial effects in youths with ASD. Horses have natural benefits over human therapists: they are experts in nonverbal communication, do not judge, offer unconditional acceptance, provide tactile comfort and give multiple real opportunities to gain experience with new behavior. EAT involves specific rehabilitative goals, individual tailored objectives and is directed by a healthcare professional (Fuller & Kaiser, 2020; Reichow, 2012; Trzmiel, Purandare, Michalak, Zasadzka, & Pawlaczyk, 2019). Patients can project and express their emotions and use the horse as a metaphor or a model. Just recently, a systematic review was published demonstrating highly potential effectiveness of EAT, especially with regard to both social functioning and emotion dysregulation (Trzmiel et al., 2019). However, this review also pointed towards the need for further high-quality research, as methodological shortcomings hampered prior studies (please see literature review).

Adolescence: a (perhaps last?) 'Window of opportunity' to intervene

In ASD literature, early childhood is considered to be the first 'window of opportunity' to intervene }(Fuller & Kaiser, 2020; Reichow, 2012). During this period, significant treatment-related improvements in the core features of ASD have been reported (Peters-Scheffer, Didden, Korzilius, & Sturmey, 2011). We argue that a second (and perhaps last?) 'window of opportunity' for intervention is adolescence for two reasons: Firstly, like early childhood, adolescent age is related to major genetically preprogrammed neurological changes with a high sensitivity for environmental input across species (Malter Cohen, Tottenham, & Casey, 2013; Mengler et al., 2014). In humans, myelination and synaptic pruning in the prefrontal cortex increases sharply, and neural connections between the prefrontal cortex and other regions of the brain are strengthened; this improves the efficiency of information processing, controlling impulses, planning ahead, enhanced decision-making, processing of emotional experiences and social information, and helps to more accurately evaluate rewards and risks (Burnett Heyes et al., 2015). Secondly, recently published longitudinal studies show that, particularly during adolescence, significant ASD symptom change may occur (Fein et al., 2013; Sussman et al., 2015). Offering intervention during this time window may therefore be the last opportunity to improve adulthood outcomes.

Aim of this research

For people with ASD, daily life is highly stressful and traumatic with many unpredictable events that evoke emotion dysregulation: a strong difficulty with appropriately regulating negative affect. Heightened levels of negative emotions may influence social dysfunctioning in ASD and vice versa. Treatment targeting this maladaptive cycle may improve outcomes for individuals with ASD. A part of the patients does not respond to care as usual. There currently are no guidelines or treatment options for this subgroup of severely impaired, therapy-resistant individuals with ASD. The 'dooming' perspective of a lifelong disease may significantly damage parents and patients by creating a self-fulfilling prophecy and a chronic cascade towards lifelong psychiatric disorders. Particularly problematic is that these patients often lack motivation for typically initiated forms of therapy, thereby further limiting their chances for a more favorable outcome. A highly promising method that may prove effective for therapy-resistant individuals with ASD is Equine-Assisted Therapy (EAT). While often met by prejudgment and skepticism, reports from parents and therapists as well as a recent systematic review (Trzmiel et al., 2019) suggest that EAT may have beneficial effects in youths with ASD. In addition to early childhood, a second (and perhaps last?) 'window of opportunity' for intervention is adolescence because of the major genetically pre-programmed neurological changes occurring in this period, which heighten the sensitivity for environmental input.

2. OBJECTIVES

Primary Objective:
The primary objective is to significantly improve emotion dysregulation in a therapy-resistant group of youths with ASD (aged 11-18) after a 15-week treatment with Equine-Assisted Therapy (EAT).

Secondary Objective(s):
1. is to significantly improve emotion dysregulation in a therapy-resistant group of youths with ASD (aged 11-18) at the long term (1 year) after treatment with Equine-Assisted Therapy (EAT).

2. is to significantly improve outcome measures of ASD, symptom severity, quality of life, self-esteem, global and family functioning and goal attainment as well as qualitative outcomes in (a) the short term (15 weeks) and (b) the long term (1 year).

3. is to quantify the cost of EAT therapy when compared to a scenario including continued CAU.

3. STUDY DESIGN

In order to be able to assess the effectiveness of the intervention, we designed a mixed-methods strategy consisting of three elements: a randomized, multiple-baseline single-case design, a qualitative study and a cost-effectiveness study (see figure 1).

The randomized, multiple-baseline single-case design

Since heterogeneity is a major aspect of the population that could benefit from EAT, classic RCT models assuming homogeneity are less useful. In addition, an RCT comparing EAT with CAU would not be feasible if the treatment of choice (EAT) is withheld for a substantial amount of time from the group randomized to CAU. This group would either withdraw from participation immediately and/or seek EAT elsewhere, thereby directly undermining the validity of an RCT.

Therefore, in line with the recommendations of ZonMw (Eindrapport 'Alternatieven voor RCT bij de evaluatie van effectiviteit van interventies!'), a randomized, multiple-baseline single-case design will be carried out in 35 adolescents (11-18 years) with therapy-resistant ASD in n=4 therapy centers. Participants will complete repeated measurements during a baseline phase (phase A, 2-6 weeks), an intervention period (phase B, 15 weeks) and a post-intervention period (phase C, 2-6 weeks). Phase A acts a control and will be compared with phase B. By applying multiple baselines of varying length, observed effects of the treatment can be distinguished from effects due to chance, thus increasing internal validity. The total duration of phases A and C is 8 weeks for each participant and, consequently, participants with a longer phase A will have a shorter phase C. Participants will be randomly assigned to a baseline duration by an independent researcher not involved in EAT nor the patient. During the total study period of 23 weeks, parents of the participants will complete an emotion regulation questionnaire (Emotion Dysregulation Inventory) three times a week to answer the primary outcome. Other assessments (ASD symptom severity, quality of life, self-esteem, global and family functioning and goal attainment) will take place at the baseline (T0), at the end of phase A (T1), after completion of phase B (T2), after the end of phase C (T3) and after one year (T4) (please see section outcomes).

The qualitative study

In the qualitative study, the aim is to obtain information from a varied group of adolescents, caregivers and professionals not only to gain insight into their individual experiences, but also into the effects of the intervention. Purposive sampling will identify those participants

who are most likely to provide rich information about their experiences. The study population consists of three groups:

1. Adolescents referred to the intervention (n = 8-10)
2. Caregivers of the adolescents referred to the intervention (n = 8-10)
3. The professionals delivering the intervention (n = 4)

We will use a maximum variation sampling approach of purposefully selecting a wide range of participants (Patton, 2015). The sample size is particularly dependent on the complexity of the research and the heterogeneity of the relevant characteristics. However, it is expected that 8-10 participants will be included in each group (adolescents and caregivers). The final number of subjects, however, will be determined on the basis of saturation. We will interview four members of the professionals group.

The cost-effectiveness study

We will use the 'Outcome generator' (https://effectencalculator.nl/) to answer questions on cost-evaluation and the effect of improved well-being. This instrument helps to make statements on innovative interventions based on dialogue. Using the 'Outcome generator' we will be able to analyze the social benefits of EAT and the costs saved by society due to positive results of EAT. Each year we will randomly select two cases from the participants to the EAT intervention. The evaluation sessions will take place approximately four months after the intervention is finished.

