Assessment and treatment of planning skills in adolescents with ADHD

Boyer, B.E.

Publication date
2016
Document Version
Final published version

Chapter 5

Qualitative treatment-subgroup interactions in treatments for adolescents with ADHD: What cognitive behavioral treatment works for whom?

Based on:

Boyer*, **B. E.**, Doove*, L. L., Geurts, H. M., Prins, P. J. M., Van Mechelen, I., & Van der Oord, S. (*accepted).* Qualitative Treatment-Subgroup Interactions in Treatments for Adolescents with ADHD: What Cognitive-Behavioral Treatment Works for Whom? *PLOS ONE*.

* These authors contributed equally to this work

Abstract

Objective. This study investigated qualitative treatment-subgroup interactions within data of an RCT with two cognitive behavioral therapies (CBT) for adolescents with ADHD: a planning-focused treatment (PML) and a solution-focused treatment (SFT). Qualitative interactions imply that which treatment is best, differs across subgroups of patients, and are therefore most relevant for personalized medicine.
Methods. Adolescents with ADHD (n=159) received either PML or SFT. Pre-, post- and three-month follow-up data were gathered on parent-rated ADHD symptoms and planning problems. Pretreatment characteristics were explored as potential qualitative moderators of pretest to follow-up treatment effects, using an innovative analyses technique (QUINT; Dusseldorp & Van Mechelen, 2014). In addition, qualitative treatment-subgroup interactions for the therapeutic changes from pre- to posttest and from post- to follow-up test were investigated.
Results. For the entire time span from pretest to follow-up only a quantitative interaction showed up, while from posttest to follow-up qualitative interactions were found: Adolescents with less depressive symptoms but more anxiety symptoms showed more improvement when receiving PML than SFT, while for other adolescents the effects of PML and SFT were comparable.
Discussion. Whereas subgroups in both treatments followed different trajectories, no subgroup was found for which SFT outperformed PML in terms of the global evolution from pretest to three months after treatment. This implies that there is no need for personalized treatment allocation with regard to the CBTs under study for adolescents with ADHD. However, for a subgroup with comorbid anxiety symptoms but low depression PML clearly appears the treatment of preference.

Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is a childhood neurodevelopmental disorder that affects approximately 5–10% of the children (Faraone, Sergeant, Gillberg, & Biederman, 2003; Polanczyk, Silva de Lima, Lessa Horta, Biederman, & Rohde, 2007). Children with the disorder have symptoms of inattention, symptoms of hyperactivity/impulsivity or both, resulting in three subtypes of the disorder (American Psychiatric Association, 2000): the inattentive subtype, the hyperactive/impulsive subtype and the combined subtype. The trajectories of these symptom-clusters differ, as symptoms of inattention usually persist during development, while symptoms of hyperactivity/impulsivity often diminish over time. About 65% of children with ADHD still meet criteria for the diagnosis in adolescence (Biederman, Mick, & Faraone, 2000; Hill & Schoener, 1996). In addition, about 62% of children with ADHD below 19 years have at least one comorbid disorder, whereas 34% even have two or more comorbid conditions (Yoshimasu et al., 2012). Prevalent comorbid disorders are oppositional defiant disorder (ODD), conduct disorder (CD), anxiety- and mood-disorders, adjustment disorder and substance use disorder (Yoshimasu et al.). Also, more boys than girls have ADHD (3:1 respectively; Szatmari, Offord, & Boyle, 1989; Tuithof, Ten Have, Dorsselaer, & De Graaf, 2014) and girls appear to have less severe inattention, hyperactivity and impulsivity, but greater intellectual impairments than boys with ADHD (Gershon, 2002). In addition, girls tend to have more internalizing comorbid disorders than boys, whilst boys with ADHD are at higher risk for externalizing psychiatric comorbidities than girls (Yoshimasu et al., 2012). Taken together, the group of individuals with ADHD is a heterogeneous one.

This heterogeneity makes it is unlikely for one treatment to fit all and increases the need for personalized treatment. Randomized clinical trials (RCTs) can be valuable in revealing moderators of therapeutic change, investigating for whom what treatment is most effective (Kraemer, Wilson, Fairburn, & Agras, 2002). For example, in a recent RCT the effectiveness of two cognitive behavioral therapies (CBTs) for adolescents with ADHD was investigated (see **chapter 3** and **4**). Even though some are critical of the use of CBT in children with ADHD (Sibley, Kuriyan, Evans, Waxmonsky, & Smith, 2014), others presume that adolescents might have enough cognitive capacity to benefit from CBT (Antshel & Olszewski, 2014). Because adolescents with ADHD have planning problems in daily life that can cause im-

pairment in school, family- and social functioning (Abikoff et al., 2013) and evidence-based nonpharmacological treatments for adolescents with ADHD are lacking (Evans, Owens, & Bunford, 2014), a CBT was developed focusing on planning skills: Plan My Life (PML; Kuin, Boyer, & Van der Oord, 2013; see **appendix**, p. 151). In PML, every session a fixed, planning skills focused, subject and strategy is discussed and trained (e.g., a to-do list). In a multi-site RCT (*n*=16 sites, *n*=56 therapists, see **chapter 3**), this treatment was compared to a control CBT, without the proposed active element of enhancing planning skills: a solution-focused treatment (SFT; Boyer, Oberink, Kuin, & Van der Oord, 2014). Both PML and SFT are individual, manualized treatments consisting of 8 adolescent sessions and 2 parental sessions. Whereas in PML every week planning skills are actively learned by discussing a fixed subject, in SFT the adolescent/parent chooses a problem that is discussed using fixed questions in a solution focused manner, to lead the adolescent to a solution for the problem. To reduce drop-out, motivational interviewing is integrated within both treatments.

Pre-, post- and three-month follow-up data were gathered in 159 adolescents with ADHD (12 to 17 years), with parent-rated ADHD symptoms and planning problems as primary outcomes. Results showed a significant improvement of primary outcomes as well as comorbid symptoms, functioning and impairment (with large effect sizes) from pre- to posttest with maintenance of effects to three months after treatment on most measures, also when controlling for medication use. In addition, 15.2% of adolescents showed normalization of functioning at follow-up. However, only marginally significant treatment differences were found, in favor of PML: PML showed more reduction of parent-rated planning problems compared to SFT, and higher treatment satisfaction of parents and therapists. And, on the primary outcome ADHD symptoms there was no difference between both treatments (see **chapter 3**). Also, due to the lack of an adequate control-group like a waitlist or a treatment as usual group, effectiveness of both treatments could not be proven. The lack of differences in treatment outcome in this RCT could be due to heterogeneity of treatment effects in different subgroups of adolescents with ADHD: When for some subgroups of adolescents one treatment is better than the other, while for other subgroups the reverse is true, this can result in comparable mean outcomes for the two treatments.

