UiT
THE ARCTIC
UNIVERSITY
OF NORWAY

Faculty of Health Sciences
Department of Psychology

Change and predictors of rate of change in the treatment of children and adolescents with emotional disorders

A naturalistic observational study in two child and adolescent mental health services in Norway

—

Toril Sørheim Nilsen

A dissertation for the degree of Philosophiae Doctor –April 2017

Change and predictors of rate of change in the treatment of children and adolescents with emotional disorders

A naturalistic observational study in two child and adolescent mental health services in Norway

Toril Sørheim Nilsen

Department of Psychology

Faculty of Health Sciences

UiT The Arctic University of Norway

Dissertation for the degree of Philosophiae Doctor (Ph.D.) April 2017

Table of Contentsi

Acknowledgements

The work presented in this dissertation started in 2006 as I was accepted within the "The National Program for Integrated Clinical Specialist and PhD-training for Psychologists" in Norway. This program was a joint cooperation between the Universities of Bergen, Oslo, Tromsø, The Norwegian University of Science and Technology (Trondheim), the Regional Health Authorities, and the Norwegian Psychological Association. The program was funded jointly by The Ministry of Education and Research and The Ministry of Health and Care Services. I would like to thank the National Project Committee and my local program coordinator, Ingunn Skre, for giving me the opportunity to realize this project and for the support and encouragement all the way to the finalization of the dissertation.

First of all, this project would never have been realized had it not been for the participating children, adolescents, parents, clinicians and administrative staff at the child and adolescent mental health services (CAMHS) in Alta and Silsand. There would be no study or results to present without your contribution. Thank you!

To Martin Eisemann and Siv Kvernmo, my two outstanding, stable and always-supportive supervisors: Thank you for being there for me all the way from the beginning in 2006 to the finalization of this work in 2017. You have both encouraged and inspired me in different ways, and you have shared willingly of your extensive and broad experience as researchers, clinicians and clinical researchers.

My supervisor, Martin: I appreciate that you have always been available and present, wherever in the world you have been. You have given concrete and solid advice to whatever questions I have raised, as well as to my ideas and my work. It would be great to collaborate with you in the future.

My supervisor, Siv, who was also the Head of the CAMHS North study: your expertise in the field, your enthusiasm and stream of new ideas for promoting the field of child and adolescent mental health is of great inspiration to me. I am very grateful for the good discussions and constructive feedback to my writings. I hope we will find common projects together in the future.

Bjørn Helge Handegård, statistician at the Regional Centre for Child and Youth Mental Health and Child Welfare (RKBU North): You have been a very pedagogical, confident and thorough consultant supervisor with much knowledge in the mental health research field. Your contributions on methodological and statistical issues have been important for the realization of paper 2 and 3. Thank you!

I would like to thank the “Odd Berg Gruppens Medisinske Fond” (The Odd Berg Group Medical Fund) for allocating me a scholarship so that I could finalize this dissertation during the summer of 2016. I would like to thank the Department of Psychology, at the UiT Arctic University of Norway, and Head of Department Ingunn Skre specifically, for additional support in the work on this dissertation. I would also like to thank my current employer, The Child and Adolescent Mental Health Clinic (BUP Tromsø) at the University Hospital of North-Norway and in particular head of section Lene Danielsen, for supporting the finalization of my doctoral thesis.

Colleagues and fellow PhD students at the Department of Psychology and the research group of clinical psychology in particular: Thank you for inspiring conversations, for your sharing great enthusiasm and inspiring ideas in research and for being such a good support a long the way. I had some great years working with you.

Colleagues at the University Hospital of North-Norway: Thank you for sharing your clinical knowledge and expertise, for promoting the important field of child and adolescent mental health every day and for the personal and professional support that you have given me these years. I look forward to spend my continuing working days together with you.

Last, but not least: To my family and friends, my beloved crowd, who are always there for me and who have supported and encouraged me all along the way: Thank you for being in my life! And to Trygve, my beloved son: Thank you for reminding me every day what life is all about.

Abstract

Anxiety and depressive disorders are among the most prevalent mental health problems among children and youth, although a considerable proportion of this group do not receive any help to relieve the problems. Anxiety and depression can lead to considerable problems with daily- functioning in the family, with friends, with leisure activities and with learning in school. There is a risk that the difficulties can worsen, become chronic, and that the individual develops additional problems with substance misuse and dropping out of school. The goals of this dissertation was to evaluate the rate of change for children and adolescents with anxiety and/or depression (hereafter referred to as emotional disorders) that had been treated within two Child and Adolescent Mental Health Services (CAMHS) in the North of Norway. Another goal was to investigate associations between different demographic and clinical characteristics of the child with rate of change during the CAMHS treatment.

The empirical studies included in this dissertation are based on The Child and Adolescent Mental Health Services North (CAMHS North) Study. The CAMHS North study was a regional multicenter study including four CAMHS outpatient clinics in the northern part of Norway, and it was carried out in the time period between 2002 and 2008. The study was of naturalistic observational type where data from clinical instruments were collected as part of ordinary clinical practice. When this multicenter study was planned and designed, few other naturalistic and clinical studies about this health service existed in Norway. The study sample consisted of 84 children and youth with emotional disorders treated at two CAMHS in the North of Norway. The Strengths and Difficulties Questionnaire (SDQ), the Health of the

Nation Outcome Scale (HONOSCA) and the Children's Global Assessment Scale (CGAS) were used as measures of change.

Paper 1 is a literature review aimed at examining the extant research studies concerning pre-treatment child and adolescent characteristics as predictors and moderators of outcome in anxiety and depression psychotherapy treatment trials. In total, 44 published studies (32 anxiety studies and 12 depression studies) meeting the predefined methodological criteria were identified. The results showed that none of the pre-treatment demographic variables, such as age, gender and ethnic background, were found to consistently predict or moderate treatment outcome in both the anxiety and depression treatment trials, suggesting that the psychosocial treatments for anxiety and depression in children and youths are equally effective for girls and boys, young children and adolescents and across ethnic backgrounds. In the anxiety studies, no clinical factors (e.g. severity, comorbidity) were associated with differences in treatment change in the majority of studies. Among the depression treatment studies, higher levels of baseline symptomatic severity were associated with worse outcome in the majority of depression treatment studies that evaluated severity as predictor (3 out of 4 studies). There was also some evidence that comorbid anxiety and depression can be negative prognostic factors of treatment response. The conclusion drawn from the literature review was that existing studies of pre-treatment patient variables as predictors and moderators of anxiety and depression treatment outcome provide little consistent knowledge concerning for whom and under what conditions treatments work.

Paper 2 aimed at describing changes in symptomatic and functional impairment for children and youth with emotional disorders treated at two CAMHS in Norway. The results showed that children and youth with emotional disorders experience a statistically significant improvement per month during outpatient treatment according to nearly all the measures of change. For the clinician-rated scores, change rates during active assessment/treatment were larger than during the waitlist period. Evaluating change from the perspective of clinical significance showed that only a small proportion of the subjects had statistically reliable and clinically significant change scores. Whether an actual change has occurred is uncertain for the majority of patients.

Paper 3 aimed at examining demographic and clinical characteristics as potential predictors of change for children and youth with emotional disorders treated at two CAMHS in Norway. There was a difference between the two clinics in clinician rated functional impairment, where the CAMHS Silsand group was rated as more functionally impaired and with a steeper rate of change compared to the CAMHS Alta group of patients. The results showed that, according to clinician-rated total severity, youths with a diagnosis of depression had statistically higher symptom severity levels at baseline and significantly lower change rates as compared to youths with an anxiety disorder. No other variables were associated with differences in rate of change. The results point to a special need to improve clinical care for depressed children and adolescents.

There is little knowledge about how the health service in Norwegian CAMHS works for different patient groups, and little is known about factors that may influence change during treatment. The current studies add to this limited knowledge about children and adolescents with emotional disorders who have received health care in two Norwegian CAMHS.

In sum, the results of these studies point to the importance of tracking change during CAMHS treatment and to study change through different approaches and from the viewpoint of different informants. Evaluating change both as statistically significant change at the group level, in comparison with waitlist rates of change and from the perspective of clinical significance at the individual level seems warranted. In addition, it is important to have sufficient data to reliably track change as rated both by children themselves, their

parents/other caregivers and clinicians. Focusing on prognostic factors for change during CAMHS service provision is also an important further venue. Many methodological limitations, which impact the interpretation of findings, are discussed. Important limitations affecting the external validity of the study concern selection bias, missing data, a small study sample, change measures, the lack of a control group, and clinician bias in reporting.

Sammendrag (abstract in Norwegian)

Angstlidelser og depresjon er blant de mest vanlige psykiske lidelsene hos barn og ungdom samtidig som studier viser at et mindretall mottar hjelp for disse vanskene. Angstlidelser og depresjon kan medføre betydelige vansker med å fungere i hverdagen, i familieliv, med venner, i fritidsaktiviteter og på skolen. Det er en risiko for at vanskene kan forverres, bli kroniske, at personen utvikler tilleggsvansker som rusproblemer eller ikke fullfører skolegang. Målet med studiene i denne avhandlingen var å evaluere endringsrate hos barn og ungdom med angst og depresjon som har mottatt hjelp i Barne- og ungdomspsykiatrisk poliklinikk (BUP), samt å undersøke hvorvidt det er bestemte egenskaper hos barnet eller ungdommen som kan ha betydning for prognose i behandlingen.

De empiriske studiene som utgjør grunnlaget for denne doktorgradsavhandlingen baserer seg på data fra BUP Nord studien. BUP Nord var en multisenterundersøkelse som ble gjennomført ved fire barne- og ungdomspsykiatriske poliklinikker (BUP) i Helse Nord i perioden 2002 – 2008. Da studien ble utformet fantes det få naturalistiske og kliniske studier om denne helsetjenesten i Norge. Utvalget bestod av 84 barn og ungdom med angst og/eller depresjon som hadde mottatt et tilbud ved to BUP klinikker. Endringsmålene var Sterke og svake sider (SDQ-Nor), the Health of the Nation Outcome Scale (HONOSCA) og the Children's Global Assessment Scale (CGAS).

Artikkel 1 er en litteraturstudie, der en gjennomgår psykoterapeutiske behandlingsstudier som har fokusert på demografiske (eks. alder, kjønn) og kliniske egenskaper (eks. alvorlighetsgrad, komorbiditet: samtidig tilstedeværelse av flere psykiske lidelser) hos barn med potensiell betydning for endring under behandling. Artikkelen hadde til formål å skape en oversikt over forskningsfeltet i forkant av de empiriske studiene som er inkludert i denne avhandlingen. Totalt ble 44 publiserte studier inkludert (32 studier av

angstlidelser og 12 studier av depresjon). Resultat av litteraturgjennomgang viste at faktorer som alder, kjønn, og etnisk bakgrunn ikke var forbundet med forskjeller i behandlingsutfall for barn og ungdom med angst eller depresjon. Det var forskjell mellom studiene med fokus på primære angstlidelser, og studier av depresjonsbehandling. Ingen av de kliniske faktorene viste sammenheng med forskjell i behandlingsutfall for angstlidelser. Høyere alvorlighetsgrad ved oppstart viste sammenheng med lavere behandlingsrespons hos barn og ungdom med depresjon. Samtidig angstlidelse var forbundet med lavere behandlingsrespons i noen studier av depresjon. Konklusjonen av litteraturgjennomgangen var at en ut fra eksisterende studier har lite kunnskap om egenskaper ved barn og ungdom som kan ha betydning for endring under behandling, og en vet lite om for hvem og under hvilke betingelser behandling for angst og depresjon virker.

I artikkel 2 ble endringer i symptomnivå og funksjonsnivå i løpet av et utrednings- og behandlingstilbud i BUP evaluert. Mål på endring av symptomer og funksjonsnivå fra barn og ungdom selv, foreldre og fagpersoner i BUP ble vurdert. Tre tilnærminger ble brukt for å evaluere endring, nemlig statistisk signifikant endring på gruppenivå, med ventelistekontroll, og ut fra klinisk signifikant endring på individnivå. Resultatene viste at barn og ungdom med angst og depresjon oppnår statistisk signifikant bedring per måned på gruppenivå i følge nesten alle endringsmål, i løpet av tidsperioden da de har hatt et tilbud ved BUP. I følge endringsmål fra fagpersoner i BUP så er endringene i perioden for utredning/behandling større enn tilsvarende endring under ventelisteperioden. Tilsvarende resultat for mål rapportert av ungdom og foreldre viser ikke forskjell mellom disse to tidsperiodene, men lav svarprosent gjør at disse resultatene må tolkes med varsomhet. Evaluering av endring ut fra klinisk signifikant endring viser at kun et fåtall individer kan klassifiseres som ”bedret”. For majoriteten er grad av endring usikker.

I artikkel 3 testes det hvorvidt ulike demografiske og kliniske egenskaper ved barnet/ungdommen har sammenheng med endringsrate i behandlingsforløpet. Resultatene viste en forskjell mellom klinikkene, der utvalget fra en av poliklinikkene var vurdert med signifikant lavere psykososial fungering ved inntak, og også en større grad av endring for fungering under behandlingsforløpet. Egenskaper som kjønn og alder ved inntak viste ikke sammenheng med grad av endring. Resultatene viste at barn og ungdom med depresjon hadde høyere verdier for klinikervurdert alvorlighetsgrad ved inntak, og de opplevde mindre grad av endring sammenlignet med barn og ungdom med angstlidelser.

Det er begrenset kunnskap om hvordan helsetilbudet i BUP virker for ulike pasientgrupper, og egenskaper som kan virke inn på endring under behandling. Studiene innenfor denne avhandlingen er ett bidrag til å øke kunnskapen om barn og ungdom med angst og depresjon som mottar helsehjelp i BUP. Samlet indikerer funnene av disse studiene at det er viktig å måle endring under behandling i BUP og å evaluere endring ved hjelp av ulike statistiske og metodiske tilnærminger. Statistiske signifikant bedring på gruppenivå kan være en nyttig indikator på organisasjonsnivå, men sier ikke noe om helsetilbudets relative effekt. De fleste individer har høye skårer på henvisningstidspunktet, og mange vil oppleve en spontan bedring som ikke er relatert til selve helsetilbudet. Det er viktig å ta i bruke evalueringsmetoder som kan estimere helsetilbudets relative betydning for individenes endring. Ventelistekontroll er en måte å kunne estimere en slik effekt i de tilfeller hvor en mangler en kontrollgruppe. For den enkelte kliniker er det å evaluere endring ut fra klinisk signifikans på individnivå trolig en mer nyttig tilnærming. Denne avhandlingen viser også at det er behov for å undersøke faktorer som kan påvirke endring under behandling, og at kunnskapen man har om dette per i dag er lite entydig. Resultatene fra studiene i denne avhandlingen drøftes og sammenlignes med andre relevante studier. Metodiske begrensninger ved studiene i denne avhandlingen drøftes, og er viktige i forhold til tolkning av resultatene.

List of papers

1. Nilsen TS, Eisemann M, Kvernmo S (2013) Predictors and moderators of outcome in child and adolescent anxiety and depression: a systematic review of psychological treatment studies. *European Child and Adolescent Psychiatry* 22:69-87.