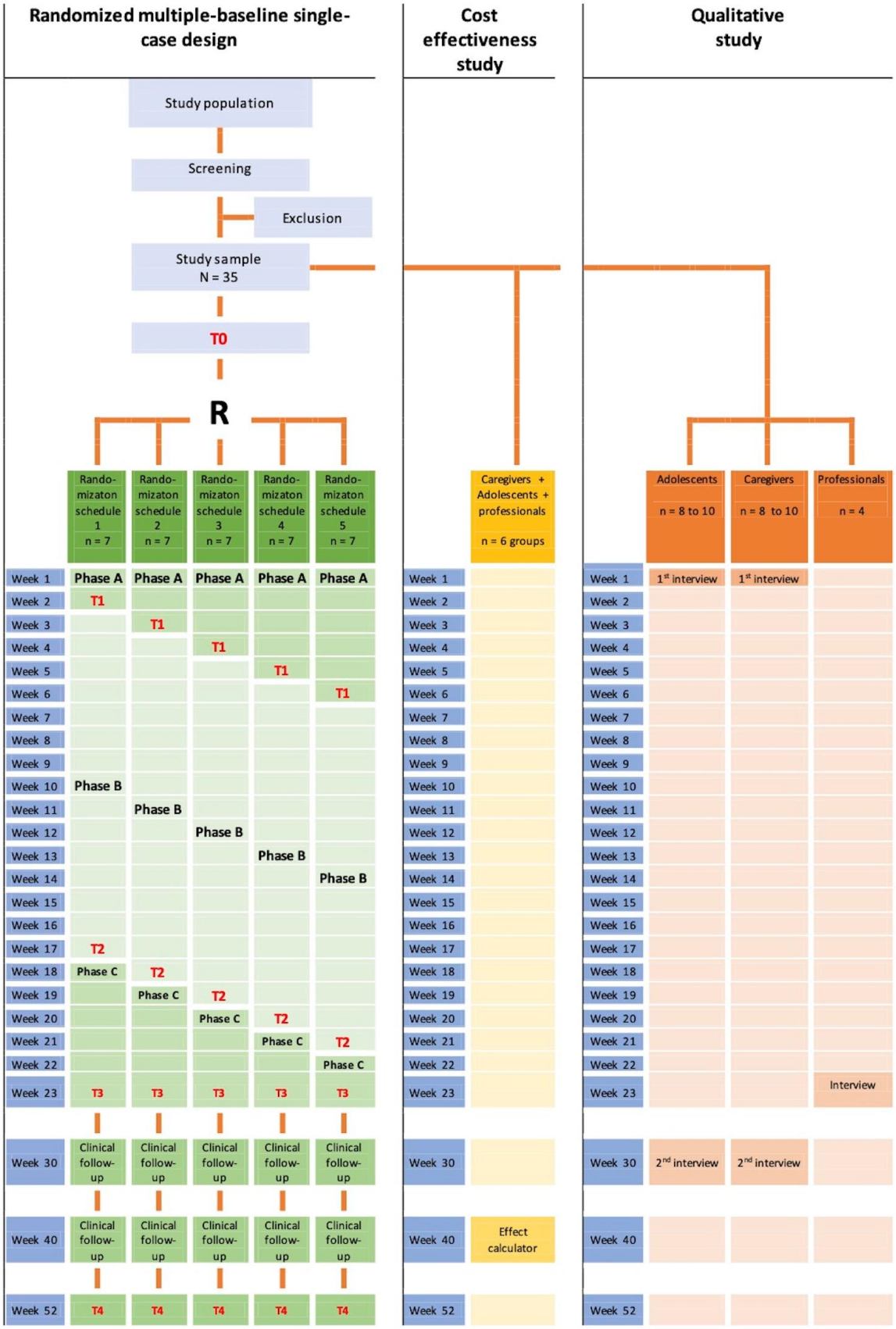

Figure 1. Design Pegasus study

4. STUDY POPULATION

4.1 Population (base)

Participants are adolescents between 11-18 years, with a diagnosis of autism spectrum disorder. Recruiting n=35 participants within four years is highly feasible, since recruitment is done both from within Karakter (where n=3200 patients with ASD are yearly treated) as from other sources (EAT centers, social media, patient organizations).

4.2 Inclusion criteria

In order to be eligible to participate in this study, a subject must meet all of the following criteria:

- between 11-18 years old;
- a clinical diagnosis of autism spectrum disorders according the DSM 5 as diagnosed by a BIG registered healthcare professional;
- insufficient emotion regulation after regular therapy for at least 1,5 years as indicated by a score above clinical cut-off (T-score = 65) on the EDI
- (T-score = 65) on the EDI;
- comorbidities are allowed

4.3 Exclusion criteria

A potential subject who meets any of the following criteria will be excluded from participation in this study:

- unable to respond to questions (parents or adolescents);
- no access to an Internet connection;
- insufficient mastery of Dutch language in parents or adolescents;
- physically incapable to work with the horses;
- Significant change in medication use;
- total IQ equal to or below 80 on the WISC-III-R, or WISC-V or WAIS-IV;
- allergic or phobic to horses;
- insufficient regulation to safely handle the horses;
- therapy with horses within the last two years.

4.4 Sample size calculation

The tool https://architecta.shinyapps.io/SingleCaseDesignsv02/ was used to calculate if the anticipated sample size had sufficient power. The following input was used:

- the maximum number of participants is 30

- the effect size is 0.4
- the number of measurements is 69 (three times a week for 23 weeks)
- minimum number of baseline measurements is 2 (once a week; phase A)
- minimum number of intervention measurements is 15 (once a week; phase B)
- autocorrelation is 0.15
- percentage outliers is 10%
- percentage missing is 35%
- the type 1 error rate (alpha) is 0.05

In order to be able to do the power simulations, we specified the number of samples that have to be drawn (400 samples assignments), and the number sample simulations (100 sample simulations). We found that with a sample size of 30, the estimated power is 0.94. To account for possible dropout, we added another five participants, thus including a total of 35 participants. For the secondary outcomes we will correct for multiple testing by the False Discovery Rate (FDR)

4.5 Study population for the Qualitative study

Purposive sampling will identify those participants who are most likely to provide rich information about their experiences. The study population consists of three groups:

1. Adolescents referred to the intervention (n = 8-10)
2. Caregivers of the adolescents referred to the intervention (n = 8-10)
3. The professionals delivering the intervention (n = 4)

We will use a maximum variation sampling approach of purposefully selecting a wide range of participants (Patton, 2015). The sample size is particularly dependent on the complexity of the research and the heterogeneity of the relevant characteristics. However, it is expected that 8-10 participants will be included in each group (adolescents and caregivers). The final number of subjects, however, will be determined on the basis of saturation. We will interview four members of the professionals group.

4.6 Study population for the cost effectiveness study

Each year we will randomly select two cases from the participants to the EAT intervention. The evaluation sessions will take place approximately four months after the intervention is finished.

5. TREATMENT OF SUBJECTS

5.1 Investigational product/treatment

During the first session(s) of EAT, the personal goals of the adolescent are determined. The goals of the adolescent may differ from the goals of caregivers. This is an essential part of the therapy, as most of the adolescents have been sent to therapies with goals set by adults. The therapist can help the patient to set personal goals by means of observations made and by asking if the patient recognizes what happens during the session with respect to real-life. People act in the same way in most situations, so if the patient does not immediately recognize the pattern, the behavior will soon be repeated and the therapist can help the patient recognize the pattern. Examples of personal goals verbalized by adolescents are: “How can I get friends and keep them?”, “I want to go to school”, “I want to join the army, so I have to be able to control my anger”, ”How can I deal with a traumatic event?”.