This phenomenon, where the optimal treatment for one subgroup is different from that for another subgroup, is referred to as a *qualitative* treatment-

subgroup interaction, as opposed to a *quantitative* treatment-subgroup interaction, where the optimal treatment is the same in all subgroups but the size of the between-treatment difference differs across subgroups (Byar, 1985; Dusseldorp & Van Mechelen, 2014). In both cases, the patient characteristic(s) defining the subgroups in question are called moderators of treatment effect (Kraemer et al., 2002). For example, Ogrodniczuk, Piper, Joyce, and McCallum (2001) compared two forms of short-term individual psychotherapy (interpretative, supportive) for individuals suffering from depression. The authors found that for males interpretative therapy outperforms supportive therapy, whereas for females the reverse is true. This is a classic example of a *qualitative* treatment-subgroup interaction, where gender acts as a moderator of treatment effect. In the case of a *quantitative* treatment-subgroup interaction, the same treatment, for example interpretative therapy, would have outperformed supportive therapy in both genders, but for males the difference between the two treatments would have been larger than for females.

In the present study, a new and innovative statistical technique is used to explore these *qualitative* interactions: Qaalitative Interaction Trees (QUINT; Dusseldorp & Van Mechelen, 2014). Earlier work on the detection of treatment-subgroup interactions primarily refers to situations in which clear a priori hypotheses exist about which subgroups of clients are involved in the interactions, or situations that involve a small number of potential moderator variables only. Examples include factorial analyses of variance (ANOVA), with one factor pertaining to treatment methods and another one to subgroups (Shaffer, 1991), and regression analyses with suitable interaction terms being included in the regression model (see e.g., Dixon & Simon, 1991; Hayward, Kent, Vijan, & Hofer, 2006). In contrast, QUINT does not require a priori hypotheses or a limited number of potential moderator variables but rather induces subgroups involved in treatment-subgroup interactions during the actual data analysis. The goal of QUINT is to find a partition of the total group of adolescents with ADHD, based on their pretreatment characteristics, into two or three mutually exclusive subgroups that are characterized as follows: In the first subgroup PML outperforms SFT; in the second subgroup the reverse is true; and in the third (optional) subgroup, the adolescents assigned to PML would show more or less the same outcome as the adolescents assigned to SFT. Each of the subgroups is defined by a combination of one or several dichotomized patient pretreatment characteristics

(for which associated cut-off scores are being provided by QUINT). For example, a combination of pretreatment scores for intellectual impairment and externalization larger than some critical values may imply that SFT harvests better results than PML, whereas for other subgroups the reverse holds true. As such, the QUINT results may have straightforward implications for optimal treatment assignment strategies to support healthcare decision makers of adolescents with ADHD. For an example of a recent successful application of QUINT in the field of psychotherapy research, see Doove, Van Deun, Dusseldorp, and Van Mechelen (2015).

The present study is the first to investigate *qualitative* interactions in treatments for adolescents with ADHD, comparing two types of CBT. To our knowledge, there are no studies on qualitative interactions in treatment of ADHD and no treatment moderation studies have been conducted in adolescents with ADHD. However, two research groups *have* investigated a wide range of treatment moderators (child and family characteristics) in younger children with ADHD (up to 12 years of age), using mainly traditional moderation analyses such as ANOVA in which clear a priori hypotheses about potential moderators are required (Jensen et al., 2001; March et al., 2000; MTA cooperative group, 1999b; Owens et al., 2003; Rieppi et al., 2002; Van den Hoofdakker et al., 2010, 2012, 2014). Although solely based on behavioral (parent) treatment of children with ADHD as compared to other types of treatment (i.e., medication, combined treatment or regular care), the following characteristics appeared to positively influence the effects of behavioral (parent) treatment on ADHD-symptoms: having no or one single comorbid disorder (Van den Hoofdakker, 2010), in particular anxiety (Jensen et al., 2001; MTA cooperative group, 1999b), being older of age, having a mother with high parenting self-efficacy (Van den Hoofdakker, 2010), or having no or one single DAT1 10-repeat allele (Van den Hoofdakker, 2012). Due to a lack of treatment moderation studies in adolescents with ADHD, and taking the heterogeneity of our sample into consideration, we measured a broad range of pretreatment patient characteristics to explore as potentially relevant moderator variables in the present study: age, gender, full-scale IQ, medication use, parental education, ADHD subtype, ADHD severity, comorbid ODD/CD-, depressive- and anxiety symptoms, and overall impairment. In addition, planning problems were included as a potential moderator, as one treatment aimed at enhancing planning skills (PML), whereas the other treatment did not (SFT). Reasons for including moderators in this

study were 1) to be comparable to previous studies (e.g., gender, medication use, IQ, age, parental education, comorbid internalizing disorders like depressive and/or anxiety, externalizing disorders like ODD and/or CD), and 2) because one can presume one CBT to be a better fit that the other. For example, one could argue that older adolescents, adolescents with higher IQs, or less ADHD symptoms or impairment or adolescents with parents who have higher education, have more cognitive capacity or support at home, and therefore do better in a more open treatment like SFT, while younger adolescents, or with lower IQs, more severe ADHD or impairment and less educated parents fare better with a more structured treatment like PML. Outcome measures in this study were the primary outcomes from our RCT: parent-rated ADHD-symptoms and planning problems (see **chapter 3**).

To our knowledge, previous studies on moderation of treatment effects in children with ADHD have focused mainly on moderators of short-term effects of treatment between pre- and posttest (Jensen et al., 2001; March et al., 2000; MTA cooperative group, 1999b; Owens et al., 2003; Rieppi et al., 2002; Van den Hoofdakker et al., 2010, 2012, 2014; *but* Jensen et al., 2007). However, trajectories of treatment effects may differ between phases of treatment. The best-known example thereof is the MTA-study, where children who had received medication only or in combination with behavior therapy improved more from pre- to posttest than children who had received behavior therapy only or community care (MTA cooperative group, 1999). However, 10 months after treatment these differences had diminished and approximately 2 years after treatment, groups did not differ significantly on any measure, an effect that remained stable until 8 years after pretest (Jensen et al., 2007; Molina et al., 2009; MTA cooperative group, 2004). Also, even though in the MTA-study moderators of treatment effect were found from pre- to posttest (Jensen et al., 2001; March et al., 2000; MTA cooperative group, 1999b; Owens et al., 2003; Rieppi et al., 2002), no moderators of treatment effect were found from pretest to 36 months after treatment (Jensen et al., 2007).