2. Nilsen TS, Handegård BH, Eisemann M, Kvernmo S (2015) Evaluating change in symptomatic and functional level of children and youth with emotional disorders: a naturalistic observation study. *European Child and Adolescent Psychiatry* 24:1219-1231.

3. Nilsen TS, Handegard BH, Eisemann M, Kvernmo S (2016) Predictors of rate of change for children and youth with emotional disorders: a naturalistic observational study. *Child and Adolescent Psychiatry and Mental Health* 10:11-11.

Abbreviations

BCS - The Bergen Child Study

CAMHS - child and adolescent mental health services

CAMHS North study – The Child and Adolescent Mental Health Services North (CAMHS North) Study (multicenter study conducted in the North of Norway)

CBT - cognitive-behavioural therapy

CGAS - the Children's Global Assessment Scale

CORC - Child Outcomes Research Consortium

GAD - generalized anxiety disorder

DSM V - the Diagnostic and Statistical Manual of Mental Disorders

HONOSCA - the Health of the Nation Outcome Scale

K-SADS-PL - The Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime version (6-18)

RCT - randomized controlled trials

REK - The Regional Committee for Medical Research Ethics

RTM - regression toward the mean

SAD - separation anxiety disorder

SOC - social phobia

SDQ - The Strengths and Difficulties Questionnaire

SPSS - Statistical Package for the Social Sciences

TAU - treatment as usual

UC - usual care

WISC III - Wechsler Intelligence scale

Introduction

Background

Back in 2006, when I started working as a clinical psychologist at the child and adolescent mental health outpatient services (CAMHS) in Tromsø. I became increasingly interested in working clinically with the group of children and adolescents with internalizing response patterns and with anxiety and depressive disorders (hereafter referred to as emotional disorders). At that time, the CAMHS Tromsø formed part of a large research program, The Incredible Years focusing on children with externalizing disorders, namely conduct disorders, oppositional defiant disorder, and hyperkinetic disorder. Little was known about the group of children with emotional disorders within the CAMHS system of care. Focusing on the group of children with emotional disorders was in line with both my clinical and research interests.

In the present dissertation two major themes have been studied, namely change during CAMHS treatment and prognostic factors associated with rates of change. The target group is children and adolescents with emotional disorders. There is limited knowledge about rates of change during treatment as usual (TAU) in naturalistic settings. It is important to know whether children and adolescents accessing the service seem to be better off at follow up, and to be able to identify negative outcomes, or no change. Further, little is known about how the health service in Norwegian CAMHS works for different patient groups and about factors influencing change during treatment. The study of prognostic factors may contribute to a better understanding of what kind of treatment is effective for whom, and may facilitate the process of individualizing treatment interventions (Kraemer, Wilson, Fairburn, & Agras, 2002; La Greca, Silverman, & Lochman, 2009).

The Child and Adolescent Mental Health Services North (CAMHS North) Study was a regional multicenter study including four CAMHS outpatient clinics in the northern part of Norway: CAMHS Alta, CAMHS Silsand, CAMHS Storslett and CAMHS Tromsø. The study was of a naturalistic observational type, where data from clinical instruments were collected as part of the ordinary clinical practice. The study was carried out in the time period between 2002 and 2008. The primary aim of the multicenter study was to evaluate clinical procedures for assessment by implementing and evaluating structured instruments, to evaluate change during treatment, to investigate factors affecting waiting time, and to investigate user satisfaction. The outpatient clinics covered both urban and rural areas. The empirical studies of this dissertation only included participants from CAMHS Alta and Silsand, since these were the only clinics with follow-up data. CAMHS Alta collected data between 2002 and 2005, while CAMHS Silsand collected data between 2004 and 2006.

Child- and Adolescent psychiatry has been gradually developing from the 1960`s onwards, and is therefore a rather new service in Norway (Statens helsetilsyn, 2000). In Norwegian mental health services, 95 % of children and youth are treated in outpatient clinics (Brofoss, Larsen, Friis, & Norges forskningsråd, 2009). Child and adolescent psychiatry was targeted by the Norwegian government plan on improving the mental health care of the population during the decade between 1998 and 2008 (Brofoss et al., 2009). The goal of Norwegian health care authorities is that the outpatient CAMHS for children and adolescents shall be staffed to provide services for 5 % of the youth population (Tuseth, Sverdrup, Hjort, & Friestad, 2006). According to White Paper 25 (1996-97) and 39 (1998) it is an official requirement that the health care system must be based on research-based practice (Tuseth et al., 2006). Still, very few studies have been conducted regarding the quality of these services in Norway. Despite the focus on quality assurance and research-based practice, most CAMHS do not evaluate patient change systematically. To what extent the treatment provided within

Norwegian child psychiatry results in desired change for children and youth continues to be unknown.

Emotional disorders: Anxiety and depression

Emotional disorders are among the most prevalent mental health problems experienced by children, and are also among the largest groups within CAMHS. As results from The Bergen Child Study (BCS) indicate, Norwegian parents and teachers tend to under-recognize this group of children (Heiervang, Goodman, & Goodman, 2008), and only 13 % of this group receive specialized mental health care. Among depressed adolescents, less than 20 % had been in contact with specialized mental health services (Sund, Larsson, & Wichstrøm, 2011).

There are arguments, both pros and cons, as to whether one should study anxiety and depression together or separately (Wilkinson, 2009). The classification of symptoms of anxiety and depression into "emotional disturbance" in childhood psychopathology was firstly made by Hewitt and Jenkins in 1946 (reference in Wilkinson, 2009). The development of the diagnostic classification systems ICD (Helsetilsyn & World Health Organization, 1999) and DSM (American Psychiatric Association, 2013) has since that time contributed to a greater complexity and differentiation of childhood mental disorders, and the distinction between mood and anxiety disorders has been included in all subsequent versions. As stated in the introduction section of the DSM V (American Psychiatric Association, 2013) larger groupings of disorders have received more scientific support than the validation of individual diagnostic categories, and internalizing and externalizing factors are considered as a scientifically supported framework (Achenbach, Ivanova, Rescorla, Turner, & Althoff, 2016).

Anxiety and depression are regarded as distinct but related disorders (Cummings, Caporino, & Kendall, 2014), and there is an ongoing debate as to how the different anxiety and depressive disorders should be classified, categorized and grouped (e.g Beesdo, Pine, Lieb, & Wittchen, 2010; Mohr & Schneider, 2013). The principal emotions distinguish depressive disorder (depressed mood) and anxiety disorder (anxiety), while the secondary symptoms overlap considerably (e.g. difficulty with sleep, reduced concentration, rumination) (Wilkinson, 2009). In particular, there is considerable overlap between depressive disorder and generalized anxiety disorders (Kendler, 1996), and between social phobia and depressive disorders.

Comorbidity refers to the co-occurrence of two or more distinct disorders in one individual. Anxiety and depressive disorders often co-occur (Melton, Croarkin, Strawn, & McClintock, 2016), and epidemiological studies show that it is common for children diagnosed with one anxiety disorders to display symptoms of other anxiety disorders and/or depressive symptoms (Costello, Mustillo, Erkanli, Keeler, & Angold, 2003). Anxiety often occurs prior to the onset of depression (Strauss, Last, Hersen, & Kazdin, 1988). Rates of comorbidity between depression and anxiety disorders have ranged from 15.9 % to 61.9 % (in Beidel & Alfano, 2011, pp 33), but comorbidity is likely underestimated in children and adolescents (Melton et al., 2016). A meta-analysis of community epidemiological studies showed that 32 % of children/youth with major depression also had an anxiety disorder, and 24 % of adolescents with an anxiety disorder had major depression (Angold, Costello, & Erkanli, 1999). An epidemiological study showed that during a 3-month period, 28 % of the youngsters with a depressive disorder also had an anxiety disorder (Costello et al., 2003). Generalized anxiety disorder and social phobia co-occur more frequently with depression compared to separation anxiety disorder and specific phobias (Verduin & Kendall, 2003). Comorbid anxiety and depression have a worse prognosis than either conditions alone, being associated with more severe symptoms and greater functional impairment (Melton et al., 2016). Neuroimaging research show that both distinct and common brain structures and neurocircuitry are involved in depression and anxiety, and some preliminary findings suggest that adolescent with anxiety with comorbid depression may have distinct patterns of structural and functional abnormalities compared to adolescents with either depression or anxiety disorders (Melton et al., 2016).

Anxiety disorders

Worldwide prevalence rates of anxiety disorders in children and adolescents show that these disorders are the most prevalent psychiatric conditions in childhood and adolescence (Beidel & Alfano, 2011), that rates increase with age (Beesdo, Knappe, & Pine, 2009; Esbjorn, Hoeyer, Dyrborg, Leth, & Kendall, 2010), and that more females than males have anxiety disorders. Epidemiological studies vary considerably in prevalence rates. Among pre-adolescent children, epidemiologic studies have found prevalence rates ranging from 2.6 % to 41.2 % (Cartwright-Hatton, McNicol, & Doubleday, 2006), and point prevalence rates ranging between 12 and 20 % in children and adolescents with different types of anxiety disorders (Costello, Egger, & Angold, 2005). The average age of onset differs depending on the specific anxiety diagnosis, but anxiety disorders can occur at any time during childhood and adolescence. Boys with anxiety disorders were referred at a younger age than girls to the CAMHS in Denmark (Esbjorn et al., 2010). The age groups of 7–12 year olds (boys) and 13–19 years old (girls) were more frequently referred for treatment than boys and girls in other age groups.

Anxiety disorders are sub-classified by the stimuli that trigger the anxiety. The most prevalent anxiety disorders among children and adolescents are separation anxiety disorder (SAD), generalized anxiety disorder (GAD), specific phobia (SP) and social phobia (SOC)

(Beesdo et al., 2009). Panic disorder and agoraphobia are rare in children, with an increasing prevalence during adolescence. In SAD, the child is overly anxious when faced with the potential separation from caregivers, from home, or by the concern of something happening to their caregivers. In SP, excessive fear is triggered when facing specific stimuli (e.g. certain animals, heights, small rooms). Generalized anxiety disorder (GAD) is characterized by excessive worry related to different topics and situations. Social phobia is characterized by a marked and consistent fear of negative evaluations in social and performance situations. Physiological symptoms of anxiety, such as increased heart rate, rapid breathing, trembling, muscle tension, are central diagnostic markers. Avoidance of the feared situation or stimuli is a behavioural diagnostic criterion in all anxiety diagnosis. The diagnosis of anxiety disorders lie on a continuum, where the level of distress and impairment caused by the anxiety (especially avoidance), is crucial for diagnostic evaluation. The aspect of normal development is important in the assessment of normal to maladaptive fear. Research suggests that anxiety disorders in childhood seem to function as a pacemaker for later psychiatric conditions (Wehry, Beesdo-Baum, Hennelly, Connolly, & Strawn, 2015)

Depressive disorders

The World Health Organization is ranking depression as the leading cause of disability worldwide, and as a major contributor to the overall burden of disease globally (WHO., 2017). Depression affects around 3 % of children, and between 5-8 % of adolescents (Costello, Erkanli, & Angold, 2006). By the end of adolescence, however, around 28 % may have experienced a depressive episode (Lewinsohn, Rohde, & Seeley, 1998). Gender differences in depressive symptoms become apparent around the age of 13 with statistically significantly higher scores for girls than for boys (Twenge & Nolen-Hoeksema, 2002). The prevalence of depression increases with age for both genders, but with the double rate for girls. Results from the Youth and Mental Health Study in Central Norway showed that the different depressive disorders all had long duration of episodes, and that the duration of episodes varied considerably (Sund et al., 2011). The mean episode duration of depression has been estimated to be between 4 to 9 months among clinically referred adolescents (Birmaher, Arbelaez, & Brent, 2002; Emslie, Mayes, Laptook, & Batt, 2003). In a Finnish study, approximately 50 % of depressive episodes lasted longer than the one year follow-up time (Karlsson et al., 2008). Around 2/3 of adolescents with a diagnosis of depression at age 15 remain depressed at age 20, and the stability of depressive disorder was apparent in both genders (Agerup, Lydersen, Wallander, & Sund, 2014).

Depressive disorders are syndromes of enduring emotional, biological and psychological symptoms with accompanying psychosocial impairment (Wilkinson, 2009). Major depression (DSM IV, American Psychiatric Association, 2000) and depressive episode (World Health Organization, 1992) are the most prevalent type of depressive disorders. The core symptoms of depression are depressed or irritable mood, and diminished interest or pleasure in almost all activities. Other common symptoms are changes in appetite/weight, changed sleep pattern, psychomotor agitation or retardation, reduced concentration or indecisiveness, reduced energy, increased fatigability, excessive feelings of worthlessness and guilt, recurrent thoughts of death or suicidal ideation. In the DSM IV, a diagnosis of major depression is characterized by at least five depressive symptoms that have been present most of the time for at least two weeks, and that are associated with a reduction of prior functioning. In the ICD-10, an assessment of the severity of the condition, whether it can be classified as mild, moderate or severe, is evaluated on the basis of numbers of symptoms and degree of functional impairment. Dysthymic disorder is characterized by a more chronic, yet less severe condition, with depressed mood most of the time accompanied by a minimum of two other symptoms of depression. Minimum duration is one year. In DSM IV, depressive disorder not otherwise specified is characterized by a depressive condition that does not meet the full set of diagnostic criteria.

The research field of depression in childhood and adolescence is relatively new, from the 1980's onward. The study of depression and its treatments in children and adolescents has lagged behind advances in our understanding of other areas of psychopathology for this age group (Weisz, McCarty, & Valeri, 2006).

The evaluation of change in mental health services

A common definition of outcome in mental health is a measurable change in the mental health condition of the individual/group/population that can be attributed to an intervention or a service (Patterson, Matthey, & Baker, 2006). Several recommendations for the measurement and implementation of change data within CAMHS are presented in the literature (Wolpert et al., 2014a, 2014b). In short, the evaluation of change should be multidimensional (including changes in e.g. both symptoms and adaptive functioning), and multi-informant (both the clinician and the youth/caregiver perspective should be represented) (Wolpert, 2008). Measures included must have good psychometric properties, should be simple and quick to complete, be cost-effective and easy to interpret (Hall et al., 2014). Also, change should be evaluated and reported by taking into account both statistically- and clinically significant changes (Campbell, 2005), and to apply the most appropriate approach for the problem to be investigated, whether at the service level or at the individual level (Wolpert, Goerzig, et al., 2015). Different approaches to change evaluation may lead to different conclusions regarding the effectiveness of services, an observation reported in paper

2 of this dissertation, and from the Child Outcomes Research Consortium (CORC) data (Wolpert, Goerzig, et al., 2015).

If we want to improve the mental health care of children, we need to know how it is currently working, what is functioning well and what is in need of improvement. Studies of the effectiveness of community-based usual care (UC) show that there is no convincing evidence of a strong clinical impact of outpatient specialty mental health care for children (Garland et al., 2013). Concerns about the limited effectiveness of usual care have influenced policy and research initiatives internationally with the goal of improving the quality of the services. In Great Britain, child and adolescent mental health services (CAMHS) are expected to evaluate change during service provision by systematically measuring outcome (Hall et al., 2013). Through collaborative efforts such as the Child Outcomes Research Consortium (CORC), a non-profit learning collaboration formed in 2002 in Great Britain, routine outcome measurement is becoming more common in Great Britain, and is also spreading gradually to other European countries. Still, most CAMHS do not track patient change. In Norway, no such national initiative has been raised to evaluate the CAMHS service.