After this, the therapist determines the smaller steps that are needed to reach the goal. For example, in case of the goal "How can I get friends and keep them?", the first step is to observe other living beings (in this case horses interacting with each other). Observations can be made from a safe distance and the patient learns to be aware of the nonverbal signals given by the horses. The transfer to real life is facilitated by a homework assignment to observe nonverbal behavior of other people and successful when the adolescents spontaneously starts to observe the nonverbal signals people make. The second step in making friends is to make contact, a skill that can be practiced over and over with the horses. The horses, being a prey and a herd animal, will give immediate and neutral feedback as they are used to carefully observing the surroundings and reacting to each other. The herd is extremely important for their safety and horses observe every change of their fellow horses and, in this case, the patient. Changes can be in muscle tension, breathing, facial expression, posture, etc. Patients can experiment with different behaviors and see, hear and feel the reactions that are evoked in the horse. If the behavior of the adolescent is calm and straightforward, the horse will respond calmly and acceptingly. A third step in making friends is joined interaction with the horse, for instance walking with the horse, grooming the horse, putting the horse on a sail or leading the horse over an obstacle. All these exercises can be made more complex by not using a leading rope. The patient can experiment with different postures, use of voice, energy, velocity, control of movements and handling all kinds of emotions such as fear, frustration, anger or sadness, but also with pride, tenderness and joy. After each activity with the horse, the therapist shares his/her observations with the patient. As a result, the patient becomes aware of his behavior and gets the opportunity to experiment with new behavior by practicing the activity. He is invited to test the new behavior

at home and at school: for example, to observe other people or to adjust his behavior in the same way he did with the horses and tell the next session what happened. We use an experience leaflet including the treatment goals, the experiences of the sessions and the homework to enhance the transfer to real life. Concrete questions are often needed to describe the session (What happened? What did you learn?). If a form of additional support is needed, we can tell the parent(s) about the homework or send a picture of the patient with the horse, etc. In case of interaction problems between the patient and the parents, which nearly always exist inherent to the core of the disorder, parents can be invited to join the therapy. We have special exercises that can be performed with more people, such as leading the horse together through a special obstacle course, grooming the horse while holding each other without talking or switching roles. Usually, two sessions with parents are scheduled.

EAT sessions will be held once a week for 15 weeks. Each session will last 60 minutes using a standardized protocol (ORS, discuss the homework, exercise, feedback, repeat the exercise, new homework, SRS). A session starts with assessing the well-being of the patient at home, at school and in total over the last week by the ORS (Outcome Rating Scale) (5 minutes). The therapist asks what the patient remembers from the last session and how the exercises at home worked out. Then a new exercise is started conform the protocol, if necessary adjusted to the goal of the patient and the possibilities of the patient and the horse. After the session, the patient completes the SRS (Session Rating Scale) questionnaire to promote awareness of the attainment of his goals, to learn to express himself in an acceptable way and to give feedback of the session and the therapist. Subsequently, an appointment is made about the homework to enhance the transfer to real life.

All therapists are BIG/SKJ registered healthcare professionals and qualified to work with horses. Every six months, a consensus meeting will be organized with all EAT-therapists to maintain treatment fidelity. Videos of EAT sessions performed by the therapists will be used for this purpose.

EAT does not seem to have a negative effect on the well-being of the horses, as indicated by unchanged plasma cortisol levels and heart rate variability (Malinowski et al., 2018).

Subject can use medication as usual. As former therapy was ineffective, no other therapy will be given.

6. INVESTIGATIONAL PRODUCT

N/A

7. NON-INVESTIGATIONAL PRODUCT

N/A

8. METHODS

8.1 Study parameters/endpoints

8.1.1 Main study parameter/endpoint

The primary objective is to significantly improve emotion dysregulation in a therapy-resistant group of youths with ASD (aged 11-18) after a 15-week treatment with Equine-Assisted Therapy (EAT).

8.1.2 Secondary study parameters/endpoints (if applicable)

Secondary study parameter is to assess emotion dysregulation in a therapy-resistant group of youths with ASD (aged 11-18) at the long term (1 year) after treatment with Equine-Assisted Therapy (EAT).

Secondary study parameters are to assess the following outcome measures of ASD, symptom severity, quality of life, self-esteem, global and family functioning and goal attainment as well as qualitative outcomes in (a) the short term (15 weeks) and (b) the long term (1 year). Another measure is to quantify the cost of EAT therapy when compared to a scenario including continued CAU.

8.2 Randomisation, blinding and treatment allocation

Participants will be randomly assigned to one of the five pre-defined baseline lengths to increase the internal validity of the design with a 1:1 allocation using permuted blocks of random sizes. The block sizes will not be disclosed to ensure concealment. Participants will be randomized using www.randomization.com, an online randomization tool. Allocation concealment will be ensured, as the person performing the randomization has no other role in the study. This person will prepare the randomization lists and seal the envelopes. The randomization code will not be released until the patient has been recruited for the trial, which will take place after completion of all baseline measurements. To ensure concealment of allocation, the steps involved in randomization, outcome measurements and treatment are separate. All patients who give consent for participation and who comply with the inclusion criteria will be randomized. Randomization will be requested by the research assistant in charge of recruitment and outcome measurement.

8.3 Study procedures for the randomized, multiple-baseline single-case design

EAT sessions will be held once a week for 15 weeks and are like regular standard EAT sessions. Each session will last 60 minutes using a standardized protocol (ORS, discuss the homework, exercise, feedback, repeat the exercise, new homework, SRS). A session starts with assessing the well-being of the patient at home, at school and in total over the last week

by the ORS (Outcome Rating Scale) (5 minutes). The therapist asks what the patient remembers from the last session and how the exercises at home worked out. Then a new exercise is started conform the protocol, if necessary adjusted to the goal of the patient and the possibilities of the patient and the horse. After the session, the patient completes the SRS (Session Rating Scale) questionnaire to promote awareness of the attainment of his goals, to learn to express himself in an acceptable way and to give feedback of the session and the therapist. Subsequently, an appointment is made about the homework to enhance the transfer to real life.

Above this, during the total study period of 23 weeks, parents of the participants will complete an emotion regulation questionnaire (Emotion Dysregulation Inventory) three times a week to answer the primary outcome. Other assessments (ASD symptom severity, quality of life, self-esteem, global and family functioning and goal attainment) will take place at the baseline (T0), at the end of phase A (T1), after completion of phase B (T2), after the end of phase C (T3) and after one year (T4) (please see section outcomes).

At admission to one of the participating centers a psychiatric evaluation will be made resulting in a diagnosis according to Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5). Clinical DSM-5 diagnoses are established by a multidisciplinary team based on information (developmental history, child observation and psychiatric assessment and review of clinical and prior records, including information available from school or other professional institutions involved with the child), gathered by a child psychiatrist and a child psychologist, resulting in a diagnosis. Emotion Dysregulation (primary measure)

Data collection

Emotion Dysregulation Inventory (EDI)

Emotion Dysregulation Inventory (EDI) (Mazefsky, Yu, & Pilkonis, 2020; Mazefsky, Yu, White, Siegel, & Pilkonis, 2018) is specifically designed to measure emotion regulation impairments in youths and adolescents with ASD. The EDI-short form is a validated, change-sensitive, 13-item caregiver report measure of emotion regulation impairment for individuals who are at least 6 years of age. The EDI was developed using the item response theory (IRT) analysis and none of the final items had evidence of differential item functioning (e.g., psychometric biases) by gender, age, intellectual ability and verbal ability, making it suitable for use across heterogeneous populations. Items on the EDI measure how problematic behaviours have been during the past day. The scale used is Not at all=0, Mild=1, Moderate=2, Severe=3, or Very Severe=4. The EDI short form includes two scales: a 7-item Reactivity Index and a 6-item Dysphoria Index. Index raw scores can be converted into t-

scores or theta scores based on a sample of 1755 individuals with ASD (Mazefsky et al., 2018) or based on a sample of 1000 youths matching the US census as general population norms (Mazefsky et al., 2020). For the purposes of this study we will administer the 13-item EDI short form three times a week (for 23 weeks).