Therefore, in the present study, apart from investigating qualitative moderators of treatment effects from pretest to follow-up, also qualitative moderators of the immediate treatment effect (from pretest to posttest) and of treatment effect consolidation (from posttest to follow-up three months after treatment) will be examined. More specifically, using QUINT, we investigated the following primary research question: Are qualitative treatment–

subgroup interactions present in the data of our RCT (from pretest to three months after treatment)? And, if so, which treatment is best for which subgroup of adolescents with ADHD? A secondary research question we investigated, was whether there are qualitative treatment-subgroup interactions in the two segments of the trajectory of therapeutic change from pretest tot follow-up, namely from pretest to posttest, and from posttest to follow-up? Given the lack of prior research on moderators of treatment effects in adolescents with ADHD, and in particular on moderators involved in qualitative treatment-subgroup interactions, we stated no specific expectations. When indeed qualitative treatment-subgroup interactions would be found from pretest to follow-up, this would be highly relevant for personalized medicine, as it would make allocation of adolescents with ADHD to specific treatments possible, resulting in higher treatment efficacy.

Methods

For a more detailed description of the sample, treatment content, procedures, research design, and approach to missing data, I refer to **chapter 3**.

Participants and procedure

All participants (n=159) were adolescents aged 12 to 17 years (M_{age}=14.4 years), who attended secondary school. Participants had received a prior DSM-IV-TR diagnosis of ADHD (American Psychiatric Association, 2000) by a child psychiatrist or certified psychologist, which was confirmed with the Diagnostic Interview Schedule for Children for DSM-IV parent version (DISC-IV; Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000). The DISC-IV is a structured diagnostic interview based on DSM-IV, which establishes ADHD group membership based on a diagnostic algorithm, including a check for the presence of cross-situational impairment. Participants had a full scale IQ (FSIQ) > 80 measured by the short version of the Dutch Wechsler Intelligence Scale for Children (WISC-III; Kort et al., 2005; Sattler, 2001). During treatment 124 adolescents used psychostimulants (no atomoxetine) and were requested to keep the dose stable until posttest. Exclusion criteria were: having a comorbid autism spectrum disorder, depression with suicidal ideations, acute familial crisis, CD or predominant addiction.

Adolescents applied for the study in one of sixteen participating mental

health care centers. After informed consent was obtained, pretest took place and adolescents were randomly assigned to either PML (n=83) or SFT (n=76). To control for therapist effects, the 56 participating therapists (who all at least had a master degree in psychology) provided both treatments. In both treatments attrition was low (n_{PML}=4, n_{SFT}=4) and the length of treatment was comparable (M_{PML}=9.3 weeks, SD_{PML}=2.6; M_{SFT}=9.1 weeks, SD_{SFT}=2.9). Adherence was high in both treatments and no treatment contamination was found. Posttest took place within a week after treatment and follow-up test approximately three months after treatment. All assessments and treatments took place in the same mental health care center where the participant applied for treatment outpatient, and were conducted by blinded research assistants. The Ethics Committee of the University of Amsterdam approved this study (2010-KP-1079).

Instruments

Treatments

Similarities. In our RCT we compared two treatments: PML and SFT (for a more detailed description of both treatments see **chapter 3** and **appendix**). Both are individual treatments that consist of 8 adolescent sessions of 45-60 minutes and 2 parental sessions. In both treatments the adolescent and therapist work together from a workbook. In both treatments MI plays an important role: the attitude of the therapist as well as wording in the workbook are based on MI. For example, the adolescents choose their own treatment goals (not those of their parents or teachers), which were written down and were always visible during treatment. Also, assignments were not presented as 'homework', but were formulated as an experiment for the upcoming week. In that way, the adolescent was free to choose a treatment strategy that fitted his/her life and the strategy could be adjusted according to their experiences.

Plan My Life (PML; Kuin et al., 2013, see appendix) is a CBT in which every session a fixed subject is discussed (e.g. using a daily planner or using a to do list). Within every subject possible planning- and organization strategies are presented of which the adolescent can compose their own strategy. The chosen strategy is than posited as an experiment that he/she will try the following week. Whenever needed, negative thoughts about the new strategy are challenged and helping thoughts formulated. All session strategies that have been tried the past week are discussed, including successes, possible

room for improvement, and associated cognitions.

In the *Solution Focused Treatment* (SFT; Boyer et al., 2014) every session the adolescent discusses a problem he/she encounters. Following fixed questions, the adolescent is guided towards a solution for the posited problem. The fixed questions are: 1) What is the subject you chose? Describe the situation, 2) How is the situation, as it is now, a problem for you?, 3) How would you like it to be?, 4) What are solutions you used in the past and what are other possible solutions to the problem?, 5) Does the situation, as it is now, have advantages?, 6) Would you like to change the situation now/later/not at all, 7) If you choose to change, what is your plan? If you choose to change later or not at all, what are your considerations (pros and cons)? In this treatment the therapist does not tell the adolescent or parent what to do and does not teach them a new skill, but guides them towards a solution using the fixed questions.

Differences between treatments. The important difference between both treatments lies in the content. Whereas in PML every week a fixed subject is discussed regarding planning and organizing and planning skills are actively learned, in SFT the adolescent/parent has to choose the subjects themselves and is guided to his/her own solution.

Outcome measures

ADHD symptoms of the adolescents are measured using the Disruptive Behavior Disorder rating scale parent version (DBD; Oosterlaan, Scheres, Antrop, Roeyers, & Sergeant, 2000; Pelham, Gnagy, Greenslade, & Milich, 1992). The DBD contains four scales composed of the DSM-IV criteria for ADHD Inattention, ADHD Hyperactivity/Impulsivity, Oppositional Defiant Disorder, and Conduct Disorder. The ADHD symptoms scale is calculated summarizing the ADHD Inattention and the ADHD Hyperactivity/Impulsivity scales (18 items, scores range from 0 to 54). Analyses are conducted on raw scores, with a higher score indicating more symptoms of ADHD.

Planning problems of the adolescent are rated by parents using the Dutch Behavior Rating Inventory of Executive Function (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000; Smidts & Huizinga, 2009): A normative behavioral rating scale for children 5 to 18 years old, designed to elicit everyday planning behavior as observed by the parents in natural everyday environ-

ments. The Plan/Organize scale measures the child's capacity to anticipate future events, set goals, develop appropriate steps to carry out associated tasks or actions, and manage current and future-oriented task demands (12 items, scores range from 12 to 36). Analyses are conducted on raw scores, with higher scores indicating more planning problems.

Outcome measures are assessed at pretest one week before treatment, at posttest within a week after treatment and follow-up test approximately three months after treatment. To investigate our primary research question, 'Are qualitative treatment–subgroup interactions present in the data of our RCT?', analyses were conducted on the difference score between pretest and three-month follow-up test on both the DBD and the BRIEF questionnaire. To answer our secondary research question, 'Do the two treatment alternatives have a different effect on the trajectories of therapeutic change of ADHD-symptoms and planning problems for different subgroups of adolescents?', we analyzed the difference between pre- and posttest on both the DBD and the BRIEF and the difference between post- and follow-up test on both the DBD and the BRIEF. The effect found between pre- and posttest is considered to be the short-term effect of treatment. The effect between post- and three-month follow-up test is considered to reflect the consolidation of treatment.