What is currently known about the effectiveness of usual care within CAMHS for children and adolescents? Several studies have reported statistically significant improvement across diagnostic groups according to different change measures from intake to follow-up while in CAMHS care (Deighton et al., 2016; Lundh, Forsman, Serlachius, Lichtenstein, & Landen, 2013; Trask & Garland, 2012; Wolpert et al., 2012; Wolpert, Goerzig, et al., 2015), and effect sizes (ES) indicate positive, but modest effects (e.g Fugard et al., 2015; Wolpert, Cheng, & Deighton, 2015). Three studies reported results for anxiety disorder, depression, emotional disorders or internalizing problems separately. In one naturalistic observational study using archival data from CAMHS in Stockholm, Sweden, clinicians rated Children's Global Assessment Scale (CGAS) indicated improvement in anxiety- and mood disorders with similar change scores. Change scores for the mood disorder group was twice as large as for the group with ADHD. The authors state that the level of improvement reported for the mood disorder group correspond to efficacy trials in depression (March et al., 2004; Wagner, Jonas, Findling, Ventura, & Saikali, 2006), but no such comparisons with relevant treatment trials are made for the anxiety disorder group. A second study, including a large representative sample served within a large county mental health system in the USA, also suggest improvement in behavioural and emotional problems with effect sizes corresponding to the small to medium range (.39 to .69) after 6 months of treatment as usual (Trask & Garland, 2012). Although effect sizes were found to be unrelated to child diagnosis, the magnitude of change was smaller for internalizing problems when compared with externalizing problems. In a study from CAMHS in London (Day & Davis, 2006) reduction in total problems and burden scores corresponded to effect sizes in the medium to large range (0.5 to 1.0) that were maintained at 1-year follow up. While these findings suggest substantial improvement, the reduction of internalizing problems did not reach statistically significant results, and a significant number of children still had problems that qualified for clinical "caseness" after intervention. For all the above-mentioned studies, the lack of a control group is a challenge if one wishes to estimate the relative contribution of the service to change.

Evaluating change from the perspective of clinical significant change and statistical reliable improvement indicates that the majority of children that receive community-based usual care (UC) do not experience clinical improvement (Garland et al., 2013; Warren, Nelson, Mondragon, Baldwin, & Burlingame, 2010). The reliable change index (RCI) is a statistics estimating if changes in psychometric measures may represent real change or occur by measurement error (Jacobson & Truax, 1991). It has been proposed as a means of meaningfully categorizing individuals as *improved*, *unchanged* or *deteriorated*. According to results on parent-rated Strengths and Difficulties Questionnaire (SDQ) from the CORC database, 16.5 % were classified as ”improved” according to the reliable change index (RCI), while 2.3 % were classified as RCI ”worse” (Wolpert, Goerzig, et al., 2015). Manteuffel et al. (2008) used the reliable change index to examine patterns of change for 8484 youths in 45 systems of care across 36 states of the USA. According to the results on parent rated data, an average of 36 % of youths improved, 50 % exhibited no reliable change and the remaining 14 % exhibited poorer outcomes following treatment. Finally, also based on parent-rated data, Warren, Nelson and Burlingame (2009) reported that over 50 % did not achieve positive outcomes in therapy, and 21 % had significantly higher symptoms at the end of treatment, and an additional 30 % did not achieve a reliable change in symptom levels.

Considering the lack of convincing evidence of a strong clinical impact of usual care within CAMHS the question raises of how usual care performs when compared to evidence based treatments? Weisz (2013) conducted a multi-level meta-analysis based on 52 studies comparing evidence based psychotherapies (EBI) with usual care. The results showed overall better outcomes for EBI´s, but the difference was modest and several usual care services outperformed EBI´s. Studies using exclusively diagnosed samples (d=0.09), and studies that focused on clinically referred youths (d=0.17) showed low and non-significant ES values, which seem to support the argument that EBIs have mostly been tested with subclinical youths, and therefore might not apply well to real-world clinical settings. The EBI – UC difference was also smaller in studies conducted outside North America, indicating that EBI is challenged by cultural differences. Outcome data from 30 child and adolescent mental health services in the UK showed both the evidence-based practice and non-evidence-based practice groups improved over time, with moderate effect sizes (Deighton et al., 2016). There were greater improvements associated with evidence-based practice (CBT) for children with emotional disorders, based on child self-reported symptoms, but not on parent reports. Thus, findings provide tentative support for evidence-based practice for the treatment of emotional disorders in routine care settings - based on child self-report. In another meta-analysis (Spielmans, Gatlin, & McFall, 2010) the modest benefit of evidence-based treatments for youth was reduced when controlling for potential confounding factors (e.g. treatment dose, supervision, caseload). According to Ng and Weisz (2016), the relatively limited superiority of EBI compared to UC may be due to UC being highly individualized.

What is the effect of psychological treatment for anxiety and depression in more controlled settings? Recent meta-analysis suggest that there has been a substantial decline in depression treatment effect sizes over time (Weersing, Jeffreys, Do, Schwartz, & Bolano, 2016), that effects are modest and that the effects of evidence based treatments for depression lag behind the effects for other youth disorders (Weisz et al., 2006). The decline in effect sizes may be partly due to actual progress within the field of depression treatment studies with more recent trials involving active control conditions. A meta-analysis on depression psychotherapy treatment trials published 10 years ago (Weisz et al., 2006), found an average effect size of 0.34, falling in the small to medium range according to Cohen's benchmark (Cohen, 1988), but there were considerable variations in effect sizes between studies. Also, beneficial changes showed no lasting effects, although few studies assessed effects at long-term follow-up (Weersing et al., 2016; Weisz et al., 2006). Also, effects of CBT have been found to decline in the context of more severe, complex and clinically impaired depressed youths (Emslie, Kennard, & Mayes, 2011; March et al., 2004; Rohde, Waldron, Turner, Brody, & Jorgensen, 2014). In their review of predictors of depression treatment, Emslie et al. (2011) report 30 – 40 % experiencing remission of the depressive episode and recurrence rates within 1 year as high as 50 %, among depressed children receiving evidence-based treatments Thus, the need for further developments of available depression treatments have been highlighted, and especially to increase the understanding of mechanisms of change in depression treatment (Weersing et al., 2016).

One recent review states that the evidence for psychological treatments of depressed children aged 12 years and below is inconclusive, probably due to the limited number of studies conducted in this age group (Forti-Buratti, Saikia, Wilkinson, & Ramchandani, 2016). A recently published review (Weersing et al., 2016) of 42 randomized controlled trials and of previously reviewed literature, points out that evidence for child depression treatment is weaker than for adolescent treatment. Cognitive behavioural therapy (CBT) was classified as *possibly efficacious* for children, while no child intervention could be classified as *well established*. For adolescents however, both CBT and interpersonal psychotherapy (IPT) could be classified as *well-established* interventions. The authors state that the dominance of CBT rests on the larger treatment outcome literature for CBT compared to other treatments (e.g 27 RCT studies of CBT versus 6 for IPT), and not only on the strength of findings (Weersing et al., 2016). A recent meta-analysis concludes that CBT and interpersonal therapy (IPT) should be viewed as the best psychotherapies available at present, but several alternative psychotherapy approaches for depression are understudied (Zhou et al., 2015). In one review (Restifo & Bogels, 2009) the comparison of individual with family focused approaches to the treatment of adolescent depression yielded mixed results. Treating parent depression may be a promising approach for alleviating adolescent depression according to one review (Gunlicks & Weissman, 2008).

For the treatment of anxiety disorders in children and adolescents, the majority of approaches with strong support are CBT interventions and exposure based interventions (Higa-McMillan, Francis, Rith-Najarian, & Chorpita, 2016). Three meta-analytic reviews of evidence based treatments for anxiety disorders found clear beneficial effects of CBT when compared to passive controls (Ewing, Monsen, Thompson, Cartwright-Hatton, & Field, 2015; James, James, Cowdrey, Soler, & Choke, 2015; Reynolds, Wilson, Austin, & Hooper, 2012).

While the meta-analytic review by Reynolds (2012) found significant but small effect sizes favouring CBT when compared with active control conditions, the most recent Cochrane-review found no evidence that CBT outperforms active controls such as non-CBT or treatment as usual (TAU) (James et al., 2015). Further, the meta-analytic review by Reynolds (2012) found that the overall effect sizes for generic CBT programs (targeting different anxiety disorders) were moderate, while the disorder specific treatments showed medium to large effect sizes. Also, anxiety treatments for adolescents showed large effects sizes, while treatments for children showed small to medium effect sizes. One meta-analysis found that individual treatment was more effective than group interventions (Reynolds et al., 2012), while other meta-analysis found no difference between individual and group intervention (Bennett et al., 2013; James et al., 2015). Two meta-analyses conclude that parental interventions were not associated with improved effects for anxious children and adolescents (James et al., 2015; Reynolds et al., 2012), and another meta-analytic review even found a small and non-significant effect size favouring child-only treatment (Thulin, Svirsky, Serlachius, Andersson, & Ost, 2014). A dose-response effect was also demonstrated suggesting that providing five or more sessions were associated with small effects, while providing nine or more sessions was associated with medium to large effect sizes (Reynolds et al., 2012).

To sum up the results concerning outcomes in CAMHS usual care, results reported at the group level indicate positive results, although to a moderate extent. In most studies, surprisingly few individuals obtain clinical significant improvements. One venue further is to intensify efforts in understanding usual care practice, and to conduct rigorous research in naturalistic settings (Garland, Bickman, & Chorpita, 2010). One important step in individualizing treatments, and for improving the understanding of treatment effects and effectiveness, is to study prognostic factors that may impact change during treatment.

The study of predictors and moderators of change

According to La Greca et al.'s (2009, pp 376) *"Predictors of treatment outcome are variables that account for outcome irrespective of the treatment condition (main effect model)"*. In paper 1, all studies were RCT studies, and predictors included were baseline characteristics of the child or adolescent that were associated with post-treatment diagnostic status or symptomatic level regardless of condition (in line with the approach presented in Kraemer et al., 2002). In paper 3, demographic and clinical characteristics of the individual measured at baseline were tested as potential predictors of rate of change per month at the group level, in line with the approach of Singer and Willet (2003). Clinically, predictors of treatment response may serve as prognostic variables and may highlight targets for further treatment development.

Moderators of treatment outcome represent variables influencing the association between the intervention and the outcome (Baron & Kenny, 1986). A moderator variable affects the relationship between two variables, so that the nature of the impact of the predictor on the criterion varies according to the level or value of the moderator, i.e. an interactive effect. In paper 1 of the present dissertation, treatment moderators are baseline or pre-randomization characteristics associated with differential responses to an intervention. This definition of moderator variable is applicable to RCT design where comparison with a control group is possible. Clinically, the identification of treatment moderators helps to specify for whom and under what circumstances the treatment works (Kraemer et al., 2002). For treatment development, moderators may help to identify limitations of treatments and areas of further development within the mechanisms of action of specific interventions.

As noted in paper 1, predictors and moderators of change have been primarily investigated in randomized controlled trials (RCT) of cognitive-behavioural therapy (CBT) in research clinics. When conducting an updated literature search for the preparation of this summary, seven naturalistic observation studies examining predictors and moderators of change were identified. The studies were all of patients with different diagnosis, and yielded mixed results. Gender was found to be unrelated to treatment outcome in most studies (Fuggle, 2015; Gordon, Antshel, & Lewandowski, 2012; Warren et al., 2009; Warren et al., 2010), except one study which found that boys improved more on measures of internalizing behaviours than did girls (Trask & Garland, 2012). Age did not predict treatment outcome in most studies (Gordon et al., 2012; Trask & Garland, 2012; Warren et al., 2009; Warren et al., 2010), although one study found that younger children have better outcomes than adolescents (Fuggle, 2015). Child's diagnosis was not related to differential treatment outcome in two studies (Gordon et al., 2012; Trask & Garland, 2012), while one study found large differences in improvement depending on diagnostic group and the change score of the mood disorder group being twice as large as for the group with hyperkinetic disorder (Lundh et al., 2013).

An updated literature review of studies examining predictors and moderators of outcome within the anxiety treatment literature showed that demographic factors, such as age and gender, do not predict or moderate differential treatment response rates or diagnostic recovery in most studies (Compton et al., 2014; Hudson, Keers, et al., 2015; Lundkvist-Houndoumadi, Hougaard, & Thastum, 2014; Wergeland et al., 2016). Age was a moderator of treatment response in one study (Peris et al., 2015), and younger children showed steeper rates of improvement following exposure tasks compared to adolescents, indicating increased benefit of exposure. Several recent studies have found higher baseline anxiety severity (Compton et al., 2014; Lundkvist-Houndoumadi et al., 2014; Wergeland et al., 2016) to be associated with higher end-state anxiety levels, but not with the degree of improvement in most studies. Higher pre-treatment anxiety was associated with higher symptom scores throughout treatment in a 16- to 20 week CBT program (Chu, Skriner, & Zandberg, 2013). Higher pre-treatment functional impairment was associated with less favourable outcomes post-treatment in one study (Wergeland et al., 2016). A diagnosis of social phobia has been found to predict less favourable outcomes of general CBT programs for anxiety in several studies (Compton et al., 2014; Hudson, Keers, et al., 2015; Hudson, Rapee, et al., 2015; Kerns, Read, Klugman, & Kendall, 2013; Wergeland et al., 2016). One study found that youth with social anxiety symptoms or diagnosis were significantly less improved also at 7.4-year follow-up (Kerns et al., 2013). The results indicate that children and adolescent with social anxiety may need more tailored interventions targeting explicit social skills training. One study found better outcomes for social phobia following a CBT treatment targeting cognitive characteristics of social phobia when compared to a general anxiety treatment program (the Coping Cat) (Ingul, Aune, & Nordahl, 2014). Findings regarding comorbid depression as a predictor of treatment response are inconclusive. Two studies found comorbid symptoms of depression to predict less favourable outcome of CBT treatment for anxiety (Hudson, Keers, et al., 2015; O'Neil & Kendall, 2012), while two other studies did not find this association (Kley, Heinrichs, Bender, & Tuschen-Caffier, 2012; Wergeland et al., 2016). Waters and colleagues (2015) found three characteristics of the child's´ reaction pattern that was associated with more improvement following exposure training. Larger pre-treatment attention bias towards threat, greater emotional variability during exposure task, and larger habituation during the exposure training, were associated with more improvement from pre- to post-treatment.

The study of predictors and moderators of outcome within the depression treatment literature to date show that demographic factors do not seem to have a substantial impact on the response to treatment (Emslie et al., 2011; Weersing et al., 2016). In a recent review of psychosocial treatments of child and adolescent depression, the characteristics that were most consistently found to predict poor response across treatment and control conditions were higher levels of depression symptoms, poor global functioning, high levels of suicidality, comorbid anxiety, cognitive distortions, hopelessness and family conflict (Weersing et al., 2016). In the Adolescent Depression Study, a naturalistic observational study, longer time to recovery was predicted by earlier age at onset of depression, poor psychosocial functioning, having a diagnosis of depression, and longer episode duration at study entry (Karlsson et al., 2008).