Quality of life

Kidscreen-27 (parents) and Kidscreen-27 (adolescents). The KIDSCREEN-27 (https://www.kidscreen.org/english/questionnaires/kidscreen-27-short-version/) is a generic health-related quality of life (HRQOL) questionnaire for children and adolescents applicable to healthy and chronically ill children and adolescents aged between 8 and 18 years. There are two versions of the questionnaire: a self-complete version (child/ adolescent) and a proxy version (parent/proxy). The KIDSCREEN-27 consists of 27 items measuring five dimensions: physical well-being, psychological well-being, parent relations & autonomy, social support & peers and school environment. Items are answered on a five-point Likert-type scale assessing frequency (never (1), seldom (2), sometimes (3), often (4), and always (5)) or intensity (not at all (1), slightly (2), moderately (3), very (4), and extremely (5)) with a one-week recall period. Scores are coded from 1 to 5, negatively formulated items are recoded and the sum scores for respective dimensions are converted into T scores with a mean of 50 and a standard deviation (SD) of 10. Higher scores indicate better HRQOL. The KIDSCREEN-27 has been shown to have robust psychometric properties. The internal consistency of the domains was between 0.81 and 0.84, and the test-retest reliability of the domains ranged from 0.61 to 0.74. The KIDSCREEN-27 (parents and adolescents) will be assessed at baseline (T0), at the end of phase A (T1), after completion of phase B (T2), after the end of phase C (T3) and after one year (T4).

Psychometric Evaluation

Global functioning (CBCL/6-18, YRS/11-18, TRF/6-18) For assessing emotional and/or behavioural problems we will use the CBCL/6-18 (https://aseba.org/), completed by parents or substitutes, the TRF/6-18, completed by teachers and other school staff, and the YSR/11-18, completed by youths. All three questionnaires include more than 100 items assessing behavioural and emotional problems that are answered on a three-point Likert-type scale (0 = not true, 1 = somewhat or sometimes true, 2 = very true or often true) by parents. The scores will display eight problem scales (withdrawn (1); somatic (2); anxious (3); social (4); thought (5); attention (6); rule-breaking (7); aggressive (8)) and other problems. The sum of the problem scales 1, 2 and 3 form the 'internalizing behaviour' scale, whereas 7 and 8 form 'externalizing behaviour'. All subscales together count for the total problem scale. Some items contribute to more than one problem scale. T-scores are computed from raw scores;

higher scores on the syndrome scales indicate a greater severity of problems. A T-score of 63 (90th percentile) demarcates the clinical range, which is an indication that a child needs professional help. For the competence scales, lower scores indicate greater severity. A T-score < 37 indicates the clinical range. The CBCL/ 6–18 has well-established psychometric properties in clinical, non-clinical, and cross-cultural populations. The CBCL 6-18 (parents), YSR (adolescents), TRF (teacher) will be assessed at baseline (T0).

Global functioning (BPM-P, BPM-Y, BPM-T) The Brief Problem Monitor (BPM) (https://aseba.org/) is a rating instrument for monitoring children's functioning and responses to interventions and is a short version of the CBCL/YSR/TRF. The BPM consist of a minimum set of items to be completed in 1 to 2 minutes by parents (BPM-P), teachers (BPM-T) and youths (BPM-Y). Nineteen items of the CBCL/6–18 (or TRF or YSR) make up the BPM: 6 items for attention (all 6 from the CBCL/6–18 attention problems scale), 7 items for externalizing (all 7 from the CBCL/6–18 aggression scale), and 6 items for internalizing (5 from the CBCL/6–18 anxiety/depression and 1 from the CBCL/6–18 withdrawn/depressed scales). Each item is rated 0 = not true, 1 = somewhat true, or 2 = very true. Parallel items and scales on the BPM & the CBCL/6-18, TRF and YSR enable us to link comprehensive initial and outcome assessments to BPM scores. The BPM-P, BPM-Y will be assessed at the end of phase A (T1), after completion of phase B (T2), after the end of phase C (T3) and after one year (T4). The BPM-T (teachers) will be assessed after the end of phase C (T3).

Communication and social functioning (SRS-2)
The Social Responsiveness Scale – Second edition (SRS-2) (Constantino et al., 2003) measures deficits in social behaviour associated with ASD and can be used to assess the severity of symptoms in ASD (Frazier et al., 2012). The questionnaire will be completed by multiple raters (parents and teachers). The SRS-2 consists of 65 items scored in a Likert-like scale format ranging from not true=1, sometimes true=2, often true=3 to almost always true =4. It is designed to identify social impairment intrinsic to ASD and to quantify its severity across the duration of the treatment. A total score and five treatment subscale scores (Social Awareness; Social Cognition; Social Communication; Social Motivation; and Restricted Interests and Repetitive Behaviour) are obtained. The accepted diagnostic criteria (cut point) for the SRS-2 for the association with a diagnosis of ASD are: <=59 (normal); 60–75 (mild to moderate ASD); and =>76 (severe ASD). The SRS-2 (parents) will be assessed at baseline (T0), at the end of phase A (T1), after completion of phase B (T2), after the end of phase C (T3) and after one year (T4). The SRS-2 (teachers) will be assessed at the baseline (T0) and after the end of phase C (T3).

Rosenbergh self-esteem (RSES)

The Rosenberg Self-Esteem Scale (RSES) (Rosenberg, 1979) will be used to assess self-esteem. It is a widely used 10-item Likert-scale self-esteem measure. Items are answered on a four-point scale — ranging from strongly agree to strongly disagree — measuring positive and negative feelings towards the self. The Dutch version of the RSES is found to be a one-dimensional scale with high internal consistency and congruent validity and a Cronbach's alpha of 0.89 (Franck, De Raedt, Barbez, & Rosseel, 2008). Rosenberg Self-Esteem Scale will be assessed at the baseline (T0), at the end of phase A (T1), after completion of phase B (T2), after the end of phase C (T3) and after one year (T4).

Family functioning, prior beliefs, satisfaction

Family functioning

Family functioning is assessed using a validated questionnaire: the Family Functioning Questionnaire (VGFO, 34 items) (https://www.praktikon.nl/wat-we-doen/vragenlijsten/vgo). The Family Functioning Questionnaire can be answered on a four-point scale ranging from 1 (not applicable) to 4 (completely applicable) with lower scores indicating more problems in family functioning. This questionnaire will be assessed at the baseline (T0) and after the end of phase C (T3).

Prior beliefs and motivation for therapy, satisfaction

Parents and adolescents prior beliefs about the short-term and long-term success and burden of the intervention are evaluated using a 6-item questionnaire in which parents and adolescents rate their prior believes on a 10-point scale ranging from 1 (completely disagree) to 10 (completely agree). This measurement will also be used to examine if randomization is successful by verifying if the five groups do not differ in believes about the effectiveness of the intervention at the baseline. Prior beliefs of parents and adolescents will be assessed at the baseline (T0). Parents and adolescents can rate their overall treatment trajectory on a scale of 0 to 10. Satisfaction will be measured after the end of phase C (T3).