Patient Characteristics

In total, 12 patient characteristics were assessed as potential moderators involved in qualitative treatment-subgroup interactions. Demographic variables included were gender, age, and medication use. Age was evaluated as a continuous variable while gender and medication use (yes/no) were treated categorically. Also, the average education level of both parents was assessed on a 4-point Likert-scale, with 1 representing the lowest educational level and 4 the highest. Parental education was treated ordinal. Several additional baseline adolescent characteristics were also included in the analyses as moderators:

Full Scale IQ was measured by two subtests of the Dutch Wechsler Intelligence Scale for Children (WISC-III-NL; Kort et al., 2005) that correlate highly with the Total IQ: Vocabulary and Block Design (FSIQ; Sattler, 2001). The variable full scale IQ was treated continuous.

ADHD symptoms of the adolescent were measured using the DBD rating scale parent version (Oosterlaan et al., 2000; Pelham et al., 1992). The DBD

contains four scales composed of the DSM-IV criteria for ADHD Inattention, ADHD Hyperactivity/Impulsivity, Oppositional Defiant Disorder, and Conduct Disorder. The ADHD symptoms scale is calculated, summarizing the ADHD Inattention and the ADHD Hyperactivity/Impulsivity scales, with a higher score indicating more symptoms of ADHD and was thus treated continuous.

ADHD subtype was assessed by the DISC-IV, a structured diagnostic interview, that establishes ADHD group membership and subtype status based on a diagnostic algorithm, including a check of cross-situational impairment (Shaffer et al., 2000). Based on the DISC-IV the adolescents were classified into the inattentive subtype, the hyperactive/impulsive subtype and the combined subtype. This variable was treated categorically.

Planning problems of the adolescent were rated by parents using the BRIEF (Gioia et al., 2000; Smidts & Huizinga, 2009). This is a normative behavioral rating scale for children 5 to 18 years old, designed to elicit everyday EF as observed by the parents in natural everyday environments. In this study the Plan/Organize was used: higher scores on this subscale indicates more planning problems, and was therefore treated continuous.

Depressive symptoms were measured using the self-reported Child Depression Inventory (CDI; Sitarenios & Kovacs, 1999; Timbremont & Braet, 2002). Scores range from 0 to 54, in which higher scores indicate more depressive symptoms, and was therefore treated continuous.

Anxiety symptoms were evaluated using the Screen for Child Anxiety Related Emotional Disorders (SCARED; Birmaher et al., 1997; Muris, Bodden, Hale, Birmaher, & Mayer, 2007). Adolescents filled in 69 items, which added up to a total anxiety score, ranging from 0 to 138: Higher scores indicate more anxiety symptoms, and was thus treated continuous.

ODD/CD symptoms were measured with the 24 items of the ODD and CD scales of the parent rated DBD (Oosterlaan et al., 2000; Pelham et al., 1992; for description see outcome measure ADHD symptoms). Scores range from 0 to 72, in which higher scores indicate more externalizing symptoms, and was therefore treated continuous.

General Impairment was measured using the Impairment Rating Scale (IRS; Fabiano & Pelham, 2002). Parents answered six questions on a scale from 0 to 10, resulting in a total score ranging from 0 to 60, with higher scores indicating high general impairment, and was thus treated continuous.

Table 5.1: Pre-treatment adolescent characteristics.

	All (n=159)		PML (n=83)		SFT (n=76)		
	M/N	SD/%	M/N	SD/%	M/N	SD/%	Comparison
Gender (n boys)	117	73.6	63	75.9	54	71.1	$\chi^2(1)$=.48, φ=.06
Age in years	14.4	1.2	14.4	1.2	14.4	1.3	$t(157)$=.19, d=.00
FSIQ	103.4	11.8	102.5	11.6	104.3	11.9	$t(157)$=−.97, d=.15
Parental education [a]	3.1	0.8	3.2	0.7	3.1	0.8	$t(146)$=.55, d=.13
n Medication	124	78.0	62	74.7	62	81.6	$\chi^2(1)$=1.09, φ=.08
ADHD Subtype							$\chi^2(2)$=5.06, φ=.18
Inattentive	112	70.4	62	74.7	50	65.8	
Hyperactive/impulsive	8	5.0	6	7.2	2	2.6	
Combined	39	24.5	15	18.1	24	31.6	
Other measures							
ADHD symptoms	25.3	9.3	25.6	9.5	25.0	9.2	$t(157)$=.40, d=.06
Planning problems	28.2	4.4	28.0	4.5	28.4	4.4	$t(157)$=−.54, d=.09
ODD/CD symptoms	6.6	5.1	6.8	5.3	6.5	5.0	$t(157)$=.37, d=.06
Depression symptoms	10.0	6.1	10.1	6.5	10.0	5.7	$t(157)$=.13, d=.02
Anxiety symptoms	25.3	9.3	27.0	19.4	24.1	18.3	$t(157)$=.99, d=.15
Overall Impairment	31.5	11.5	32.3	11.7	30.6	11.3	$t(157)$=.93, d=.15

ADHD: Attention Deficit Hyperactivity Disorder; CD: Conduct Disorder; FSIQ: Full Scale IQ; ODD, Oppositional Defiant Disorder; PML: Plan My Life; SFT: Solution Focused Treatment. *a* On this variable data are missing. * p<0.05; Cramer's φ effect size: 0.10 is small, .30 is medium, .50 is large; Cohen's *d* effect size: .20 is small, .50 medium, .80 large.

In the analyses, raw scores were used. Table 5.1 shows the adolescents' scores on all variables; independent *t*-tests and chi-squared tests showed no group differences on all patient characteristics between both treatments.

Statistical analyses

To address our research questions, we subjected the data to an analysis with the recently developed method QUINT (Dusseldorp & Van Mechelen, 2014). We will first elaborate on the conceptual basis of QUINT and later on in this section explain how we set out the analysis of our data using this method. As mentioned before, the goal of QUINT is to find the best partition of the total group of adolescents with ADHD, based on their pre-treatment characteristics, into two or three mutually exclusive subgroups that are characterized as follows: In the first subgroup, $\wp_1$, the adolescents assigned to treatment PML would show a clearly better outcome than the adolescents assigned to SFT; in the second subgroup, $\wp_2$, the reverse is true; in the third (optional)

subgroup, $\wp_3$, the adolescents assigned to PML would show more or less the same outcome as the adolescents assigned to SFT. The subgroups may comprise one or several types of adolescents as defined by different (combinations of) dichotomized pre-treatment characteristics (moderators).