In paper 1, the review of predictors and moderators of treatment response in the Treatment for Adolescents with Depression (TADS), the Treatment of Resistant Depression in Adolescents (TORDIA) and the Adolescent Depression Antidepressants and Psychotherapy Trial (ADAPT) was described (Emslie et al., 2011). The TADS, the TORDIA and the ADAPT trials were all large-scale RCT-studies of medication, CBT or a combination of CBT/medication. Overall, demographic characteristics did not predict or moderate treatment outcome with a few exceptions. In the TADS, younger adolescents had better response to treatment than older adolescents. In the TORDIA, the older age group (18-19 years old) had better response to combination treatment. Gender had no overall effect on treatment response, but long-term outcome of the TADS study showed that girls were more likely to experience relapse over time. Across studies, greater improvement following acute treatment was associated with less severe depression at baseline, better functioning, shorter depression duration, fewer comorbid psychiatric disorders, less suicidal ideation and less hopelessness at baseline. In the TADS and the ADAPT, the absence of a comorbid anxiety disorder was associated with more improvement. In the TORDIA, anhedonia was the only dimension that predicted longer time to remission, and was also the only dimension to predict fewer depression-free days (McMakin et al., 2012). The authors state that anhedonia may represent an important negative prognostic factor among treatment-resistant depressed adolescents. Since the CBT condition did not seem to be successful in alleviating symptoms of anhedonia, the authors raise the question of whether the current CBT treatments address anhedonia adequately. Further, in the TADS and the TORDIA, less family conflict and stress was also associated with more improvement. In another study, better baseline family functioning scores predicted a greater decrease in the adolescent depressions symptom scores, and also that families that had improved family functioning also showed significantly greater decreases in the youths' CDI scores.

To sum up the results on predictors and moderators of treatment effects, very few studies reported age and gender to be prognostic factors for anxiety and depression treatment outcome. Most investigations do not include large numbers of ethnic and racial minority youth, which limit the possibility to detect potential effects. Further, high pre-treatment severity of anxiety and depression predicts higher end-state levels of anxiety and depression, but is not necessarily predictive of the degree of improvement as seen in several studies. Another prevailing finding within both the anxiety and depression treatment literature to date is that very few of the studied variables show significant associations with change during treatment. Limitations of previous studies on predictors and moderators were described in paper 1 and will be further elaborated in the general discussion.

Aims of the thesis

There is limited knowledge about rates of change during treatment as usual (TAU) in naturalistic settings and careful examination of treatment outcome and change processes in naturalistic settings of usual care is needed. There is also little knowledge about prognostic factors influencing change during treatment. The identification of predictors and moderators of change can help clarify which subgroups respond well, and which groups that respond less well to the services (La Greca et al., 2009).

In the empirical studies of the present dissertation, the rate of change per month during CAMHS treatment is evaluated for children and adolescent with emotional disorders, and predictors of rates of change are assessed.

The following main research questions were addressed in this dissertation:

1. Based on the extant research literature review: To what extent have associations between child modalities and treatment outcome been confirmed?
2. What are the changes in symptom severity, functional impairment and emotional problems as reported by adolescents and parents, and in symptom severity and functional impairment as reported by clinicians at the group level, as compared to waitlist control and from the perspective of clinical significant change?
3. Are characteristics such as age, gender, baseline symptom severity or functional impairment, type of emotional disorder, comorbidity, prosocial characteristics and problem with peers associated with differences in rate of change over time?

Methods of paper 1

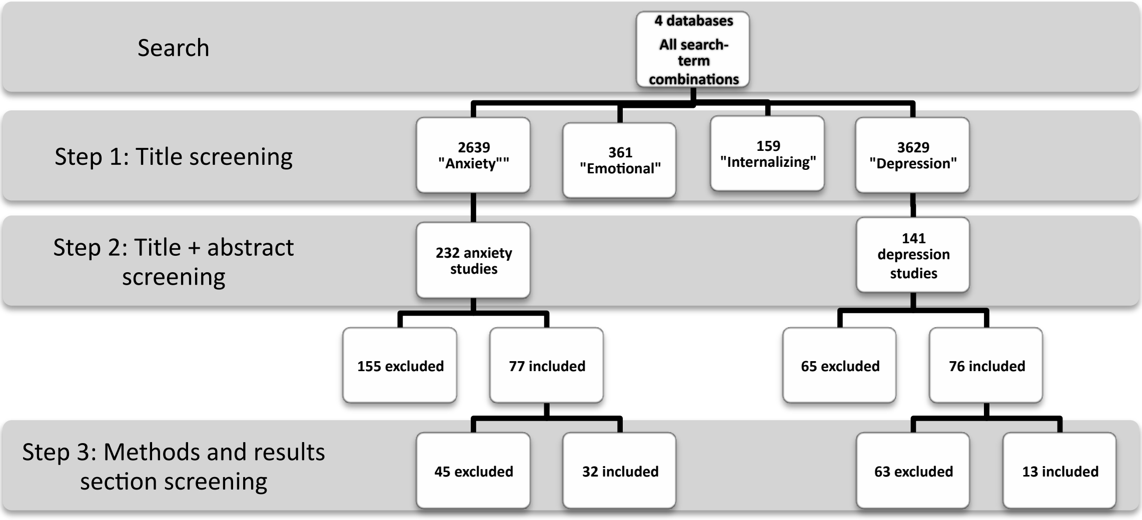
Literature search:

The aim of the literature review was to examine pre-treatment child and adolescent characteristics as predictors and moderators of outcome in anxiety and depression psychotherapy treatment trials. In order to identify relevant studies for inclusion in the review, I conducted literature searches in the ISI Web, PubMed, Medline (1985 through week 1 March 2011), PsycINFO (1967 through week 1 March 2011). For the literature search in the databases ISI Web, Pub Med, Medline and PsycINFO lsearches, four key terms related to diagnosis (“anxiety disorder*” or ”depression” or ”emotional disorder*” or “internalizing disorder”) were paired with each of six key terms related to treatment (“predictor* treatment outcome” or “mechanism* of change treatment outcome” or “moderator * treatment outcome” or “treatment outcome” or “treatment”), all paired with “Child* Youth* Adolescent*”. These search terms combined in the four databases yielded several thousand hits. For the PubMed and ISI Web databases I received weekly updates of relevant studies automatically. The literature search was conducted mainly during the spring of 2011 followed by some updated searches during the preparation and publication process in order to see whether new studies were published. Figure 1 displays a flowchart of the number of studies assessed for the present review. 1

---

[1] Details concerning the literature search process and selection of studies are described in supplemental material attached in appendix In step 3, 21 anxiety studies and 19 depression studies with statistical examinations of predictors and moderators were excluded because they did not meet other inclusion or exclusion criteria. List of excluded studies in Step 3 can be obtained from the corresponding author.

Figure 1. Flow chart of the search process



The following criteria were used to select studies: The study must 1) conduct and present a statistical examination of child pre-treatment characteristics as possible predictors or moderators of treatment outcome, 2) be a treatment study with randomized controlled designs or controlled designs. Prevention studies were excluded, 3) be published in peer-reviewed journals from 1985 onwards, 4) target psychological treatments. Pharmacotherapy or combination studies were excluded, 5) be conducted in an outpatient setting, i.e. studies of inpatient treatment were excluded, 6) include a minimum of 50 subjects in post treatment analysis, 7) the study population must consist of children and/or adolescents 4 - 18 years of age, 8) include children and adolescents with a primary diagnosis of anxiety or depression with or without comorbid conditions. All diagnostic categories pertaining to Anxiety disorders according to DSM IV and ICD 10 criteria were included, excluding Post Traumatic Stress Disorder (PTSD) and Obsessive Compulsive Disorder (OCD). All diagnostic categories for unipolar depression according to DSM IV and ICD 10 criteria were included, excluding bipolar depressive disorder, 9) the study must report an outcome measure of anxiety and/or depression symptoms from multiple informants and/or diagnostic status of anxiety and/or depression, 10) the outcome measures must be conducted post-treatment. Studies reporting follow-up or maintenance outcomes only were excluded. In addition, only predictors and moderators that have been examined in more than one study were included in order to increase confidence in specific findings and to be able to compare the results of studies.

Data analyses:

A meta-analytic review would have been preferable for the summary of findings. Due to the variability in methodologies used to investigate predictors and moderators in the studies, both in terms of outcomes used, use of different informants, employing different statistical analyses and statistical corrections, it was not feasible to calculate pooled effect sizes for the degree of association between variables. The analytic strategy adopted for the review in paper 1 was the box-score method, in which counts of statistically significant results were compared to non-significant results.

Summary of studies:

A total of 44 treatment trials were found. Thirty-two published papers based on 21 treatment studies for anxiety disorders and twelve published papers based on eight treatment studies for depression were selected for review. An overview of key characteristics of the studies is presented in Table 1 (anxiety studies) and Table 2 (depression studies) in paper 1. All anxiety treatment studies represent comparisons of different protocols and formats of cognitive behaviour therapy (CBT) with other active treatments, or with a waitlist control condition. Only five anxiety studies were conducted in community mental health clinics, while the rest were carried out in university outpatient clinics. Among the publications on depression ten compared different protocols and formats of CBT with other active treatments or waitlist controls, and three publications investigated the effectiveness of Interpersonal Therapy (IPT-A) compared to treatment as usual. Four of the studies were conducted in university outpatient clinics, one was conducted in community mental health clinics, and two studies had a mixed procedure with some patients treated in community outpatient clinics and some in university-based clinics.

Methods of papers 2 and 3

Sampling and participants

The Child and Adolescent Mental Health Services North (CAMHS North) Study was described in the introduction. All individuals aged 5 to 18 years referred to the Alta Child and Adolescent Mental Health Outpatient Service at the Finnmark Hospital Trust (CAMHS Alta) and to the Silsand Child and Adolescent Mental Health Outpatient Service at the University Hospital of Northern Norway (CAMHS Silsand) were eligible for inclusion to the CAMHS North study. Alta is a town and administration center for the municipal of Alta, situated in the Western part of the county of Finnmark. Alta is also the largest town in Finnmark, and an important administrative, commercial and educational intersection for this northern most county in Norway. In 2002, when the study was initiated, the number of inhabitants in Alta was 17159. Silsand is a village close to the town of Finnsnes, which is situated in the municipal of Lenvik in the central part of the county of Troms. Finnsnes is an administrative and commercial center for the region of central Troms, and in 2004 when the study was initiated at CAMHS Silsand the number of inhabitants in the region was 11080. Both centres cover semirural and rural areas.

Among the 320 clients eligible for this part of the multi-centres study, only 276 patients had data for the relevant change measures on one or more measurement occasions. A total of 284 patients were registered in the CAMHS Alta and CAMHS Silsand database. Eight patients were excluded in the first step due to the lack of demographic and clinical information. Among the eight patients, the only data available was results for the Wechsler Intelligence scale (WISC III) (n = 3), for the WISC III and some anamnestic information (n = 2) and some anamnestic information (n = 3). Thus, a total of 276 individuals (CAMHS Alta N = 153, CAMHS Silsand N = 123) were included. In the next step, among the 276 patients with demographic and clinical information, three patients (1 boy, and two girls) were not included in the data analysis due to missing data for the targeted outcome measures. Thus, 273 patients had ratings registered for one or several of the outcome measures, for one or more measurement occasions.

In the studies of the present dissertation, the target group was a subsample of 84 patients with emotional disorders treated at two CAMHS in the north of Norway, CAMHS Alta (n = 56) and CAMHS Silsand (n = 28). Data from the self- and parent reported Strengths and Difficulties Questionnaire SDQ were not available for the entire sample. In Table 1 (paper 2 and paper 3), characteristics of the study sample with available data for the different change measures are presented. Characteristics of the CAMHS Alta, the CAMHS Silsand sample, and the multicentre sample are presented in Table A1 (supplemental material of paper 3). The sample consists of 56 girls (66.7 %) and 28 boys (33.3 %). The mean age of the sample was 12.49 years at intake and the girls (*M* = 13.21, *SD* = 2.65) were significantly older (t (82) = - 3.24, *p* < .01) than the boys (*M* = 11.04, *SD* = 3.38). Twenty-seven patients (32.2 %) were assessed as depressed (4 boys and 23 girls), 38 patients (45.2 %) as having one or more anxiety disorders (18 boys and 20 girls), and 19 patients (22.6 %) were assessed as having both depressive and anxiety problems (6 boys and 13 girls).

Sample characteristics and representativeness

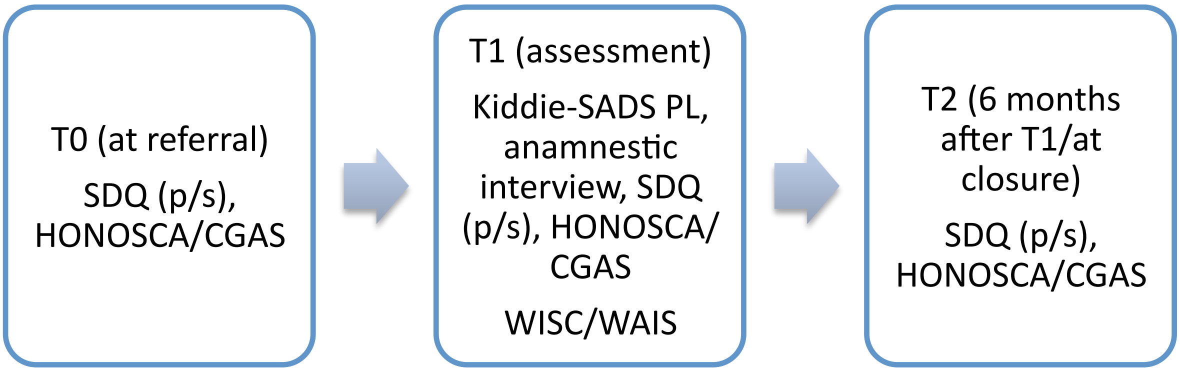
Concerning age distribution, mean age for the group of children with depression was 14.6 years at intake (88.9 % was in the age group 13-18 years). For the group with one or several anxiety disorders the mean age was 10.9 (57.8 % was in the group 6-12 years). For the group with both anxiety and depressive disorder the mean age was 12.6 (and age groups 6-12 and 13-18 were equally represented). Gender and age distribution in our sample is consistent with other studies of anxiety disorders (Kendall et al., 2010), and of depressive disorder (March et al., 2005). Also, the age and gender distribution for depression correspond with a recent longitudinal epidemiological study of depression in Central Norway (Larsson, Ingul, Jozefiak, Leikanger, & Sund, 2016). Gender and age distribution of our total sample for children and adolescents with anxiety and depressive disorders is consistent with Norwegian National statistics of CAMHS from the time period of the CAMHS North study (Lidal, 2007) and recent Norwegian national statistics for CAMHS (Krogh, 2016) and to statistics of the Child Outcomes research Consortium (CORC) (Fugard et al., 2015).

Children and adolescent with emotional disorders constituted 30.4 % of the total sample. This is similar to numbers reported from recent Norwegian National statistics of CAMHS (Krogh, 2016), as well as with statistics from the CORC (M. Wolpert et al., 2012).

Parents provided information about their ethnicity, family arrangements (living with one or both biological parents, or other caregivers), parental judicial responsibility (both parents or one parent alone) and mother tongue. The great majority of families were Norwegian, the largest proportion of children with emotional disorders were living with one of their parents, and the great majority defined Norwegian as their mother tongue.