Goal attainment and adherence

SRS/ORS

For collecting client feedback we will use two brief questionnaires, the Outcome Rating Scale (ORS) and Session Rating Scale (SRS), which will be easily administered on a regular basis during treatment (https://www.scottdmiller.com/ scholarly-publications-hand-outs-vitae/).This allows treatment sessions to be evaluated at any time to ascertain whether or not individual treatments are 'on the right track' towards a successful outcome. The ORS is primarily focused on the well-being of the client and is administered at the beginning of the treatment

session. The SRS is administered at the end of the session and deals with how the client has experienced the treatment session. The outcomes of the questionnaires are reflected in a graph on an iPad (or on paper when an iPad is not available) per interview to allow the height of the score and progress to be visualized during the sessions. The ORS and SRS both have a high internal consistency and an average correlation with other outcome measurements. The therapist will assess the SRS/ORS each session.

Goal Attainment Scale

Goal Attainment Scale (Kiresuk & Sherman, 1968) is a method of scoring the extent to which the patient's individual goals are achieved in the course of the intervention. In effect, each patient has his own outcome measure, but this is scored in a standardized way as to allow statistical analysis. Each goal is rated on a six-point scale, capturing the degree of attainment for each goal area: Patient starts on -2. When he achieves somewhat better this is scored -1. When the patient achieves the expected level, this is scored 0. When he achieves a better than expected outcome this is scored +1 (somewhat better) or +2 (much better). Achieving a worse than expected outcome is scored -3. In this study, a maximum of 3-4 goals are identified, which are incorporated into a single GAS score. The GAS will be assessed in the fifth, tenth en fifteenth sessions.

8.4 Study procedures for the qualitative study

The first contact of adolescents and caregivers will be done by the professional involved in the intervention. Then the qualitative researcher will be introduced by the professional. Caregivers will receive a formal letter explaining the qualitative part of the study and an informed consent form. The qualitative researcher will only contact those who have given their written informed consent to schedule an interview. All interviews will be assessed by the qualitative researcher together with an experienced expert parent or a young experienced expert with autism (depending on the subgroup). Interviews will last for 45–60 minutes. Caregivers and adolescents will be interviewed separately. Each participant will be interviewed twice (before the start of the intervention and 30 weeks after the intervention).

The professionals will be interviewed at their workplace at a convenient time for them. All interviews will be assessed by a qualitative researcher and an experienced mental health professional. Interviews will last for 45–60 min. Each professional will be interviewed twice (at the start of the project and when all participants are included).

Data collection

A total of n=8-10 adolescents and their caregivers will be invited to participate in individual semi-structured in-depth interviews of one hour. The interview will consist of open questions in order to obtain a complete as possible picture of the participants' experiences (Kvale & Brinkmann, 2009). We work with an iterative topic list. The topic list will be exploratory in nature and focus on gathering knowledge and experiences about the topic. Possible examples of questions for the adolescents and their parents include: What are your expectations of the intervention? What should be maintained or strengthened in the intervention? What are facilitators, barriers? What benefits have you derived from the intervention? How did you experience the intervention? What was your experience with the therapist?, etc.

The EAT therapists will be interviewed to describe and evaluate therapeutic elements and implementation processes. Possible examples of questions are: What characterizes the referral procedure for the intervention? What are the barriers and facilitators for referral to the intervention? How do participants perceive the intervention offered? What aspects are missing from your perspective? How is the intervention implementation experienced by EAT therapists?, etc.

8.5 Study procedures for the cost-effectiveness study

We will organize three evaluation sessions. In each session a participant of the main study will be invited. All sessions will have a trained panel chairman. Furthermore, five to eight persons are invited to participate. The persons are all affected by or have contributed to the intervention and add to a variety of perspectives (for example, parents, teacher, professional, social worker, etc.). Depending on the ability, the adolescent may attend the session. The conversations will take 3 hours at most, including several breaks.

Data collection

Two questions are central to this conversation: "What has changed in the lives of the participants after the EAT intervention?" and "What would (hypothetical) have happened if the EAT intervention was not conducted in terms of well-being, lifetime costs and care expenses?"

In the first part of the conversation - the reconstruction - those present will make a reconstruction of the period from admission to the intervention. The moderator will make sure that there will be a complete overview of the events, which will be composed as much as possible prior to the meeting based on available data in the patient clinical files. During the conversation, information will be structured according to experiences and events (What

happened in the child's life? How did the child experience the interventions over time?) and actions (Which interventions and actions have been used?). In the second part of the conversation the (added) value of the intervention for the participant will be discussed. To do so, a reference situation will be drawn (Where would the participant have ended up without the intervention? or Which (care) facilities did the participant previously make use of and What facilities were addressed in this period without the intervention?).

At the end of the meeting, five of the more reflective questions will be presented to those involved.

1. Please rate: How do you rate the intervention?
2. Social contacts: What has the intervention contributed to your social life?
3. Thresholds: What made it difficult to participate in the intervention?
4. Cooperation: How did the people of the intervention work together with other parties?
5. Points for improvement: If you could improve the intervention, what would you suggest?

8.6 Withdrawal of individual subjects

Subjects can leave the study at any time for any reason if they wish to do so without any consequences. The investigator can decide to withdraw a subject from the study for urgent medical reasons.

8.6.1 Specific criteria for withdrawal (if applicable)

n/a

8.7 Replacement of individual subjects after withdrawal

Individual subjects are replaced if they leave the study till 35 participants have participated in more than 10 sessions. We will investigate the reason why participants leave the study so we can adjust the EAT or the measurements, if possible without methodological consequences.

8.8 Follow-up of subjects withdrawn from treatment

Regular therapy will be continued after subjects have left the study

8.9 Premature termination of the study

The study will only be terminated when the research team, therapists and the patient panel team agree to stop.

Table. Measures and Frequency

Domain	Measure	Screening		Phase A		Phase B Intervention													Phase C						follow-up
		Admission		1	2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1	2	3	4	5	6	
		to Karakter	T0		T1															T2					T3	T4
Parents																											
Descriptives	Intake questionnaire	X																									
	IQ screener	X																									
Emotion Dysregulation (primary measure)	Reactivity and Dysfory		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Reactivity			2X	2X	2X	2X	2X	2X	2X	2X	2X	2X	2X	2X	2X	2X	2X	2X	2X	2X	2X	2X	2X	2X	2X	
Quality of Life	Kidscreen-27 parents as informant		X		X															X						X	X
Psychometric Evaluation	CBCL 6-18		X																								
	BPM-P				X															X						X	X
	Communication and social functioning (SRS-2)		X		X															X						X	X
Family functioning, prior beliefs, satisfaction	Family functioning (VGFO)		X																							X	
	Prior beliefs		X																								
	Satisfaction																									X	

Domain	Measure	Screening		Phase A		Phase B Intervention													Phase C						follow-up
		Admission		1	2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1	2	3	4	5	6	
		to Karakter	T0		T1															T2					T3	T4
Adolescents																											
Quality of Life	Kidscreen-27 kids		X		X															X						X	X
Psychometric Evaluation	YSR		X																								
	BPM-Y				X															X						X	X
	Rosenbergh selfesteem		X		X															X						X	X
Prior beliefs/ Satisfaction	Prior beliefs		X																								
	Satisfaction																									X	

Domain	Measure	Screening		Phase A		Phase B Intervention													Phase C						follow-up
		Admission		1	2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1	2	3	4	5	6	
		to Karakter	T0		T1															T2					T3	T4
Teachers																											
Psychometric Evaluation	TRF		X																								
	BPM-T																									X	
	Communication and social functioning (SRS-2)		X																							X	

Domain	Measure	Screening		Phase A		Phase B Intervention													Phase C						follow-up
		Admission		1	2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1	2	3	4	5	6	
		to Karakter	T0		T1															T2					T3	T4
Professionals																											
Descriptives	Psychiatric diagnoses DSM-V	X																									
Goal attainment and adherence	Intervention adherence					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X							
	Outcome Rating Scale (ORS) (each session)					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X							
	Session Rating Scale (SRS) (each session)					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X							
	Goal attainment				X					X					X					X							

Note. The lenght of phase A and C differers depending on the group the participant is randomized into.