QUINT is looking for an optimal partition of the total group of adolescents so that the qualitative treatment-subgroup interaction that is related to that partition has the largest possible practical significance. To achieve this, two conditions with regard to the subgroups $\wp_1$ and $\wp_2$ need to be satisfied: (a) In both subgroups the difference in outcome between PML and SFT should be substantial, and (b) each of the two subgroups should comprise a sufficient number of adolescents. QUINT uses a weighted compound criterion that implies that these two conditions are optimized simultaneously. The difference in outcome between the PML and SFT treatment conditions included in condition (a) can be formalized in terms of either a difference in treatment means or a treatment effect size (Cohen's *d*; Cohen, 1988).

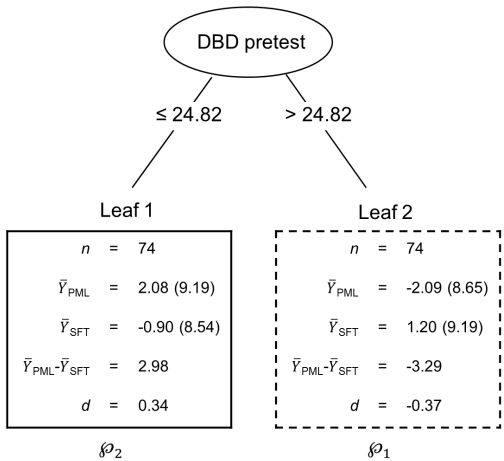
To optimize this criterion, QUINT uses a stepwise tree-building algorithm. We refer to King and Resick (2014) for a conceptual introduction to tree methods in psychological treatment research, and note that QUINT is an innovative member of this family of methods that is custom-made to identify qualitative treatment-subgroup interactions. The QUINT algorithm sequentially splits the total group of adolescents into subgroups, while optimizing in each step the weighted compound criterion mentioned above, with the resulting series of splits being represented by a tree structure like Figure 5.1 (which will be further discussed in the Results section, p. 110). The QUINT procedure uses three types of criteria to stop the tree building process: Firstly, after the first split it tests the presence of a qualitative interaction on the basis of a so-called qualitative interaction condition, which reads that in each of the two leaves, the absolute value of the treatment effect size exceeds a critical minimum effect size value (d_{min}). If QUINT does detect a qualitative treatment-subgroup interaction, the first split will be executed and the stepwise binary splitting procedure is continued. Secondly, the algorithm stops if no longer a split can be found that implies a higher criterion value than in the previous step. Thirdly, QUINT takes into account some additional stopping criteria including a maximum value for the number of leaves, and the fact that each leaf should contain a minimum number of adolescents assigned to PML and SFT.

The tree growing procedure may result in a large tree with a high crite-

rion value for the data at hand, which may not be replicable with future data. QUINT controls for this so-called overfitting by a pruning procedure, which implies that the maximal tree is pruned back to an optimal subtree. This procedure yields a sequence of nested subtrees, for which the number of leaves (L) can be considered a complexity parameter. For each of these subtrees, QUINT has determined a criterion value. Yet, these values are positively biased since they have been determined on the basis of the same data as the ones that were used to build the tree. To overcome this problem, QUINT relies on a bootstrap procedure that yields an estimation of the biases, which further allows calculation of bias-corrected criterion values. The final tree selected by QUINT then is the one with the highest bias-corrected criterion value. Otherwise, QUINT also allows for a bootstrap-based bias correction procedure for the differential treatment effect sizes in the leaves of the finally selected tree, which may give insight into the generalizability of the QUINT solution. For a formalization and detailed description of the above, we refer to Dusseldorp & Van Mechelen (2014). To analyze RCTs with QUINT, the R-package quint has been developed by Dusseldorp, Van Mechelen & Doove (2015). In this package some freedom is left to the user (e.g., regarding parameters related to the qualitative interaction condition and tree complexity). The package can be freely downloaded from the Comprehensive R Archive Network (CRAN) and can handle continuous and dichotomous background characteristics. For the analyses in the present paper, we used a slightly extended version of this package that can also handle polytomous categorical background characteristics. This extension can be obtained from the authors.

In this study, the QUINT analyses were performed using the default effect size criterion. When using this criterion, Dusseldorp & Van Mechelen (2014) propose to set the weight for the first constituent of the criterion equal to $1/\log(1+d_{max})$, where d_{max} is the realistic maximum for treatment effect size. Based on a recent review and meta-analyses on psychological interventions for ADHD (Daley et al., 2014; Sibley, Kuriyan, Evans, Waxmonsky, & Smith, 2014; Sonuga-Barke et al., 2013), we set d_{max} equal to 1. As value for the weight of the second constituent of the optimization criterion, Dusseldorp & Van Mechelen (2014) propose $1/\log(0.5n)$, where n is the sample size. For reliable estimation of the mean outcome in a treatment group, we set the minimum number of adolescents assigned to PML and SFT in each leaf equal to 20. We used the default values for the critical minimum value in the quali-

Figure 5.1: Result of the application of QUINT on posttest-to-follow-up difference in parent-rated ADHD symptoms (on the Disruptive Behavior Disorder rating scale parent version; DBD)



Smaller outcome values are preferred as these refer to more reduction of parent-rated ADHD symptoms. The ellipsis in the figure represents the root node, which corresponds to the complete group of clients. This ellipsis contains the split variable, with below it the corresponding split point. The rectangles represent the leaves of the tree, that is, the final subgroups of adolescents; each rectangle contains the sample size of the corresponding subgroup, the outcome means (and standard deviations) for the Plan My Life (PML) and Solution Focused Treatment (SFT) conditions $\bar{Y}_{PML}$ and $\bar{Y}_{SFT}$, the uncorrected difference in means $\bar{Y}_{PML}-\bar{Y}_{SFT}$, and the uncorrected corresponding effect size d. Assignment of the leaves to the partition classes is represented by $\wp_1$ and $\wp_2$.

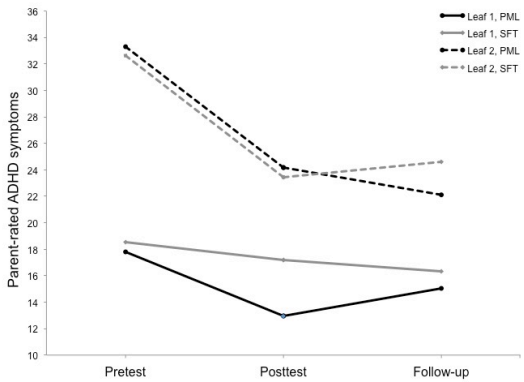
tative interaction criterion for the absolute value of the standardized mean difference in treatment outcome (d_{min}=0.30, based on results of a simulation study by Dusseldorp & Van Mechelen, 2014), and the maximum number of leaves. Lastly, we set the number of bootstrap samples in the pruning procedure equal to 200, which is a relatively large number leading to more stable results than the default of 25 bootstrap samples.