A follow-up (T2) assessment was not completed by the clinicians for 32.1 % (n = 27) of the sample for the HONOSCA and 38.1 % (n = 32) for the CGAS. The reason for non-completion is unknown. The group of patients without follow-up data was not different from the rest of the sample as regards gender composition, mean age, age grouping and type of emotional disorder (depression, anxiety or mixed). Test statistics for comparison of the groups are presented in Table A2 and A3 (supplemental material of paper 3, attached in Appendix).

Figure 1: Procedure of the CAMHS North study



Procedure:

Children and adolescents between the ages of 5 and 18 referred to the clinics between 2002 and 2005 were asked to participate through both written and oral information. Before inclusion to the study, a written consent was obtained (see more detailed description in “Ethical considerations”). The only exclusion criteria were referral acute (e.g. concerning suicidality, suspicion of psychosis) and age < 5 years. Refusal to participate in the study did not affect type of service offered. Non-participants were assessed and treated according to the same procedures as study participants. On three occasions, measures by means of questionnaires were repeated. The SDQ (self-, mother- and father- reported), the clinician rated HONOSCA and the CGAS were administered at intake (T0), during assessment/treatment (T1) and approximately six months after T1 (T2) or at termination.

The procedure of the CAMHS North study is presented in figure 1. Standardization of the assessment procedures (T1) was one goal for the implementation of the CAMHS North Study. Also, other goals were to improve the efficiency of the assessment phase, and to complete the assessment before entering into treatment. Patients (both study participants and non-participants) were assessed through the same assessment means. The T0 was an admission meeting with the child/adolescent and parents/caregiver. The SDQ, HONOSCA and the CGAS were completed after the meeting. Informed consent was obtained. In the assessment phase (T1), the Kiddie-SADS PL was used to aid diagnostic evaluation (Ambrosini, 2000; Kaufman et al., 1997) and was conducted separately with the parents and children > 11 years of age. An anamnestic interview was conducted with the parents. The SDQ, HONOSCA and the CGAS were repeated during the assessment phase. Further assessment of mental health, development and cognitive functioning was conducted on indication. In the clinics, the diagnostic evaluation following assessment was consensus based, and a specialist in clinical psychology or psychiatry approved the evaluations. Next, the clinicians met with the child/adolescent and caregivers to give feedback concerning the evaluation, the diagnostic assessment, and to agree upon further treatment plans. It was also accustomed to arrange feedback meetings with the school.

Treatment

Treatment practice was not changed because of the observational study, and the treatment given can be classified as “treatment as usual” (TAU). The content, type and the extent of the treatment provided were not recorded in this study. Verbal accounts from clinicians’ in retrospect indicate that “treatment as usual” at the clinics was not predominated by any particular therapeutic or theoretical approaches, but was rather chosen according to the individual clinicians’ competence. Both cognitive-behavioural- and psychodynamic approaches were used, and both individual and family-based interventions were offered. For anxiety disorders, there was an agreement that exposure based cognitive-behaviour therapy should be the treatment of choice. For the treatment of depression, cognitive- behavioural approaches, narrative approaches, and family-based and systemic interventions seemed to be treatment alternatives. For both depression and anxiety disorders, medical treatment was not first line treatment, but was in a few cases offered as additional treatment.

Measures

Diagnostic interview: The Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime version (6-18) (K-SADS-PL) (Kaufman et al., 1997) is a well-established semi-structured interview that assess present and past episodes of psychopathology according to DSM-IV criteria for axis 1 diagnosis (American Psychiatric Association, 2000). The interview provides DSM-IV diagnoses of a wide range of psychiatric disorders, and was administered by a trained professional to a primary caregiver and to the child/adolescent.

In the studies of the present dissertation, we only addressed the present episode of psychopathology. The administration of the K-SADS includes a short background interview (including anamnestic information), a screening interview, targeted supplemental interviews if indicated by the screening, and assessing the diagnostic scheme. Most symptoms are evaluated on a scale from 0 (no information), to 1 (not present) through 3 (clinical threshold). Both probes and objective information is used to rate individual symptoms. After the screening interview, the clinician evaluates if further assessment is indicated through targeted supplements. Each supplement has a list of symptoms, probes, and criteria to assess current and the most severe past episodes of disorder. The summary diagnostic information is evaluated through the synthesis of the data from all sources. The following diagnosis was of interest in the present study: Major Depression, Dysthymia, Depressive disorder NOS, Adjustment disorders with depressed mood, Panic Disorder, Separation Anxiety Disorder, Simple Phobia, Social Phobia, Agoraphobia, Generalized Anxiety, Adjustment disorders with Anxiety, Obsessive Compulsive Disorder, Post-Traumatic Stress Disorder, and Acute Stress Disorder. The psychometric evidence for the Norwegian K-SADS-PL is limited. A review (H. S. Kornør, G., 2016) conclude that excellent inter-rater reliability for central diagnostic groups support clinical usage of the instrument by qualified health personnel.

For our sample the most common supplements were: 1. Affective disorders (n= 29), and/or supplement 3. Anxiety disorders (n = 34). Other supplements conducted in this sample were: Supplement 4. Behavioural disorders (n = 18), Supplement 5. Drug abuse and other disorders (n = 5), and Supplement 2. Psychotic disorders (n = 2).

The Health of the Nation Outcome Scale (HONOSCA): The HONOSCA is a 15-item clinician rated measure of mental health symptoms, and it was developed to evaluate treatment in child and adolescent mental health services (CAMHS) (Gowers et al., 1999). The HONOSCA have moderate to good inter-rater reliability (Brann, Coleman, & Luk, 2001;

Gowers et al., 1999; Hanssen-Bauer, Gowers, et al., 2007), acceptable validity (Bilenberg, 2003; Brann et al., 2001; Garralda, Yates, & Higginson, 2000; S. G. Gowers et al., 1999; Manderson & McCune, 2003), and is sensitive to change in clinical populations (Brann & Coleman, 2010; Brann et al., 2001; Holzer et al., 2006; Kisely, Campbell, Cartwright, Cox, & Campbell, 2010; Kisely, Campbell, Crossman, Gleich, & Campbell, 2007). The HONOSCA total score has proven as a good quantitative measure of clinical severity (Brann et al., 2001; S. G. Gowers et al., 1999) and correlates well with the CGAS (Lundh et al., 2013; Wolpert, 2008). In this study, the sum score of the first 13 scales, the total score (range 0 – 52), was used to assess the overall severity of mental health problems, as well as change in the overall severity over time. The scoring of the HONOSCA is based on the information that the clinicians have about the client, and is not the result of a specific interview or a specific assessment. The clinician is instructed to rate the items based on the most severe occurrence relevant for each scale for the past 2 weeks. Scores of 2 considered as the clinical cut-off. There is no clinical cut-off for the total score. Cronbach's α for the HONOSCA total score at intake was 0.52 for the current sample.

The Childrens' Global Assessment Scale (CGAS): The CGAS (Shaffer et al., 1983) is a uni-dimensional, clinician rated score of general psychosocial functioning for children and adolescent between 4-16 years, in the range from 0 (needs constant supervision) to 100 (superior functioning in all areas). In the study of the present dissertation, the most impaired level of functioning for the last month was rated. The clinician rates the CGAS based on a broad range of clinical data regarding the child and adolescent symptomatology and behaviour at home, in school and in social activities. The inter-rater reliability of the CGAS in comparable studies, have shown intra-class correlation (ICC) of ICC = 0.61 (Hanssen-Bauer, Aalen, Ruud, & Heyerdahl, 2007) and ICC = 0.73 (A. Lundh, Kowalski, Sundberg, Gumpert,

& Landen, 2010), which according to the standards presented in Shrout (1998) correspond to moderate reliability. In the study of the present dissertation, only one rater had scored the CGAS and therefore the ICC could not be calculated. The CGAS has been used as one “gold standard” for psychosocial functioning when validating other instruments (Schorre & Vandvik, 2004). National normative data have not been collected for the CGAS (Winters, Collett, & Myers, 2005). The CGAS threshold between 61 to 70 has been found to best distinguish cases from non-cases (Bird et al., 1988; Bird et al., 1990). Scores below 61 was considered a “definite” case, while scores between 61 and 70 were considered a “possible” or “probable” case, and scores above 70 were considered non-cases. A cut-off score of 70 have been shown to differentiate normal functioning from problems in need of treatment also in a German study (Steinhausen, 1987).

The Strenghts and Difficulties Questionnaire (SDQ): The SDQ (Goodman, 1997) has become a widely used tool for the screening of mental health symptoms, peer relations, prosocial behavior and functional impairment. The characteristics of the SDQ, as well as studies on the psychometric properties, are presented in the *Methods* section of paper 2. Some information is added here. The SDQ total score is considered a usefulness outcome measure (Mathai, Anderson, & Bourne, 2003; Vostanis, 2006), and a reasonable compliment to the HONOSCA (Mathai et al., 2003). The impact score has been described as a reliable and easily obtained measure of change over time (Stringaris & Goodman, 2013). The psychometric properties of the Norwegian self-report SDQ (SDQS) have been evaluated (Kornør & Heyerdahl, 2013). The review concludes that the SDQ is uncertain as a screening instrument, national norms are lacking, the internal consistencies of all scales are below adequate levels (except for the emotional problem scale), and the construct validity is acceptable. The parent version is currently being reviewed.

The HONOSCA, CGAS and SDQ as change measures

Along with the HONOSCA and the CGAS, the parent- and youth-reported SDQ (Goodman, 1997) is extensively used in CAMHS (Wolpert, 2008) as a measure of both symptom severity, domain specific severity and a separate measure of the impact of problems. In three studies, change in the HONOSCA was shown to be comparable in magnitude and direction with the Children's Global assessment scale (CGAS) (Bilenberg, 2003; Garralda et al., 2000; Gowers, Levine, Bailey-Rogers, Shore, & Burhouse, 2002), and to a lesser extent with the parent-rated Strengths and Difficulties Questionnaire (SDQ). To evaluate service level outcomes the CORC recommends the SDQ to be used alongside with e.g. goal based measures, measures that are more specific to the presenting difficulties, and HONOSCA and CGAS representing the clinician perspective (Wolpert, 2008). A more recent review of the clinical use of the SDQ self-report version conclude that its utility as a feedback measure is unclear, and that the general scope of the SDQ makes is perceived by clinicians as less relevant for individual clients (Wolpert, Cheng, et al., 2015).

Reliability

Clinicians trained and experienced with the Kiddie SADS conducted the interviews. The clinicians attended courses in conducting the Kiddie-SADS PL, and received supervision regarding the use of the interview during the study period. Inter-rater reliability tests were not done in the clinics.

Clinicians were offered training in the scoring of the HONOSCA and CGAS at several occasions during the project period. The training consisted in the scoring of vignettes followed by discussion aimed at reaching consensus of the scoring. In table 2, descriptive data for the clinicians' account of training and experience with completing the HONOSCA and CGAS is presented. Unfortunately, the ICC was not calculated for the scored vignettes during the training. Clinicians from the CAMHS Alta participated in the study by Hanssen-Bauer, Aalen, Ruud and Heyerdahl (2007). The ICC for the HONOSCA total score was 0.81 (single scales 0.47 – 0.96) and for the CGAS was 0.61. The ICC for the 10 clinics participating in the study was not reported separately, but there were not statistically significant differences in the ICC between clinics. The raters profession or experience did not influence the scores in this study. Recently, Ketil Hanssen-Bauer calculated the ICC for the clinicians participating from the CAMHS Alta on my request. A total of eight clinicians rated the HONOSCA total score (ICC = 0.80) and four of the clinicians rated the CGAS (ICC = 0.55), while the remaining four clinicians rated the Global Assessment of Psychosocial Disability (GAPD).

Table 1: Descriptive data for the clinicians account of their experience with the HONOSCA/CGAS

	T0 assessment % (n) n = 73		T1 assessment % (n) n = 71		T2 assessment % (n) n = 57	
	yes	no	yes	no	yes	no
HONOSCA/CGAS training	72.6 (61)	14.3 (12)	57.1 (48)	27.4 (23)	54.8 (46)	13.1 (11)
Completed/ handed in scored test vignettes	66.7 (56)	20.2 (17)	56 (47)	28.6 (24)	50 (42)	17.9 (15)
Scoring after consultation with colleague	7.1 (6)	81 (68)	3.6 (3)	79.8 (67)	1.2 (1)	66.7 (56)
No of HONOSCA/ CGAS scorings						
0-5	11.1 (10)		16.7 (14)		-	
6-10	13.1 (11)		16.7 (14)		2.4 (2)	
11-15	7.1 (6)		17.9 (15)		13.1 (11)	
16 – 20	8.3 (7)		11.9 (10)		17.9 (15)	
>20	46.4 (39)		19 (16)		34.5 (29)	

Handling of missing data

The term missing data refers to an intended measurement that could not be obtained. Missing data in longitudinal studies may result in unbalanced data over time, in a loss of information and may cause a reduction in the precision of variable estimates (Fitzmaurice, 2004). By using the Mixed Models approach, some of the problems with missing data can be accounted for since this method allows for the inclusion of subjects with missing data. A presumption for the use of the mixed models approach is that data must at least be *missing at random* (MAR), (i.e., missing data is related to an observed variable, but not to the variable of interest) and preferably also be *missing completely at random* (MCAR), (i.e., missing data is unrelated to any other variable in the study).

In the studies of the present dissertation, missing data includes both missing information for variables tested as predictors, and also missing data for the repeated measures. There is a difference between the CAMHS clinics, where CAMHS Alta have considerably more complete data both for predictor variables, and for the repeated questionnaires at baseline and at follow-up for all the different measures. A flow-chart of participants was presented in Figures 1 (clinician rated Health of the Nation Outcome Scale - HONOSCA) and 2 (clinician rated Children's' Global Assessment Scale - CGAS) in paper 3, and available data for all the change measures were presented in Table 1 of paper 2 and 3. The amount of follow-up data for the clinician rated measures was acceptable (CGAS 61 % and HONOSCA 73 %), while the amount of follow-up data for the child- and parent-rated measures was problematic (SDQS N = 20, 32 %; SDQM N= 24, 31 %; SDQF N= 12, 24 %). In paper 2 the SDQ was used as a measure of change despite this limitations, while in paper 3 the amount of data for predictor-outcome pairings were too small to conduct reliable analyses. Inclusion of predictor-outcome pairings with small sample sizes could have resulted in low accuracy of the parameter estimates of those pairings (Kahn, 2011). As a consequence, SDQ was removed as a change measure in paper 3, and predictor-outcome evaluations were performed on the clinician rated measures only. For the preparation of paper 3, I counted predictor –outcome pairings to evaluate which predictors that had a reasonable amount of data. First, I calculated the maximum amount of available data for the HONOSCA (n = 80; 3 x 80 = 240 measurement occasions in total), and for the CGAS (n = 82; 3 x 82 = 246). Next, available data concerning pre-treatment characteristics of the child was counted. When preparing for paper 3, several potential predictors were available and relevant, but could not be tested due to low proportions of predictor –outcome pairings. The variables omitted due to missing data were: Intelligence Quotient (IQ) from the Wechsler Intelligence Scale III (Full scale IQ, Verbal IQ and Performance IQ), having learning difficulties (yes/no), experience with physical or sexual abuse (yes/no), SDQ-item regarding duration of problems (mother- or self-report), being in contact with child protective service (yes/no), being under the care of child protective service (yes/no).