9. SAFETY REPORTING

9.1 Temporary halt for reasons of subject safety

In accordance to section 10, subsection 4, of the WMO, the sponsor will suspend the study if there is sufficient ground that continuation of the study will jeopardise subject health or safety. The sponsor will notify the accredited METC without undue delay of a temporary halt including the reason for such an action. The study will be suspended pending a further positive decision by the accredited METC. The investigator will take care that all subjects are kept informed.

9.2 AEs, SAEs and SUSARs

9.2.1 Adverse events (AEs)

Adverse events are defined as any undesirable experience occurring to a subject during the study, whether or not considered related to the experimental intervention. All adverse events reported spontaneously by the subject or observed by the investigator or his staff will be recorded.

9.2.2 Serious adverse events (SAEs)

A serious adverse event is any untoward medical occurrence or effect that

- results in death;
- is life threatening (at the time of the event);
- requires hospitalisation or prolongation of existing inpatients' hospitalisation;
- results in persistent or significant disability or incapacity;
- is a congenital anomaly or birth defect; or
- any other important medical event that did not result in any of the outcomes listed above due to medical or surgical intervention but could have been based upon appropriate judgement by the investigator.

An elective hospital admission will not be considered as a serious adverse event.

The investigator will report all SAEs to the sponsor without undue delay after obtaining knowledge of the events.

The sponsor will report the SAEs through the web portal *ToetsingOnline* to the accredited METC that approved the protocol, within 7 days of first knowledge for SAEs that result in death or are life threatening followed by a period of maximum of 8 days to complete the initial preliminary report. All other SAEs will be reported within a period of maximum 15 days after the sponsor has first knowledge of the serious adverse events.

9.2.3 Suspected unexpected serious adverse reactions (SUSARs)

n/a

9.3 Annual safety report

n/a

9.4 Follow-up of adverse events

All AEs will be followed until they have abated, or until a stable situation has been reached. Depending on the event, follow up may require additional tests or medical procedures as indicated, and/or referral to the general physician or a medical specialist.
SAEs need to be reported till end of study within the Netherlands, as defined in the protocol

10. STATISTICAL ANALYSIS

10.1 Study parameter(s) for the randomized, multiple-baseline single-case design

Data will be analysed using SPSS V.24.0 for Windows (SPSS Incorporated), specific software for single-case experimental designs available on https://architecta.shinyapps.io/SingleCaseDesignsv02/ and the statistical software R (Comprehensive R Archive Network; https://cran.r-project.org/) for the hierarchical linear modelling. Descriptive statistics will be performed for baseline characteristics of the study population. Parametric data will be presented as means with SD, and non-parametric distributed variables as median with IQRs. For the primary outcome we will combine visual analyses, randomization tests and effect sizes for a comprehensive analysis of intervention based on the state-of-the-art knowledge on analysing a randomized multiple-baseline single-case design (Heyvaert et al., 2017).

Time series analyses

Firstly, data will be visually inspected within and between phases with respect to 1) level 2) trend 3) variability 4) immediacy of the effect 5) overlap and 6) consistency of data across similar phases (Bulté & Onghena, 2012; Kratochwill et al., 2010). Secondly, the intervention effect will be statistically tested by applying a sequential approach. We will test the null hypothesis of no treatment effect for any of the cases with the non-parametric approach to combine randomization test p values (Manly, 2007). If it is determined that the intervention has indeed a statistically significant effect on emotion regulation, we will use the hierarchical linear modelling (Van Den Noortgate & Onghena, 2003) approach for analysing the data in closer detail and to obtain the parameter estimate of and to further model the average treatment effect and individual treatment effects. This average treatment effect is an effect size and will as such be compared with the meta-analytic effect size as obtained in other studies. There will be no statistical testing in this final step. The interaction between treatment effect and time will be included in the hierarchical linear model in order to model differential trends.

Single time points analyses

Pre-treatment scores on the EDI will be compared to post-treatment and follow-up scores on the EDI and other secondary measurements using a non-parametric Wilcoxon signed rank test. An effect size will be obtained from this analysis, with the follow-up versus pre-treatment comparison as input for this effect size.

10.2 Study parameter(s) for the qualitative study

The interviewer will audiotape the interview with the participant's permission. Immediately after each interview participants', non-verbal behaviors and emotions will be logged.

Furthermore, a reflexive journal will record the overall process of data collection. The data collection and analysis will occur concurrently. After each interview, the researcher will transcribe the recording verbatim and then code the transcript to reveal broad or initial categories or themes. Interview transcripts and observation narratives will be coded thematically by two researchers independently.

Atlas-ti will be used for coding, management and retrieval of data. Throughout the analysis, researchers will use memos and reflective diaries to engage with the data and refine emergent themes through an iterative and inductive process. Data triangulation will be achieved by comparing interview data from adolescents and caregivers to explore the intervention from different perspectives.

10.3 Study parameter(s) for the cost-effectiveness study

We will use the 'Outcome generator' (https://effectencalculator.nl/) in the cost-effectiveness study to answer questions on cost evaluation and the effect of improved well-being. This instrument helps to make statements on innovative interventions based on dialogue. Using the 'Outcome generator' we will be able to analyze the social benefits of EAT and the costs saved by society due to positive results of EAT.

The outcomes are placed on a timeline and described in terms of social benefits (such as: went back to school, less reactivity and improved emotion regulation, less violent/getting angry, etc.) and saved costs (less admission to institutions, less therapy, future perspective on work, etc.), including a corresponding price tag. The result is a visual overview of the social benefits and cost savings offered by the intervention.

10.4 Mixed methods

We will use an embedded mixed-methods methodology to triangulate both qualitative, quantitative methodologies and the cost-effectiveness in this study (Creswell & Plano Clark, 2011). Qualitative interviews will be embedded in the randomized, multiple-baseline single-case design to evaluate the effectiveness of the EAT intervention. Data will be collected concurrently during the project. Using a mixed-methods design we enable us to combine the results of the three components and explain agreements and disagreements between the findings. Furthermore, the use of qualitative data will provide complementary data about the nature, acceptability, perceived usefulness, barriers and issues of the EAT intervention. It will enable us to draw more meaningful conclusions about the EAT intervention.

11. ETHICAL CONSIDERATIONS

11.1 Regulation statement

The current study is conducted according to the principles of the "Declaration of Helsinki", adopted by the 18th WMA General Assembly, Helsinki, Finland, June 1964 and amended by the 64th WMA General Assembly, Fortaleza, Brazil, October 2013, in accordance with the Medical Research Involving Human Subjects Act (WMO) and in accordance to the Guidelines for Good Clinical Practice (GCP) (CPMP/ICH/135/95 – 17th of July 1996).
The protocol of the current study will be submitted to the local "Commissie Mensgebonden Onderzoek" (CMO) Arnhem-Nijmegen, an official METC in The Netherlands. Possible participants and their parents will not be approached before formal approval has been granted.