Results

ADHD symptoms

To answer the first research question, 'Are qualitative treatment–subgroup interactions present in the difference scores between pretest and three-month follow-up?', we analyzed the difference scores for parent-rated ADHD

Figure 5.2: Evolution of ADHD symptoms from pretest to follow-up up for the two subgroups resulting from the application of QUINT on posttest-to-follow-up difference in parent-rated ADHD symptoms (DBD).



Leaf 1 is represented in solid lines, and Leaf 2 in dashed lines. The Plan My Life and Solution Focused Treatment groups within each leaf are represented in black (PML) and grey lines (SFT).

(DBD). The test of the qualitative interaction condition after the first split (i.e., QUINT's first stopping criterion, see Statistical analyses section) revealed that no qualitative treatment-subgroup interaction is present in the data. This result implies that there is no need for personalized treatment allocation for adolescents with ADHD when it comes to ADHD symptoms.

Next, we addressed for parent-rated ADHD the second research question on qualitative treatment-subgroup interactions in the two segments of the trajectory of therapeutic change from pretest tot follow-up, namely from pretest to posttest, and from posttest to follow-up. With regard to the short-term effect of treatment, that is, the change in ADHD-symptoms from pre- to posttest, QUINT again revealed that no qualitative treatment-subgroup interaction is present. However, regarding the consolidation of treatment (in other words, the change in ADHD-symptoms from post- to three-month follow-up test), QUINT *did* reveal a qualitative treatment-subgroup interaction in the data: The analysis resulted in a tree with two leaves (shown in Figure 5.1). These results show that adolescents with less ADHD symptoms (DBD pretest score $\leqslant 24.82$), should preferably receive SFT over PML (Figure 5.1, Leaf 1); indeed, for these adolescents, improvement on ADHD symptoms between post- and follow-up test was on average 2.98 points higher when as-

signed to SFT compared to PML (d=0.34). On the other hand, adolescents with more ADHD symptoms at baseline (DBD pretest score > 24.82) take more advantage from PML compared to SFT (Figure 5.1, Leaf 2). For these adolescents, improvement on ADHD symptoms was on average 3.29 points higher when assigned to PML compared to SFT (d=−0.37). After applying the bias-correction procedure to the two leaves, we found that the improvement on ADHD symptoms for the adolescents in Leaf 1 was on average 0.54 points higher when assigned to SFT compared to PML (d=0.05), and that the improvement on ADHD symptoms for the adolescents in Leaf 2 was on average 1.17 points higher when assigned to PML compared to SFT (d=−0.11). The values of these bias-corrected effect sizes imply that the detected qualitative treatment-subgroup interaction is small.

In order to more clearly understand the subgroup-specific trajectories across all three assessments, we examined the evolution of the mean outcome values in the two leaves across the entire time span from pretest to follow-up. Figure 5.2 shows this evolution for the ADHD symptoms for the induced subgroups of adolescents, while further drawing a distinction between the two treatment conditions. The figure suggests that whereas from posttest to follow-up the adolescents in Leaf 1 benefit relatively more from SFT and the adolescents in Leaf 2 benefit relatively more from PML, this does neither apply to the time from pretest to posttest, nor for the entire time span from pretest to follow-up. From pretest to posttest the subgroups rather seem to be involved in a *quantitative* treatment-subgroup interaction (with PML outperforming SFT in both subgroups, but with the advantage of PML being larger for the adolescents in Leaf 1). This interaction, however, appeared not to be significant in a post-hoc factorial analysis of variance on the difference score between pre- and posttest (with a first factor pertaining to treatment and a second one to the leaves, $F(1,144)=1.84$, $p=.177$, $\omega_p^2=.006$. Furthermore, the relatively larger advantage of PML for the adolescents in Leaf 1 in the evolution from pretest to posttest, is undone by the qualitative treatment-subgroup interaction in the evolution from posttest to follow-up (leading to a non-significant treatment-leaf interaction with regard to the evolution from pretest to follow-up, $F(1,144)=1.03$, $p=.311$, $\omega_p^2=.000$).

Planning problems

With regard to the first research question 'Are qualitative treatment–subgroup interactions present in the difference scores between pretest and

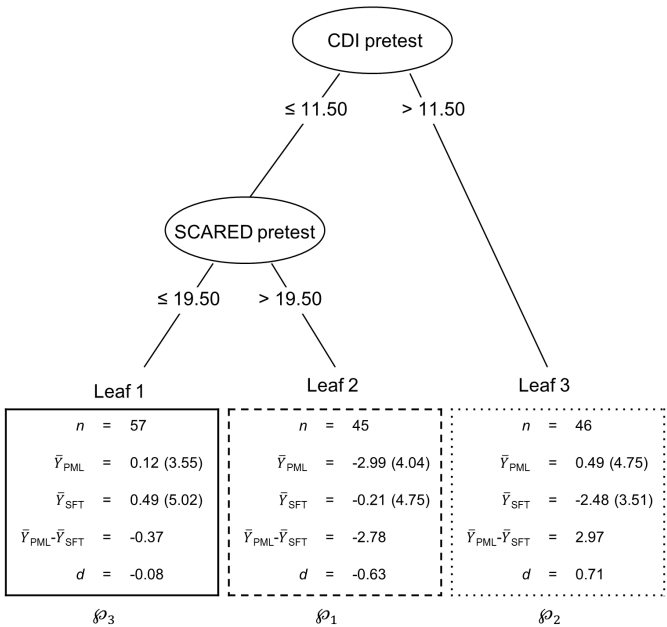
three-month follow-up?', we investigated these scores for parent-rated planning problems (BRIEF). QUINT revealed that no qualitative treatment-subgroup interaction is present in the data. This implies that there is no need for personalized treatment allocation for adolescents with ADHD when it comes to planning problems.

For the second research question on qualitative treatment-subgroup interactions in the two segments of the trajectory of therapeutic change from pretest tot follow-up, difference scores between pre-and posttest, and between posttest and follow-up were subjected to QUINT analyses. With regard to the difference scores between pre- and posttest, QUINT revealed that no qualitative treatment-subgroup interaction is present, whereas for the difference scores between posttest and follow-up a qualitative interaction *is* present in the data. Regarding the latter, QUINT constructed a tree with three leaves. The pruning procedure indicated that this was also the optimal tree size.