The mixed models approach do have the advantage that subjects are not omitted from the analysis with missing data. Regarding estimates of the rate of change, the individuals with data on all three measurements occasions will contribute the most information for the estimation of change rates (slopes) at the group level, individuals with data on two measurement occasions will contribute with some information, while individuals with data on one occasion do not contribute to the estimation of change rates. The representativeness of the estimated change rate depends on the proportion of individuals with data for at least two measurement occasions.

Selection procedure and inter-rater agreement:

Due to the problem of missing data, we had to rely on different selection procedures to identify the greatest number of relevant cases. First, we used the Kiddie-SADS interview to identify children and youth that fulfilled the criteria for a diagnosis of unipolar depression and/or a diagnosis of one or several anxiety disorder (n = 57). Two raters, Toril Sørheim Nilsen and Siv Kvernmo, rated all the interviews independently. Bjørn Helge Handegård, calculated the inter-rater agreement. The Gwet's AC2 per disorder is presented in Table A4 (supplemental material of paper 3, see Appendix 2). Gwet's AC2 was used to assess inter-rater agreement since it has been shown to provide more stable inter-rater agreement coefficients (Gwet, 2008; Wongpakaran, Wongpakaran, Wedding, & Gwet, 2013). The inter-rater agreement, as measured by Gwet's AC2 ranged from 0.90 (Major depression) to 0.99

(Panic Disorder), which according to Landis and Koch (1977) agreement criteria correspond to the range of “almost perfect” agreement (0.81-1.00). It is important to note that Landis and Koch (1977) presents their divisions as arbitrary, but useful benchmarks of agreement. There is ongoing debate about the use of such cut-offs (e.g Hallgren, 2012).

Furthermore, cases with disparate ratings were discussed and consensus based diagnoses were set. Finally, cases with missing data for the Kiddie-SADS, but with a registered axis 1 diagnosis of depression and/or anxiety, were selected (n = 27). In the clinics, diagnoses were consensus based and evaluated by a specialist in clinical psychology or psychiatry.

Ethical considerations

Research on child and adolescent patients may be one of the fields that require the most careful reflections on ethical issues. The Regional Committee for Medical Research Ethics (REK), and the Norwegian Social Science Data Services approved the study. The project was evaluated to involve minimal risk, to provide knowledge of vital importance that could not be obtained by studying other groups, and to have a favourable risk/benefit ratio. These issues were evaluated regularly throughout the project period. Although the study was approved by the REK, this does not guarantee that ethical pitfalls and dilemmas may not arise during the project, some being out of the control and some out of the awareness of the clinical researcher.

Informed consent is standard in all medical research, and in the case of research where children are participants require special attention. Before inclusion in the CAMHS North study, written consent was obtained at intake. For children < 12 years, both parents or caregivers consented. Adolescents between 12 and 16 consented together with their parents/caregivers. Adolescents > 16 could consent themselves. Information about the research project was given both orally and in writing during the intake meeting at the clinics. Unfortunately, the total number of individuals asked to participate and the number of refusals was not recorded. According to verbal accounts from the staff, very few individuals refused to participate. Refusal to participate in the study did not affect the service offered, and nonparticipants were assessed and treated through the same procedures as study participants.

Statistical analysis of paper 2 and 3

The statistical analyses in paper 2 and 3 were performed using Statistical Package for the Social Sciences (SPSS) version 22. Longitudinal multilevel analysis, also known as the mixed models approach, was used in paper 2 and 3 to evaluate the research question 1 and 2 (paper 2) and research question 1 (paper 3). Further descriptions of the use of the mixed models approach for the specific research questions are presented in the *Methods* sections of paper 2 and 3.

The reliable change index (RCI) is a statistics estimating if changes in psychometric measures may represent real change or occur by measurement error (Jacobson & Truax, 1991). The reliable change index (RCI) is a simple statistical computation, which estimate if changes in two psychometric measures are probably due to measurement error or can be accounted by other factor, such as a clinical intervention. The RCI index it's a useful method to assess if changes in psychometric measures may represent real change or occur by measurement error. When evaluating change from the perspective of clinical significance we adopted the account of Jacobson and Truax (1991). In their two-fold method, individual patients' response to treatment is evaluated both according to statistical reliability and clinical significance. For further description of the use of this approach in our study is presented in paper 2.

Definitions of the different parameters of the multilevel analysis are presented in Table 10 of the Supplemental material of paper 3 (see Appendix 2).

Summary of papers

Paper I

Predictors and moderators of outcome in child and adolescent anxiety and depression: a systematic review of psychological treatment studies.

Analyses and results

A literature search was performed using several databases, and resulted in 44 published studies (32 anxiety studies and 12 depression studies) meeting predefined methodological criteria. Ten client demographic- (age, gender, ethnicity, IQ) and clinical factors (duration, type of diagnosis, pre - treatment severity, comorbidity) were examined across studies. None of the pre-treatment demographic variables were found to consistently predict or moderate treatment outcome. Higher levels of baseline symptomatic severity were associated with worse outcome in several anxiety and depression treatment trials. There was also some evidence that comorbid anxiety and depression can be a negative prognostic of treatment response.

Discussion

The lack of association between demographic variables and treatment outcome was the most consistent finding across studies. Several interpretations of the finding were offered. Firstly, one plausible interpretation could be that psychosocial treatments for anxiety and depression in children and youths are equally effective for boys and girls, and for young children and adolescents. Secondly, most studies relied on small samples and this may have hindered the detection of small to medium effects.

Baseline symptom severity (3 of 4 studies) were negative prognostics for depression treatment outcome, while for the anxiety treatment studies this was found in a smaller proportion of studies (2 of 6 studies). Several interpretations of the findings were offered. Firstly, the apparent difference in the impact of the severity factor for anxiety and depression may be due to characteristics of the anxiety treatment studies, namely that the samples in the anxiety treatment studies were too homogenous in relation to the severity range that it was not possible to detect any effects. Another interpretation was that children and adolescents with more severe levels of depression might benefit more from a combination treatment with psychosocial treatment and medication.

Findings related to the impact of comorbidity showed that general comorbidity and externalizing comorbidity were not associated with differences in treatment response, whereas the results for internalizing comorbidity was mixed. Depressive disorder with comorbid anxiety disorder had a more negative treatment prognosis in two out of five studies, while anxiety disorder with comorbid depression was related to worse treatment outcome in two out of nine studies. The findings regarding internalizing comorbidity was regarded as inconclusive, and it was pointed out that the lack of findings and the mixed results may also be due to characteristics of the extant treatment studies considering this topic.

To progress in the understanding of the factors that impact treatment response, some future directions based on the extant research literature point to the need for prioritizing the study of predictors and moderators of treatment response, of adapting a more programmatic approach with a common conceptual framework and common measurement practices, and on increased emphasis on replicating and validating research results.

Paper 2

Evaluating change in symptomatic and functional level of children and youth with emotional disorders: a naturalistic observation study.

Analyses and results

Changes in symptomatic and functional impairment for children and youth with emotional disorders treated at child and adolescent mental health outpatient services (CAMHS) in Norway, was evaluated by means of statistically significant change, by comparison with waitlist control, and through calculation of clinical significant change. The results show that children and youth with emotional disorders experience a statistically significant improvement per month during outpatient treatment according to nearly all the measures of change. The only exception was change in parent rated functional impairment, which was not statistically significant. For the clinician rated scores, change rates during active assessment/treatment were larger than during the waitlist period. Evaluating change from the perspective of clinical significance showed that only a small proportion of the subjects had change scores that were statistically reliable and clinically significant. Whether an actual change has occurred is uncertain for the majority of patients.

Discussion

Statistically significant improvement according to nearly all measures, together with elevated change rates of the clinician rated measures in the active assessment/treatment period as compared to the waitlist period, may give an impression of positive effects during CAMHS treatment. Still, any conclusion regarding the relative effect of the CAMHS intervention is limited due to the lack of a control group. Alternative explanations of the positive effects can be regression toward the mean (RTM), the tendency for high baseline scores to follow a reductionist path, and of the active assessment/treatment period being on average longer than the waitlist period.

Evaluating change from the perspective of clinical significance showed that for the majority of patients the amount of change experienced is not sufficient to be classified as “recovered” or “improved”, and for the majority the degree of change must be classified as “uncertain”. Also, the difference in the clinician, adolescent and parent account of change became apparent through this approach to the evaluation of change. The evaluation of change from the perspective of clinical significance adds valuable clinical information at the individual level. The usage of broad measure may partly explain the uncertainty of change seen in the majority of subjects. Changes may have been more pronounced if we had included more disorder specific measures.

Paper 3

Predictors of rate of change for children and youth with emotional disorders: a naturalistic observational study.

Analyses and results

Longitudinal multilevel analyses were performed to examine demographic and clinical characteristics as potential predictors of change for children and youth with emotional disorders. For the HONOSCA total score, youths with a diagnosis of depression had statistically higher symptom severity levels at baseline and significantly lower change rates as compared to youths with anxiety disorders. For the CGAS, a difference was found between the clinics in the level of functional impairment at baseline, as well as in the CGAS rate of change. All other associations between predictors and rate of change pairings were non-significant.

Discussion

The significant difference in baseline CGAS scores between the two clinics, as well as the significant difference in CGAS rate of change may reflect actual differences in the patient groups of the two clinics, it may reflect RTM since the clinic with the lowest CGAS also had a steeper rate of change, and it may reflect differences between the clinics in the way the CGAS is interpreted and scored.

The finding that children and adolescents with a diagnosis of depression were rated as having higher total symptom severity at baseline and to experience less change when compared to children with anxiety disorders, may be interpreted as a replication of findings from other studies showing that anxiety disordered youth are more likely to recover if treated.

There were no associations between demographic factors and rate of change, which is in line with the extant research literature. The finding may reflect that the service function equally well for boys and girls and for different age groups. Missing data on follow-up measures as well as the low number of participants in this study may be one reason for not detecting plausible associations between demographic factors and rates of change.

General discussion

This thesis sought to expand our knowledge concerning change during CAMHS treatment for children and adolescent with emotional disorders, and concerning pre-treatment demographic or clinical characteristics of the child as potential predictors of change. I will discuss the main findings of the studies. Alternative interpretations of the presented results and methodological issues that may impact the interpretation of the findings are highlighted. Also, suggestions for future directions for clinical care and research in clinical settings are discussed.

Outcomes in CAMHS

Different approaches may lead to different conclusions regarding the outcomes, an observation reported in paper 2, and from Child Outcomes Research Consortium (CORC) data (M. Wolpert, Goerzig, et al., 2015). In our study, results at the group level gave an impression of positive outcomes since statistically a significant improvement was found at the group level for the great majority of change measures (Paper 2). The effect sizes, as calculated through the pseudo-R^2, showed small effects, and waitlist control data indicated that change in clinician-rated measures seem to be larger during active treatment. As presented in the introduction, results from large-scale naturalistic observational studies show significant improvement in mental health outcomes after CAMHS treatment with overall modest effects (e.g Fugard et al., 2015; M. Wolpert et al., 2012; M. Wolpert, Goerzig, et al., 2015).

The lack of a control group limits conclusions regarding the relative effect of the CAMHS interventions in our study, as well as in other naturalistic observational studies. The limitations related to the inference of results in treatment studies that are not randomized controlled trials (RCT) are several. Factors such as spontaneous recovery or the natural history of a disorder (W. Lambert & Bickman, 2004), the tendency for cases to be referred when problems are at peak followed by regression toward the mean (RTM), improvements in score due to repeated questioning, are all factors known to impact outcomes. In the absence of a control group it is methodologically challenging to rule out regression towards the mean (RTM) from the overall results, and removing RTM has been associated with a large decrease in effect size (Iachina & Bilenberg, 2012).

Evaluating change from the perspective of clinical significance added important information for the interpretation of findings in our study. It is a common observation that evaluations by means of clinically significant change yield a more modest account of the effectiveness of treatments than what is implied by standard statistical comparisons (Jacobson & Truax, 1991). Our results concerning clinical significant change did not allow an optimistic view since very few individuals could be classified as recovered, improved or both, according to the more conservative account of clinical significant change. For the majority of children, the size of the change score during assessment and treatment was either too small to be regarded as reliable, or did not move from the clinical range into the normal range. Thus, as concluded in paper 2, the amount of change registered while in CAMHS treatment was clearly questionable for the majority of subjects. It was uncertain whether the differences in pre-post scores for the majority of individuals reflect change beyond arbitrary fluctuations of the instruments. The finding of few individuals obtaining clinical significant change is in line with treatment outcome studies and meta-analysis of usual care suggesting little improvement associated with such services (Garland et al., 2013; Warren et al., 2009; Weisz, 2004; Weisz, Donenberg, Han, & Weiss, 1995). As summarized by Fuggle (2015), large scale studies of usual care from the United Kingdom and the United States suggest that complete remission of a presenting problem will occur in a minority of children around 15-20 %, that 25-30 % will not improve, while between 2 and 10 % get worse during usual care. Clearly, the proportion of individuals in our study with clinical significant change was considerably lower than what has been reported in other studies of clinician-rated outcomes (L Bickman, 2008; Day & Davis, 2006; Fuggle, 2015), and studies reporting results on parent rated change (e.g Edbrooke-Childs, Jacob, Law, Deighton, & Wolpert, 2015; Warren et al., 2009; M. Wolpert, Goerzig, et al., 2015).

How can these results be interpreted? Is the outcome of the CAMHS service in our study, as well as in other studies of usual care, really that questionable? Some methodological limitations of our study and of other comparable studies, may partly explain the lack of change observed. Firstly, choice of change measures to be used may impact the results. The use of broad change measures in our study, as well as within the CORC (e.g. M. Wolpert, Goerzig, et al., 2015) and other large scale studies (Manteuffel et al., 2008), may result in less pronounced changes due to the general scope of the questionnaire. More problem-specific measures have been found to yield larger effects than broader measures (Lee, Jones, Goodman, & Heyman, 2005), as well as the use of goal-based outcomes (Edbrooke-Childs et al., 2015). For our study, changes could have been more pronounced if we had included measures of symptomatic change in anxiety and depression. Although we did evaluate change in the SDQ emotional problem scale in order to assess change in a more specific and relevant domain, the items of the SDQ emotional problem scale are not disorder-specific and are considered a broad measure of emotional problems (Lee et al., 2005).

Secondly, small effects related to usual care may also be explained by the fact that many studies rely on a limited number of outcome domains, which may underestimate the number of youth that may have improved after treatment, a point made by Brookman-Frazee, Haine, and Garland (2006). Traditionally, the outcome domains most often studied, are changes in general symptomatic- and functional impairment levels, in disorder specific symptom levels or in diagnostic status. In one study from the CORC, 77 % of parents reported reliable improvement in progression towards goals, while the corresponding number for reliable improvement was 53 % for the SDQ impact scale and 18 % for the SDQ total score (Edbrooke-Childs et al., 2015). Thus, the choice of outcome domain clearly may affect the amount of change observed. In this study, both change in progress towards goals and in the impact score corresponded to large effects, while changes in total severity corresponded to a medium effect. For several groups of children receiving care, symptomatic change may not be pronounced and the evaluation of change in other domains, such as coping skills, may be a more appropriate target (Batty et al., 2013).