11.2 Recruitment and consent

We will recruit therapy-resistant and therapy-avoiding youths with ASD who are motivated to undergo Equine-Assisted Therapy (EAT) through screening by the therapists in the participating centers. Further inclusions and exclusions will be assessed using a checklist during an intake interview by a research assistant. Patients and their caregivers are informed about the intervention, the aims of the study and the measurements. They receive this information both orally and in print. After one week, participants aged 12 to 17 and their parents will be asked to sign a written informed consent, children aged 11 will be asked to sign a child version.

11.3 Objection by minors or incapacitated subjects

The code of conduct relating to expressions of objection by minors participating in medical research is applicable in the current study (https://english.ccmo.nl/binaries/ccmo-en/documents/publications/2002/01/01/code-of-conduct-relating-to-expressions-of-objection-by-minors-participating-in-medical-research/Code+of+conduct+relating+to+expressions+of+objection+by+minors+participating+in+medical+research.pdf).

11.4 Benefits and risks assessment, group relatedness

The potential value of this study is that this study will create an opportunity for therapy resistant adolescents to profit from EAT, especially with regard to both social functioning and emotion dysregulation.

Given that the intake procedure and the treatment sessions are part of care as usual, only the filling out of the questionnaires, and for some joining an interview, are extra. Since routine outcome monitoring is part of care as usual as well, the extra burden is relatively small. Based on a pilot study in n=6 adolescents and their parents (May/June 2020), the average time to complete the T0 questionnaires was 15 minutes for adolescents (Kidscreen-27, YSR 11-18, Rosenberg self-esteem and prior beliefs) and 25 minutes for parents (Kidscreen-27, CBCL 6-18, VGFO, EDI, SRS-2 and prior beliefs). Furthermore, we asked caregivers to complete the EDI three times a week for three weeks. None of the parents in the pilot felt completing these questions as a burden.

11.5 Compensation for injury

The investigator has a liability insurance which is in accordance with article 7 of the WMO.

The sponsor (also) has an insurance which is in accordance with the legal requirements in the Netherlands (Article 7 WMO). This insurance provides cover for damage to research subjects through injury or death caused by the study. The insurance applies to the damage that becomes apparent during the study.

11.6 Incentives

Compensation for extra traveling costs to the horse center will be given to a max of 50 euro. The EAT will be paid by municipality or the research team.

12. ADMINISTRATIVE ASPECTS, MONITORING AND PUBLICATION

12.1 Handling and storage of data and documents

The data of the current study are handled with confidentiality. The handling of personal data of subjects is in accordance with the European General Data Protection Regulation (in Dutch: Algemene Verordening Gegevensbescherming, AVG). In accordance with the AVG, the data protection officer (dpo@karakter.com) has been appointed. The research team is responsible for the processing of the data.

After a participant is included in the study, a code number (i.e. 101,102,...135) is assigned by one of the investigators. The code list will be digitally stored on the H-drive of Karakter, is password protected and only accessible to the investigators that are involved in the project. The location of the code list is different from the location of the data. Data is stored on a secure CASTOR EDC server at Radboud University Medical Centre, which is located on two separate sites and a back-up is made on a daily basis. CASTOR EDC is in accordance with the guidelines for valid data storage systems of the GCP (http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=11).

Paper documents, such as the consent forms, will be stored separately and secured for 10 years. The investigator will retain originals of all source documents for a period of 5 years. In accordance with the GCP guidelines, all study-related documents are archived for at least 15 years.

When required for urgent medical reasons, responsible medical professionals will receive access to the personal source data. When this is needed, subject and their parents/legal caregivers are informed.

After completion of the study and study reporting, data will be transferred to the Dutch National Data Archive (DANS). From DANS data are accessible for all researchers in the Netherlands. Researchers from outside the Netherlands can have access to the data according to the guidelines from DANS

12.2 Monitoring and Quality Assurance

n/a

12.3 Amendments

Amendments are changes made to the research after a favourable opinionapproval by the accredited METC has been given. All amendments will be notified to the METC that gave a favourable opinionapproval.

Non-substantial amendments will not be notified to the accredited METC and the competent authority, but will be recorded and filed by the sponsor.

12.4 Annual progress report

The sponsor/investigator will submit a summary of the progress of the trial to the accredited METC once a year. Information will be provided on the date of inclusion of the first subject, numbers of subjects included and numbers of subjects that have completed the trial, serious adverse events/ serious adverse reactions, other problems, and amendments.

12.5 Temporary halt and (prematurely) end of study report

The investigator/sponsor will notify the accredited METC of the end of the study within a period of 8 weeks. The end of the study is defined as the last patient's last measurement.

The sponsor will notify the METC immediately of a temporary halt of the study, including the reason of such an action.

In case the study is ended prematurely, the sponsor will notify the accredited METC within 15 days, including the reasons for the premature termination.

Within one year after the end of the study, the investigator/sponsor will submit a final study report with the results of the study, including any publications/abstracts of the study, to the accredited METC.

12.6 Public disclosure and publication policy

Prospective trial registration will be done in the Netherlands Trial Register (NTR). The results of the study will be submitted to a peer-reviewed journal and presented at both national and international conferences. Disclosure of results will be conducted according to the CCMO statement on publication policy. This indicates that both positive and negative results will be published.

After completion of the study, data will be transferred to the Dutch National Data Archive (DANS). From DANS data are accessible for all researchers in the Netherlands. Researchers

from outside the Netherlands can have access to the data according to the guidelines from DANS.

13. STRUCTURED RISK ANALYSIS

n/a

14. REFERENCES

Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clin Psychol Rev, 30*(2), 217-237. doi: 10.1016/j.cpr.2009.11.004

Bulté, I., & Onghena, P. (2012). When the truth hits you between the eyes: A software tool for the visual analysis of single-case experimental data. *Meth Eur J Res Meth Behav Soc Sci, 8*(3), 104-114. doi: 10.1027/1614-2241/a000042

Burnett Heyes, S., Jih, Y. R., Block, P., Hiu, C. F., Holmes, E. A., & Lau, J. Y. (2015). Relationship Reciprocation Modulates Resource Allocation in Adolescent Social Networks: Developmental Effects. *Child Dev, 86*(5), 1489-1506. doi: 10.1111/cdev.12396

Constantino, J. N., Davis, S. A., Todd, R. D., Schindler, M. K., Gross, M. M., Brophy, S. L., . . . Reich, W. (2003). Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised. *J Autism Dev Disord, 33*(4), 427-433. doi: 10.1023/a:1025014929212

Creswell, W. J., & Plano Clark, V. L. (2011). *Designing and conduction mixed methods research.* (2nd ed.). Thousand Oaks, California, USA: Sage.