The structure of the tree is shown in Figure 5.3. From this figure it appears that adolescents in Leaf 3 with more depressive symptoms at baseline (CDI pretest score 11.5) take more advantage from SFT compared to PML: For these adolescents, improvement on planning problems was on average 2.97 points higher when assigned to SFT compared to PML (d=0.71). On the other hand, adolescents in Leaf 2 with less depressive symptoms (CDI pretest score $\leqslant$ 11.5), and more anxiety symptoms (SCARED pretest score > 19.50) improvement on planning problems was on average 2.78 points higher when assigned to PML compared to SFT (d=−0.63). Finally, for adolescents in Leaf 1 with low depression and low anxiety levels, effects of both treatments are comparable. After applying the bias-correction procedure to the two leaves with the most extreme differential treatment effects, we found that the improvement on planning problems for the adolescents in Leaf 2 was on average 1.30 points higher when assigned to PML compared to SFT (d=−0.27), and that the improvement on planning problems for the adolescents in Leaf 3 was on average 1.73 points higher when assigned to SFT compared to PML (d=0.36). These bias-corrected values imply that the effect size of the detected qualitative treatment-subgroup interaction is between small and medium.

In order to more clearly understand the subgroup-specific trajectories across all three assessments, we again examined the evolution of the mean outcome values in the three leaves across the entire time span. This evolution is shown in Figure 5.4 for the parent-rated BRIEF for the three differ-

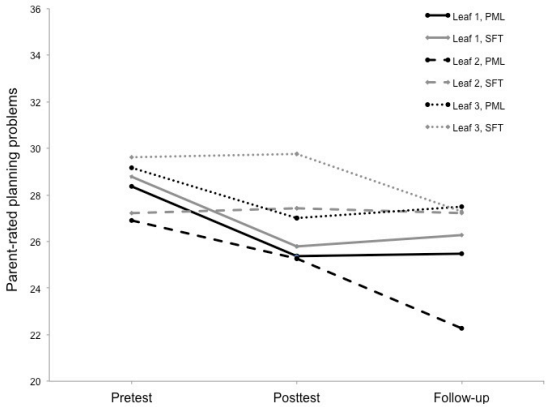
Figure 5.3: Result of the application of QUINT on posttest-to-follow-up difference in parent-rated planning problems (BRIEF)



Smaller outcome values are preferred as these refer to more reduction of parent-rated planning problems. The ellipses in the figure represent the internal nodes containing the split variables (CDI, Child Depression Invertory; SCARED, Screen for Child Anxiety Related Emotional Disorders), with below each ellipsis the corresponding split point. The upper ellipsis represents the root node corresponding to the complete group of clients. The rectangles represent the leaves of the tree, that is, the final subgroups of adolescents; the quantities contained in the rectangles are analogous to those in Figure 5.1. Assignment of the leaves to the partition classes is represented by $\wp_1$, $\wp_2$ and $\wp_3$.

ent subgroups, while further drawing a distinction between the two treatment conditions. Whereas from posttest to follow-up, the three subgroups are involved in a qualitative treatment-subgroup interaction, this is the case neither for the evolution from pretest to posttest nor for the evolution from pretest to follow-up. Rather, based on Figure 5.4, the three subgroups seem to be involved in a *quantitative* treatment-subgroup interaction with regard to the evolution from pretest to posttest, with PML outperforming SFT in leaves 2 and 3, and with no between-treatment difference in Leaf 1. In a post-hoc ANOVA, however, this interaction appeared not be significant, $F(2,142)=0.99$, $p=.375$, $\eta^2_p=.000$. With regard to the evolution from pretest to follow-up, Figure 5.4 also suggests the presence of a quantitative treatment-

Figure 5.4: Evolution of parent-rated planning problems from pretest to follow-up for the three subgroups resulting from the application of QUINT on posttest-to-follow-up difference in parent-rated planning problems (BRIEF)



Leaf 1 is represented in solid lines, Leaf 2 in dashed lines, and Leaf 3 in dotted lines. The Plan My Life and Solution Focused Treatment groups within each leaf are represented in black (PML) and grey lines (SFT)

subgroup interaction with PML outperforming SFT in Leaf 2, and with no between-treatment difference in leaves 1 and 3. In a post-hoc ANOVA, this interaction appeared to be significant, indeed, $F(2,142)=4.18$, $p=.017$, $\eta_p^2=.041$. Moreover, a post-hoc Tukey test confirmed that PML significantly outperformed SFT in Leaf 2 whereas the treatment groups did not significantly different from each other in Leaves 1 and 3. Importantly, this implies that adolescents with less depressive symptoms (CDI pretest score $\leqslant$ 11.5 and more anxiety symptoms (SCARED pretest score > 19.5) should preferably receive PML, while for other adolescents the effects of PML and SFT are comparable.

Discussion

This was the first study to investigate moderators involved in qualitative treatment-subgroup interactions in adolescents with ADHD, using an innovative statistical technique (QUINT). Our results show that no qualitative treatment–subgroup interactions were present in the data of our RCT for the evolution from pretest to three months after treatment, thereby answering our primary research question. This implies that, when comparing the two CBTs under study in adolescents with ADHD (one that aimed at enhanc-

ing planning problems and one that did not), there is no need for personalized treatment allocation when focusing on improvement of ADHD symptoms and planning problems as outcome variables. Remember further that in **chapter 3** for the same RCT a significant improvement in ADHD- and comorbid symptoms and general impairment of the adolescents was reported with large effect sizes, but only marginal differences in effect between treatments, in favor of PML. The latter lack of treatment differences is therefore not due to qualitative treatment-subgroup interactions.

Our second research question focused on the trajectories of therapeutic change for different subgroups of adolescents in both treatments, with a distinction between short-term effects of treatment and consolidation effects up to three months after treatment. Our results showed no qualitative treatment–subgroup interactions of the short-term effect of treatment. However, with regard to the consolidation of treatment from posttest to follow-up, qualitative treatment-subgroup interactions were found: In terms of improvement in parent-rated planning problems, adolescents with more comorbid anxiety symptoms but with low depressive symptoms fared better with PML, while adolescents with high depressive symptoms showed better results with SFT. Yet, when taking the evolution of symptoms of the identified subgroups into consideration from pretest to three-month follow-up, only the results in favor of PML were upheld. In addition, in terms of improvement in parent-rated ADHD symptoms, effects of PML were more positive for adolescents with more severe ADHD, while SFT showed more positive effects for adolescents with less symptoms of ADHD. Nevertheless, these results had small effect sizes only and these subgroups were not upheld in the course from pretest to follow-up. The most robust results were therefore, that adolescents with more anxiety symptoms but low levels of depression fared better with PML than with SFT with regard to improvement of parent-rated planning problems.