Thirdly, the modest effects and the few individuals with clinical significant improvement may be partly explained by relatively low baseline SDQ, and baseline CGAS scores when compared to clinical cut-offs. Thus, the range for improvement was limited at the outset. The HONOSCA has no cut-off for the total score. For the CGAS, a cut-off score of 70 has been shown to differentiate normal functioning from problems in need of treatment (Steinhausen, 1987). In our study the *Mean* CGAS at baseline was 64.03. According to British SDQ norms, the *Mean* self-reported total score and emotional problem score at baseline is in the upper normal range, while parent-rated total scores and emotional problem score correspond to the borderline range. Screening scores in the normal and borderline range at the outset could raise the question of whether the sample is sufficiently impaired to be in specialized care. As stated in the discussion in paper 2 concerning the SDQ, the general British norms may not be well suited for a Norwegian clinical sample. To be sufficiently precise, norm data should present adjusted cut-off scores for age groups and gender (Goodman, 1997) and more culture specific norms (Obel et al., 2004). Parents and teachers from Scandinavian countries tend to rate lower scores on symptom scales and tend to under-recognize emotional problems in children and adolescents (Heiervang et al., 2008). A possible explanation is that Nordic adults must observe a higher intensity and frequency of symptoms before endorsing a corresponding item. Also, results from the CAMHS North study indicate that the usefulness of the SDQ scoring algorithm (based on British norms) in detecting mental health disorders, was only partly supported, the SDQ was considered as not sufficient for clinical purposes, and emotional disorders had the lowest sensitivity (Brøndbo et al., 2011). The clinician assigned diagnosis (based on the Development and Well-Being Assessment) was higher than the "probable" screening rate for emotional disorders based on the parent- and self-reported SDQ. I have not managed to find studies evaluating the SDQ as a screening instrument in relation to the Kiddie-SADS PL. As for the CGAS, the relatively high baseline score may reflect that problems in functioning were not fully recognized in the first intake meeting, and ratings may have been influenced by limited information. As noted in Winters, Collett, and Myers (2005) the global scope of the CGAS, the mixture of objective and subjective descriptions, and the considerable time needed to gather all relevant information may be relevant disadvantages of the CGAS when the score is used to make judgements about the eligibility to services.

Fourthly, the problem of missing follow-up data for the SDQ poses a serious threat to the validity of the result for self- and parent-rated change estimates. Thus, a clear limitations of our study is that the clinician rated measures were the only outcome data with acceptable follow-up rates, which is problematic since the clinician also is the provider of the treatment, and may be a biased observer of client status and client change (Bilsker & Goldner, 2002). Another objection against using the clinician perspective is that ratings are often based on reported instead of observed symptoms or difficulties (Wolpert, 2008). In our study, clinicians were aware of the children's status in treatment, which may have influenced the ratings of the waitlist period as compared to the active assessment/treatment period. As stated by

Brookman-Frazee et al (2006), limited effects in usual care may also be explained by the reliance on a limited number of informant perspectives in many studies, with the potential loss of information regarding change during treatment. Clearly, the clinician-rated perspective is only one point of view and possibly a biased one. The clinician perspective is important from a multi-informant perspective, but should not be the only perspective on change (Fuggle, 2015). In our study, it would have been a clear advantage if the SDQ data where more complete, both in order to evaluate change from different perspectives and to gain knowledge of agreement/disagreement concerning change. Having access to data or ratings from a more objective observer perspective would also have been a valuable source of information regarding change during CAMHS treatment.

The problem of missing follow-up data is well known in naturalistic observational studies in clinical settings (e.g Baruch & Vrouva, 2010) and may comprise inferences about the effectiveness of the service, the reliability of the results, and the generalizability of the findings. To compare, within the CORC database that have been built up over a decade only 24 % of cases have meaningful outcome data (Fleming, Jones, Bradley, & Wolpert, 2016). The high proportion of missing follow-up data is a concern although the rate of outcome data has improved over the years (Hall et al., 2013). The CORC reports a situation similar to the present study, namely that measures completed by the clinician are more frequent than measures completed by the service users and that the HONOSCA was the measure that was most often used, followed by the CGAS (Batty et al., 2013; Hall et al., 2014). This was also the case in the present study.

Thus, what can be inferred about the effectiveness of CAMHS service for children and adolescents with emotional disorders in two Norwegian CAMHS? Clearly, as the discussion concerning the methodological limitations suggests, any inference will be subject to several reservations. The main finding concerning change during CAMHS treatment is that the amount of change in total severity and functional impairment is clearly questionable for the majority of subjects. It is uncertain whether the differences in pre-post scores reflect change beyond arbitrary fluctuations of the instruments, which underlines the importance of continuing efforts to improve data quality of routine outcome measurement and to continue the evaluation of clinical outcomes of the CAMHS service. The results also call for quality improvement of the service provided for this group of children. Suggestions for actions to improve both quality improvements of the service and of routine outcome measurement will be further elaborated on in the ”Future directions for clinical practice and research in clinical settings” section.

Differences in baseline CGAS and in CGAS rate of change

The sample at CAMHS Silsand was rated as more functionally impaired at baseline and had a significantly higher CGAS rate of change than the sample at CAMHS Alta. The clinics did not differ regarding the corresponding HONOSCA scores. Is this an actual difference between the two clinics in the level of functional impairment of their respective patient groups and also a difference in the rate of change for the samples? How can this finding be interpreted? Several alternative explanations are plausible. There may be a real difference in the quality of the services that might result in improved change rates. Also, the finding of higher pre-treatment functional impairment followed by a higher rate of change may reflect what is known as regression towards the mean (RTM), the tendency for high initial scores to follow a reductionist path and being closer to the mean at follow-up (Kane, 2006).

It is recommended that no ranking and no comparison between clinics should be done without careful investigation of the local context that may influence results (Fugard et al., 2015). Thus, characteristics of the two clinics, such as staff competence, caseload, turnover, working environment, may have influenced the results. One important difference between the clinics was that the CAMHS Alta had a larger sample, and had better data quality with less missing data in comparison with the CAMHS Silsand data. Some evidence suggest that service users with poorer outcomes are less likely to respond to follow-up measures (Clark, Fairburn, & Wessely, 2008), which can lead to inflated outcomes. This may result in a situation where services that succeed in collecting more complete follow-up data can get more modest accounts of change.

Also, the comparison of pre-post scores between clinics is not recommended since such information is always influenced by factors such as case-mix, sample-size etc. One study of case-mix adjustment, suggests that the important variables that relate to outcome differences among clinics have yet to be identified, and that the reasons for the differing rates of success between clinics are not currently known (Ogles, Carlson, Hatfield, & Karpenko, 2008). Parent-rated severity at baseline was the one case-mix variable that changed the rank order of clinics. High pre-treatment severity was associated with more change and adjusting for this variable changed the ranking of some clinics. In Table S1 of Appendix 2 (supplemental material of Paper 3) the descriptive information for the samples of the two clinics is presented separately. The CAMHS Silsand sample consisted of 26 individuals, and the case mix differed from the larger CAMHS Alta sample (n = 56) in respect to age distribution (majority in the 13-18 age group), gender composition (near 50 % boys), and diagnostic grouping (half of the sample with depression). There may be some characteristics of the case mix and sample characteristics that could have influenced the results. The lower baseline CGAS and the higher CGAS rate of change of CAMHS Silsand may correspond to the observation reported by Ogles et al (2008) of higher pre-treatment severity associated with higher rates of change. As described previously, in the mixed models approach, individuals with more complete data contribute more to the estimated change rate (slope) than individuals with less data. The proportion of individuals with data on two or more measurement occasions will determine the strength or representativeness of the estimates. For the sample of 26 individuals, only 14 individuals contributed to estimated rates of change.

The clinicians´ account of prior experience and training in the HONOSCA and CGAS may be important to consider when interpreting the results. Lundh et al. (2010) found moderate inter-rater reliability (ICC = 0.73) when the CGAS was used by professionals without prior training in the use of the instrument, while the ICC of a group of five experienced raters was 0.92. Compared to the experienced raters, the health-care professionals rated the vignettes significantly higher than the experienced raters. In the same study, clinical experience and prior experience of using the CGAS among the health-care professionals did not influence agreement with expert ratings. It is argued that proper training is important when introducing new rating scales. Clinicians' account of their experience with the scoring of the HONOSCA and CGAS was presented in Table 2 for both clinics. I also looked at the descriptive data of the clinicians' accounts for the two clinics separately and found no differences in the clinicians reporting about training, scoring of vignettes, or the number of HONOSCA/CGAS ratings. Apparently, there were no obvious differences in the clinicians' account of training or experience to explain differences in CGAS ratings between the clinics.

The inter-rater reliability of the CGAS (ICC = 0.61) corresponds to moderate/fair agreement, and has been found both in a study from 10 Norwegian CAMHS and in a large international study (Hanssen-Bauer, Aalen, et al., 2007; Hanssen-Bauer, Gowers, et al., 2007). Thus, differences between the clinics may also reflect local differences at the two clinics in how the CGAS scale is implemented and scored rather than an actual difference between baseline levels and rate of change of the patient groups in the two clinics.

Predictors and moderators of change

None of the demographic variables was associated with differences in treatment change, a finding in the majority of studies included in the literature review, as well as in our study. The lack of association between demographic factors such as gender and age and change during treatment is also reported in recent treatment studies of anxiety disorders (e.g Compton et al., 2014; Emslie et al., 2011; Hudson, Keers, et al., 2015; Wergeland et al., 2016) and in a review of psychosocial treatments of depression (Weersing et al., 2016). The few studies suggesting age or gender as predictors or moderators of treatment change do not show consistent findings. Thus, our findings together with the research literature to date suggest that factors such as age and gender are not associated with differential treatment change. One interpretation of these results could be that the psychosocial treatments for anxiety and depression in children and youths seem to be equally effective for girls and boys, young children and adolescents and across ethnic backgrounds.

Non-significant associations between potential prognostic variables and change during treatment have been a main finding across studies. When looking at pre-treatment demographic and clinical characteristics of the child as predictors or moderators of change, very few variables are consistently found associated with differential treatment outcome. Small sample sizes may have prevented detection of effects and result in non-significant associations for relevant prognostic variables. Small samples were a limitation described in relation to our study and in relation to the majority of the treatment literature reviewed (paper 1). In paper 1, none of the studies reviewed was originally designed to examine predictors and moderators. They were rather studies examining the relative efficacy of CBT and in a few instances for IPT. Thus, in both paper 1 and 3, statistical analyses performed on small samples may partly explain the lack of effects.

The problem of missing follow-up data, and missing information on potential predictor variables are other limitations that may result in difficulty in detecting effects in the studies of this dissertation. In paper 3, due to the problem of missing data and the small simple size, some of the tested predictor-moderator pairings relied on what is considered a minimum criterion for available data pairings in such analysis. Both the number of subjects in the sample as well as the available data for predictor-outcome pairings decides whether there is sufficient power to detect effects.

In our study, a possible limitation when examining prognostic factors was that we had a relatively heterogeneous group, with an age-range between 5 to 18 years, including individuals with depression, anxiety or both conditions. In addition, the study sample was too small to allow for analysis within subgroups. Thus, our sample may have been too heterogeneous to be able to detect any meaningful associations between prognostic factors and change. A limitation noted for previous studies on predictors and moderators concerned an opposite limitation, namely that relatively homogenous samples (e.g. age range, ethnicity, symptomatic levels) have resulted in a restricted range of target variables. For the depression treatment studies, one important limitation when examining age as a prognostic factor is that age has been tested only within samples of adolescents and not across the age range from childhood through adolescence. The null finding may therefore be partially due to restriction of range (Weersing et al., 2016). The restriction of age range prevails for the youth depression treatment literature at large. The findings on predictors, moderators, and mediators of treatment effects are largely based on samples of depressed adolescents. For the anxiety treatment studies, the relatively small variability in anxiety levels of children in previous studies, and many studies with small samples, may have hindered the detection of effects.

Several of the more recent studies on anxiety treatment are large-scale effectiveness studies that are both powered to detect effects and include more heterogeneous samples (e.g Compton et al., 2014; Hudson, Keers, et al., 2015). Another important limitation for both the depression and anxiety treatment literature assessing predictors and moderators is that the great majority of findings are drawn from trials testing the effects of CBT. Results from RCT studies about predictors and moderators of treatment change may not apply to naturalistic settings since both the treatment, the setting and client characteristics tend to differ according to some important characteristics (e. g Villabo, Cummings, Gere, Torgersen, & Kendall, 2013). Thus, this may limit the ecological validity of the findings.

As noted in paper 1, potential predictors and moderators of treatment outcome are numerous. Characteristics of the parental context (marital satisfaction, psychopathology), the familial context (family conflict, family support, socioeconomic status, living situation), child characteristics (e.g. gender, age), characteristics of the therapist (e. g experience, theoretical orientation, emphatic skills), the treatment process (therapeutic alliance, type of treatment), characteristics of the service (case-mix, case load, turnover, working environment), are all examples of potential predictors and moderators of treatment outcome. In our study, the predictor variables tested were pre-treatment demographic and clinical characteristics of the child. Other potential predictors, such as family characteristics or characteristics of the treatment, were omitted due to the lack of data for those variables.

A further limitations in the literature of both depression and anxiety treatments, is the somewhat restricted number of predictors that are studied systematically. Another plausible explanation for the finding of overall non-significant associations could be that relevant variables are not examined systematically yet. The types of variables studied are mainly the same as those studied with adults (Lundkvist-Houndoumadi et al., 2014). Developmentally related variables and functioning such as emerging social skills and friendship quality may be relevant prognostic factors for children. In the depression treatment field, future research may also capitalize on advances in the field of affective neuroscience, which can help to enhance our understanding of prognostic factors (McMakin et al., 2012).

Higher symptom severity and lower rate of change in depression than in anxiety

Depressed individuals were rated by clinicians as having higher symptom severity levels at baseline, and to experience less change in symptom severity as compared to individuals with anxiety disorders. Differences were not apparent in regards to functional impairment at baseline and in rate of change. The finding of less change in symptom severity for this group may be inferred by different explanations. The finding was discussed in paper 3, and is in line with research showing that anxiety disordered youths are more likely than depressive youths to recover if treated. Meta-analysis of CBT have found higher remission rates for youth anxiety disorder (57 % for CBT and 35 % for placebo) (Cartwright-Hatton, Roberts, Chitsabesan, Fothergill, & Harrington, 2004; James et al., 2015), than for depression (48 % for CBT and 34 % to placebo) (Watanabe, Hunot, Omori, Churchill, & Furukawa, 2007). A meta-analysis of the selective re-uptake inhibitor (SSRI) fluoxetine, showed a response rate of 61 % for depressed youth (50 % response to placebo) and 69 % response rate for anxiety disordered youth (39 % response to placebo) (Bridge et al., 2007).

Differential change rates may be due to several factors and their interaction. The depressed and the anxiety disorder groups are different regarding several characteristics, such as gender and age distribution, as well as symptom severity level. Characteristics of the depressive disorder group were that 85 % are girls and 89 % were in the age group between 13 and 18. The duration of problems was rated similarly in the two groups, but we do not have information about the recurrence of problems or prior episodes. As an example, in paper

1, we referred to one study that found interaction effects between demographic (age, gender) and clinical factors (severity) on treatment change. Mufson (2004) found interaction effects between older age and elevated symptom severity, and they found improved treatment gains for this group when treated with interpersonal therapy for Adolescents (IPT-A) compared to treatment as usual (TAU).