Fein, D., Barton, M., Eigsti, I. M., Kelley, E., Naigles, L., Schultz, R. T., . . . Tyson, K. (2013). Optimal outcome in individuals with a history of autism. *J Child Psychol Psychiatry, 54*(2), 195-205. doi: 10.1111/jcpp.12037

Fombonne, E. (2018). Editorial: The rising prevalence of autism. *J Child Psychol Psychiatry, 59*(7), 717-720. doi: 10.1111/jcpp.12941

Franck, E., De Raedt, R., Barbez, C., & Rosseel, Y. (2008). Psychometric properties of the Dutch Rosenberg Self-Esteem Scale. *Psychologica Belgica, 48*(1), 25-35. doi: 10.5334/pb-48-1-25

Frazier, T. W., Youngstrom, E. A., Speer, L., Embacher, R., Law, P., Constantino, J., . . . Eng, C. (2012). Validation of proposed DSM-5 criteria for autism spectrum disorder. *J Am Acad Child Adolesc Psychiatry, 51*(1), 28-40.e23. doi: 10.1016/j.jaac.2011.09.021

Fuller, E. A., & Kaiser, A. P. (2020). The Effects of Early Intervention on Social Communication Outcomes for Children with Autism Spectrum Disorder: A Meta-analysis. *J Autism Dev Disord, 50*(5), 1683-1700. doi: 10.1007/s10803-019-03927-z

Gillespie-Lynch, K., Kapp, S. K., Brooks, P. J., Pickens, J., & Schwartzman, B. (2017). Whose Expertise Is It? Evidence for Autistic Adults as Critical Autism Experts. *Front Psychol 8*(438). doi: 10.3389/fpsyg.2017.00438

Heyvaert, M., Moeyaert, M., Verkempynck, P., Van den Noortgate, W., Vervloet, M., Ugille, M., & Onghena, P. (2017). Testing the Intervention Effect in Single-Case Experiments: A Monte Carlo Simulation Study. *J Exp Educ, 85*(2), 175-196. doi: 10.1080/00220973.2015.1123667

Kiresuk, T. J., & Sherman, R. E. (1968). Goal attainment scaling: A general method for evaluating comprehensive community mental health programs. *Community Ment Health J, 4*(6), 443-453. doi: 10.1007/bf01530764

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single-case designs technical documentation. . Retrieved 17 may 2020, from What Works Clearinghouse Website https://ies.ed.gov/ncee/wwc/Document/229

Kvale, S., & Brinkmann, S. (2009). *Interviews: learning the craft of qualitative research interviewing*. Thousand Oaks, California, USA: Sage.

Malinowski, K., Yee, C., Tevlin, J. M., Birks, E. K., Durando, M. M., Pournajafi-Nazarloo, H., . . . McKeever, K. H. (2018). The Effects of Equine Assisted Therapy on Plasma Cortisol and Oxytocin Concentrations and Heart Rate Variability in Horses and Measures of Symptoms of Post-Traumatic Stress Disorder in Veterans. *J Equine Vet Sci, 64*, 17-26. doi: 10.1016/j.jevs.2018.01.011

Malter Cohen, M., Tottenham, N., & Casey, B. J. (2013). Translational developmental studies of stress on brain and behavior: implications for adolescent mental health and illness? *Neuroscience, 249*, 53-62. doi: 10.1016/j.neuroscience.2013.01.023

Manly, B. F. J. (2007). Randomization Tests, 4th Edition by Eugene S. Edgington, Patrick Onghena. *Int Stat Rev, 75*(2), 269-269. doi: 10.1111/j.1751-5823.2007.00015_21.x

Mazefsky, C. A. (2015). Emotion Regulation and Emotional Distress in Autism Spectrum Disorder: Foundations and Considerations for Future Research. *J Autism Dev Disord, 45*(11), 3405-3408. doi: 10.1007/s10803-015-2602-7

Mazefsky, C. A., Borue, X., Day, T. N., & Minshew, N. J. (2014). Emotion regulation patterns in adolescents with high-functioning autism spectrum disorder: comparison to typically developing adolescents and association with psychiatric symptoms. *Autism Res, 7*(3), 344-354. doi: 10.1002/aur.1366

Mazefsky, C. A., Yu, L., & Pilkonis, P. A. (2020). Psychometric Properties of the Emotion Dysregulation Inventory in a Nationally Representative Sample of Youth. *J Clin Child Adolesc Psychol*, 1-13. doi: 10.1080/15374416.2019.1703710

Mazefsky, C. A., Yu, L., White, S. W., Siegel, M., & Pilkonis, P. A. (2018). The emotion dysregulation inventory: Psychometric properties and item response theory calibration in an autism spectrum disorder sample. *Autism Res, 11*(6), 928-941. doi: 10.1002/aur.1947

McDaniel Peters, B. C., & Wood, W. (2017). Autism and Equine-Assisted Interventions: A Systematic Mapping Review. *J Autism Dev Disord, 47*(10), 3220-3242. doi: 10.1007/s10803-017-3219-9

Mengler, L., Khmelinskii, A., Diedenhofen, M., Po, C., Staring, M., Lelieveldt, B. P., & Hoehn, M. (2014). Brain maturation of the adolescent rat cortex and striatum: changes in volume and myelination. *Neuroimage, 84*, 35-44. doi: 10.1016/j.neuroimage.2013.08.034

Mogensen, L., & Mason, J. (2015). The meaning of a label for teenagers negotiating identity: experiences with autism spectrum disorder. *Sociol Health Illn, 37*(2), 255-269. doi: 10.1111/1467-9566.12208

Nader-Grosbois, N., & Mazzone, S. (2014). Emotion Regulation, Personality and Social Adjustment in Children with Autism Spectrum Disorders. *Psychology, 5*, 1750-1767. doi: 10.4236/psych.2014.515182

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice*. Thousand Oaks, California, USA: Sage.

Peters-Scheffer, N., Didden, R., Korzilius, H., & Sturmey, P. (2011). A meta-analytic study on the effectiveness of comprehensive ABA-based early intervention programs for children with Autism Spectrum Disorders. *Res Autism Spectr Disord, 5*, 60-69. doi: 10.1016/j.rasd.2010.03.011

Reichow, B. (2012). Overview of meta-analyses on early intensive behavioral intervention for young children with autism spectrum disorders. *J Autism Dev Disord, 42*(4), 512-520. doi: 10.1007/s10803-011-1218-9

Rosenberg, M. (1979). *Conceiving the self*: Basic Books.

Samson, A. C., Wells, W. M., Phillips, J. M., Hardan, A. Y., & Gross, J. J. (2015). Emotion regulation in autism spectrum disorder: evidence from parent interviews and children's daily diaries. *J Child Psychol Psychiatry, 56*(8), 903-913. doi: 10.1111/jcpp.12370

Sussman, D., Leung, R. C., Vogan, V. M., Lee, W., Trelle, S., Lin, S., . . . Taylor, M. J. (2015). The autism puzzle: Diffuse but not pervasive neuroanatomical abnormalities in children with ASD. *Neuroimage Clin, 8*, 170-179. doi: 10.1016/j.nicl.2015.04.008

Swain, D., Scarpa, A., White, S., & Laugeson, E. (2015). Emotion Dysregulation and Anxiety in Adults with ASD: Does Social Motivation Play a Role? *J Autism Dev Disord, 45*(12), 3971-3977. doi: 10.1007/s10803-015-2567-6

Trzmiel, T., Purandare, B., Michalak, M., Zasadzka, E., & Pawlaczyk, M. (2019). Equine assisted activities and therapies in children with autism spectrum disorder: A systematic review and a meta-analysis. *Complement Ther Med, 42*, 104-113. doi: 10.1016/j.ctim.2018.11.004

Van Den Noortgate, W., & Onghena, P. (2003). Estimating the mean effect size in meta-analysis: bias, precision, and mean squared error of different weighting methods. *Behav Res Methods Instrum Comput, 35*(4), 504-511. doi: 10.3758/bf03195529