The cut-off scores that were provided by QUINT (CDI: cut-off point=11.50; SCARED: cut-off point=19.50), augment the clinical relevance of this study, as they may be directly useful for future treatment assignment. What is important to note, however, is that the subgroups generated by QUINT that are based on these cut-off scores, do not represent the normal or clinical range of these questionnaires. In other words, the interpretation of these scores depends on the norms of the corresponding questionnaire. For example, effects of PML were more positive than SFT for adolescents with a score on the

CDI below 11.50, but a higher score than 19.50 on the SCARED. Depending on the age and gender of the participant, a score of 19.50 on the SCARED can be considered a low to average score as compared to the norm (Muris et al., 2007). The cut-off point of the CDI can be considered average (Timbremont & Braet, 2002). All in all, this means that for adolescents with low to average depression, and with some to high anxiety (but not necessarily a clinical anxiety level), PML is the treatment of preference.

Our findings regarding the moderating role of comorbid anxiety symptoms, are in line with results of previous studies. For example, a non-randomized study on the effects of a planning focused CBT in adolescents with ADHD (Antshel, Faraone, & Gordon, 2014) also showed that a subgroup of adolescents with comorbid anxiety improved more in comparison to adolescents with ADHD only. Moreover, in the MTA-study the subgroup of children with comorbid anxiety showed more improvement when receiving behavioral treatment (Jensen et al., 2001; March et al., 2000; MTA cooperative group, 1999b). This might indicate a subgroup within ADHD that is particularly, or at least differently, sensitive to effects of (cognitive) behavioral treatment. The MTA Cooperative Group suggested that, as anxiety reduces after treatment of ADHD, at least some anxiety in individuals with ADHD might be attributed to the stress of ADHD-related problems (MTA Cooperative Group, 1999b). This is consistent with the idea that anxiety, in individuals with ADHD, is more strongly associated with negative affectivity and disruptive behavior than with fearfulness or phobic symptoms (March et al., 2000). If this would be the case, in our study one would expect anxiety not to be a moderator, because in our RCT two CBTs were compared and after both ADHD as well as anxiety had improved (see **chapter 3**). Because the more anxious group of adolescents with ADHD showed more positive results when receiving a CBT focusing on planning skills than when receiving a CBT without such an aim, this suggests that the anxiety experienced by adolescents with ADHD could be attributed to poorer executive functions (e.g., Airaksinen, Larsson, & Forsell, 2005; Eysenck, Derakshan, Santos, & Calvo, 2007; Tucker & Derryberry, 1992). This is supported by research showing increased working memory deficits and increased rates of sluggish cognitive tempo in children with ADHD and comorbid anxiety (Schatz & Rostain, 2006). However, this mediation hypothesis needs further testing.

An alternative explanation for the superior effects of PML as compared to SFT in the subgroup with low depression but higher anxiety (see Leaf

2, Figure 5.3 and 5.4), could be the use of medication for ADHD. Remember how adolescents could not change medication status (yes/no) or dose during treatment, but could only change these after posttest. When testing this hypothesis exploratory, we found that in this specific subgroup, adolescents receiving PML increased their dose of methylphenidate after posttest, while adolescents receiving SFT decreased their dose. While this is the case for this specific subgroup, this treatment difference between PML and SFT is not found for the whole group of adolescents in this study. This could for example imply that adolescents with low depression and higher anxiety that received PML, have become more motivated to adhere to medication as compared to SFT. However, as these results are only exploratory, also this hypothesis needs further testing.

Further, a notable result is that pretreatment characteristics may influence the effects of treatment differently during different phases of treatment. In this study the pretest to follow-up evolution of the found qualitative subgroups, clearly shows that the PML and SFT treatment alternatives have a different effect on the evolution of planning problems during treatment and during consolidation for different subgroups of adolescents. This supports the idea of two clearly separate treatments; both follow different trajectories that seem to converge to comparable endpoints after three months. Moreover, these results show that it is important to consider the phase of treatment when investigating moderator-effects: in this study differential effects were found in different phases of treatment. This may imply that different characteristics determine the effect of treatment during different phases. To our knowledge, previous studies on moderation of treatment effects in individuals with ADHD, have focused primarily on effects of treatment between pre- and posttest and may therefore miss potential moderators (Jensen et al., 2001; March et al., 2000; MTA cooperative group, 1999b; Owens et al., 2003; Rieppi et al., 2002; Van den Hoofdakker et al., 2010, 2012, 2014). We would therefore advise to take the phase of treatment into account when interpreting the effects of moderators and to also investigate moderators of follow-up treatment effects in future studies.

This study has several limitations. First, in this study several adolescent characteristics but no family characteristics, such as parenting self-efficacy, were included as potential moderators. Second, because all adolescents were included in a treatment condition, their parents who completed ratings scales (i.e. outcome measures) were potentially biased by expectancy effects.

Therefore, in future studies more objective tests assessing ADHD core symptoms are preferred. Finally, a broad concern related to the goal of finding subgroups involved in treatment-subgroup interactions is that, in general, a reliable detection of interactions requires larger samples than a reliable detection of main effects (Dusseldorp et al., 2015), and perhaps considerably larger than those enrolled in traditional clinical trials in the field of behavior therapy research. As a result, QUINT (as all post-hoc methods) brings with it the risk of inferential errors, like identifying interactions that cannot be replicated in follow-up studies. Tree-based methods are especially vulnerable at this point as they rely on a very large search space based on a very huge number of covariate split-point combinations and therefore need large sample sizes (Dusseldorp & Van Mechelen, 2014). Even though this study is the largest RCT in adolescents with ADHD to date, for QUINT to show replicable results, a larger sample size would be preferred. To be sure, we made use of a bias-corrected bootstrap procedure to prevent ourselves from overfitting the data at hand. However, Dusseldorp and Van Mechelen (2014) state that the safest way to deal with QUINT, especially in case of applications with relatively smaller sample sizes, as an exploratory tool to generate meaningful hypotheses. Further, these hypotheses should be tested in follow-up confirmatory research with new RCTs that make use of a stratified sampling scheme in which the strata are constructed on the basis of the splitting variables and split points as identified by QUINT (Dusseldorp & Van Mechelen, 2014).

Taken together, this study has several clinical implications. Our RCT showed improvements with large effect sizes for both PML and SFT, but without between-treatment differences on most measures (except for parent-rated planning problems and treatment satisfaction of parents and therapists; see **chapter 3**). The present study supplements these findings by showing that there were no significant moderators of qualitative treatment-subgroup interactions of therapeutic change from pretest to three months after treatment: Both treatments harvest comparable results. There was however a *quantitative* interaction present in the data, that showed that for adolescents with ADHD with comorbid anxiety symptoms in combination with low levels of comorbid depression, PML appears to be the treatment of preference. Of additional clinical relevance are the cut-off points generated by QUINT, with which clinicians can directly allocate adolescents to the most effective treatment. All in all, personalized treatment assignment appears

unnecessary, as the effects of two CBTs are comparable for all adolescents with ADHD, except for a subgroup of adolescents for whom PML is the treatment of preference.