Several characteristics not investigated here may have an impact on the rate of change. Depression is considered a motivational disorder, characterized by thoughts of hopelessness and helplessness. Such characteristics may influence the young person's investment and belief in treatment. Also, familial-and relational problems (Garoff, Heinonen, Pesonen, & Almqvist, 2012), adverse life events and situation are often associated with the development and maintenance of depression. We lack information regarding other potential differences between the two groups, e.g. trauma experience, family context, or the presence of other risk factors known for the development of the two disorders (e.g Wilkinson, 2009), which may also impact change during treatment. According to Thapar, Collishaw, Pine and Thapar (2012), the strongest risk factors for adolescent depression is a family history of depression and exposure to psychosocial stress. There is strong evidence that family factors are important for the development, maintenance and course of depression in children and adolescents, although few studies have investigated treatment interventions targeting the family context (Restifo & Bogels, 2009). Some suggest that an increased focus on the family context and on relationship factors may be a way to improve treatment outcome for child and adolescent depression treatment (e.g Garoff et al., 2012; Restifo & Bogels, 2009). As was mentioned in the introduction, the results comparing child-only and family interventions for depression is mixed, but with some findings indicating that treating parent depression may be a promising approach for improving adolescent depression (Gunlicks & Weissman, 2008). Questions can be raised as to whether such factors are addressed and resolved sufficiently for the young person as part of treatment. From my own experience as a clinician working in the field, there is a strong consensus regarding what constitutes the best treatment for anxiety disorders, namely exposure-based cognitive behaviour interventions. My impression is that there is no strong consensus of what constitutes the best practice for depression. Cognitive behaviour therapy seems to the treatment of choice for many clinicians, but there appears to be a greater uncertainty regarding what should be done to alleviate depression. It is my impression as a clinician that depression is a condition that is regarded as both difficult to classify as well as difficult to treat.

Several reviews and meta-analyses point to a strong need for more intervention research in the field of depression treatment for children and adolescents (Cox et al., 2014; Dolle & Schulte-Körne, 2013). In their meta-analysis, Weisz et al. (2006) found that the overall effects of psychotherapy for depression lagged behind the effects of psychotherapy for other disorders in childhood and adolescence. One conclusion from this meta-analysis is that there is considerable room for improvement in psychological treatments for depression. Compared to the anxiety literature, where several candidates mediator variables have been identified, the results from the depression literature suggest that more groundwork is needed in order to understand which mechanisms are influencing change (Chu & Harrison, 2007).

Reflections on methodological issues

Several methodological limitations have been considered in relation to the interpretation of the results of the studies of the present dissertation. The problem of missing data, small sample size, the lack of a control group, clinician bias in the reporting of change, the lack of information regarding relevant predictors, are methodological issues that have already been discussed. In the following section some further methodological limitations are presented and discussed.

Selection bias and representativeness

The sample of children and adolescents included in this study comprises those with available diagnostic data, who fulfilled criteria for an anxiety and/or depressive disorder within a larger data set. The sample consisted of 27 depressed individuals, 38 individuals with anxiety disorders and 19 individuals with both anxiety and depressive disorders. In relation to the total sample of 273 individuals, the sample with emotional disorders constituted 30.8 %, which is a similar prevalence as numbers presented from the CORC database (Wolpert et al., 2016). Also, the sample has similar gender and age distribution as reported in other studies of this group (see Methods section). The individuals were enrolled at the two clinics during a time period of 4 years. Thus, for both the total sample enrolled and for the subsample with emotional disorders, not all individuals referred to the centers were included in the study. For our study, individuals were only included if they had ratings for at least one of the change measures, for at least one or more measurement occasions. A potential threat to the representativeness issue is that there is no data available concerning how many patients were eligible at the outpatient clinics and concerning the characteristics of referred individuals who did not participate in the study.

Studying anxiety and depression together versus separately

At the outset of the study, it was an intention to study children with anxiety disorders, depressive disorders and the group with both disorders separately, in order to assess potential differences in change metrics as well as in prognostic factors. The problem of missing data and a small study sample hindered such intentions. There are arguments, both pros and cons, as to whether one should study anxiety and depression jointly (Tandon, Cardeli, & Luby, 2009; Wilkinson, 2009). Commonalities and differences between anxiety and depression have been described in the literature and in the introduction section. Even though the broad band classification of internalizing/externalizing dimension, or as emotional disorder, are useful models of psychopathology, the study of individual diagnostic categories is important and relevant especially for clinical services (Achenbach et al., 2016). From his review of internalizing disorders in children and adolescents, Wilkinson (2009) concluded that anxiety and depressive disorders, despite their commonalities, should be investigated as separate disorders. For the studies in this dissertation, it would have been a clear advantage if we could have addressed our research questions by studying the diagnostic groups both jointly and separately.

Strengths of the study

The study was of a naturalistic observational type, with few exclusion criteria, which may enhance the ecological validity of the findings. The measures selected in the CAMHS North study correspond with the measures promoted and implemented by the CORC (M. Wolpert, 2008) in the UK, and also with Australian clinic studies (Patterson et al., 2006), which facilitates comparison. The design and procedure of the study also follows the procedure used by the CORC. We found few other studies reporting findings regarding predictors of change in CAMHS outpatient settings. We assessed the impact of multiple potential predictors, both demographic and clinical.

Future directions for clinical practice and research in clinical settings

Based on the findings from this naturalistic observational study and based on my experience with doing research in a clinical setting, three questions concerning future directions warrants some attention. Firstly, due to several methodological limitations affecting the interpretations of the results in our study, as well as in some other naturalistic observational studies, how can clinical studies in naturalistic settings improve the quality of the data? Secondly, our study along with other studies about change during CAMHS treatment gives reason for concern regarding the effectiveness of such services: What can be done to improve outcomes in CAMHS? Finally, considering the fact that so few prognostic factors for anxiety and depression treatment outcomes are known: How can the research on predictors and moderators of treatment outcome be improved?

Several issues regarding effort to improve data quality in naturalistic observation studies are described in the literature. Which strategies may be used that will enhance the rate of return from children, caregivers, and clinicians? For the clinicians part, lack of available resources for the collection of clinic data is one important barrier, also noted as relevant within services that are members of the CORC (Hall et al., 2013). Many factors, such as workload, few resources, etc may partly explain these completion rates. In order to improve clinician completion rates, the incentives and the clinical usefulness for completing the outcome measurement are important (L Bickman, 2008). Collecting routine outcome measurement also requires the commitment and leadership of the service management, and also a commitment from clinical and administrative staff. The collection of outcome data must be a prioritized and valued task within a busy work schedule.

Due to the comparatively lower completion rate by service users, their perspective of change is under-reported (Hall et al., 2013), which was also the case in the studies of this dissertation. Thus, it is important to find ways to improve response rates of service users. Whether data is collected during treatment or post-treatment only may impact response rate since the dropout rates among adolescents tend to be high. Therefore it can be wise to collect several data waves with shorter time intervals to ensure follow-up data from a larger proportion of individuals. With technological developments, the option of inputting data online electronically, may contribute to increased completion rates by both clinicians and service users. Instantaneous feedback on the ratings and visual presentations of change data that can be used by the patient, caregivers and therapist together in the session, may make such data a more integral part of the service provided.

What can be done to improve outcomes in CAMHS? This is of course a complex issue, and I will only comment on some issues that can be promising venues for quality improvement. Firstly, success in improving the quality of a system of care was described in Daleiden et al (2006). The results showed beneficial effects of TAU for child symptoms and functioning, and child outcomes got better during the period of a quality improvement reform (from 1996 to 2005) in child mental health services. In addition, the average length of service was reduced by 40 to 60 %. Admission and discharge characteristics of the child population was relatively stable during the same time period

Secondly, routine measurement along with frequent feedback from service users has been shown to improve outcomes of CAMHS (L. Bickman, Kelley, Breda, de Andrade, & Riemer, 2011; Gondek, Edbrooke-Childs, Fink, Deighton, & Wolpert, 2016; M. J. Lambert & Shimokawa, 2011). Another promising venue for quality improvement can be the integration of outcome data into everyday clinical practice and decision-making (Edbrooke-Childs et al., 2015). Bickman et al. (2008) argued that improvement of mental health services requires that measurements are administered frequently, which are relevant for the ongoing treatment and also provide feedback to the users of the information. In a pre-follow-up design clinicians may never see the follow-up data and do not get the chance to adjust their practice according to the data or use the data in collaboration with clients (M. Wolpert, 2014). It is important that clinicians receive feedback rapidly and supervision in interpreting the data and in contextualizing outcomes with other data.

Bridging the gap between research and clinical practice have been a goal within child mental health for decades (Weisz et al., 1995). Efforts to improve outcomes by disseminating evidence-based treatments to CAMHS continue to be important goals within the field. Evidence-based treatments protocols have tended to standardize instead of individualize treatment and assessment. There is an ongoing shift within mental health care from a focus on evidence-based treatments towards *personalized interventions* (Ng & Weisz, 2016). Personalized interventions build on evidence-based methods in translating findings from treatment trials into individual treatment plans, tailoring treatment to the needs of the individual. In mental health care there is currently insufficient empirical basis for the full implementation of such a strategy. Thus, more research in several areas is needed to translate knowledge from generic findings on the group level to clinical decision-making at the individual level. This applies both to evidence-based treatments and to usual care treatment.

How can the study of predictors and moderators of treatment be improved? Conducting studies with a sufficient number of participants that are representative of the target population, having complete data for relevant variables, and including outcome data that target relevant domains of change, seem to be important steps. The prevailing finding of non-significant associations between potential predictors and treatment change may suggest that relevant variables have not yet been assessed properly. As commented on earlier, one important step further could be to identify relevant prognostic factors for children and adolescent, such as developmentally related variables and to study those variables systematically. Examples of developmentally related variables and functioning mentioned earlier were emerging social skills, and friendship quality. Factors that are known to impact the development and maintenance of disorders may also be relevant prognostic factors for treatment change. To enhance our understanding of prognostic factors knowledge is also needed concerning negative effects of treatments, about long-term effects of treatments, regarding interaction effects of multimodal treatments (e.g. psychological-, medical and psychosocial interventions). Further, we also need knowledge concerning which can be expected change trajectories for a specific treatment, as well as knowledge concerning predictors, moderators and mediators of different change trajectories (e.g Peris et al., 2015). This may enable therapists to set more realistic expectations during the treatment process, and to assess progress more accurately through the different phases of the treatment. For the future, advances in the field of affective neuroscience may extend our understanding of how targeting specific predictors and moderators may alter these results (McMakin et al., 2012). For example, when targeting anhedonia (the inability to experience pleasure from activities usually found enjoyable) the use of neuroimaging and behavioural assessments can help to evaluate if underlying positive affective systems are altered following the intervention, which again may result in improving remission rates and recovery among youth

Conclusions

The current study adds to the limited knowledge of rate of change and predictors of rate of change in symptom severity and functional impairment for children and adolescents with emotional disorders treated within CAMHS. In conclusion, existing studies of pre-treatment patient variables as predictors and moderators of anxiety and depression treatment outcome provide little consistent knowledge concerning predictors of change during CAMHS treatment, or during RCT. Currently, there is little data that clinicians can use to guide decisions concerning who might benefit more and who more likely would benefit less to treatment (e.g Higa-McMillan et al., 2016), other than high pre-treatment symptom severity that is found to be a negative prognostic factor both within anxiety and depression treatment studies. Demographic factors such as age and gender do not seem to be a prognostic factor for differential change during treatment, neither within naturalistic observation studies within

CAMHS nor in RCT studies. Comorbid anxiety and depression also tend to be a predictor of less change, but the findings are inconclusive. In our study, children with depression had both higher total level of severity at intake and experienced less change during CAMHS treatment. This finding may support the notion that depressive disorder has a worse prognosis than anxiety disorder when treated, but may also have other plausible explanations. The results presented here point to a particular need to improve clinical care for depressed children and adolescents. Finally, rates of change during CAMHS treatment show overall improvement, but very few individuals experience change that can be defined as clinically significant. For the majority of individuals in our study, the degree of change is best classified as uncertain. Efforts to improve these services seem warranted.

The results of the studies of this dissertation point to the importance of evaluating change during CAMHS treatment and to study change through different approaches and from the perspective of various informants. The focus on prognostic factors for change during CAMHS service provision is also an important future venue. In the future, there is a need of well-planned and carefully monitored studies in naturalistic settings that keep addressing the research questions raised above.

References:

Achenbach, T., Ivanova, M., Rescorla, L., Turner, L., & Althoff, R. (2016). Internalizing/Externalizing Problems: Review and Recommendations for Clinical and Research Applications. *Journal of the American Academy of Child and Adolescent Psychiatry, 55*(8), 647-656. doi:10.1016/j.jaac.2016.05.012

Agerup, T., Lydersen, S., Wallander, J., & Sund, A. M. (2014). Longitudinal course of diagnosed depression from ages 15 to 20 in a community sample: patterns and parental risk factors. *Child Psychiatry and Human Development, 45*(6), 753-764. doi:10.1007/s10578-014-0444-8

Ambrosini, P. J. (2000). Historical development and present status of the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS). *Journal of the American Academy of Child and Adolescent Psychiatry, 39*(1), 49-58. doi:10.1097/00004583-200001000-00016

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders : DSM-IV-TR*. In (pp. XXXVII, 943 s.).

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders : DSM-5*. In (pp. XLIV, 947 s.).

Angold, A., Costello, E. J., & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry, 40*(1), 57-87.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182.

Baruch, G., & Vrouva, I. (2010). Collecting Routine Outcome Data in a Psychotherapy Community Clinic for Young People: Findings from an Ongoing Study. *Child and Adolescent Mental Health, 15*(1), 30-36. doi:10.1111/j.1475-3588.2009.00531.x

Batty, M. J., Moldavsky, M., Foroushani, P. S., Pass, S., Marriott, M., Sayal, K., & Hollis, C. (2013). Implementing routine outcome measures in child and adolescent mental health services: from present to future practice. *Child and Adolescent Mental Health, 18*(2), 82-87. doi:10.1111/j.1475-3588.2012.00658.x

Beesdo, K., Knappe, S., & Pine, D. S. (2009). Anxiety and Anxiety Disorders in Children and Adolescents: Developmental Issues and Implications for DSM-V. *Psychiatric Clinics of North America, 32*(3), 483-+. doi:DOI 10.1016/j.psc.2009.06.002

Beesdo, K., Pine, D., Lieb, R., & Wittchen, H. (2010). Incidence and Risk Patterns of Anxiety and Depressive Disorders and Categorization of Generalized Anxiety Disorder. *Archives of General Psychiatry, 67*(1), 47-57. doi:10.1001/archgenpsychiatry.2009.177

Beidel, D. C., & Alfano, C. A. (2011). *Child anxiety disorders a guide to research and treatment*. In (pp. 1 online resource (xix, 524 s.).

Bennett, K., Manassis, K., Walter, S. D., Cheung, A., Wilansky-Traynor, P., Diaz-Granados, N., . . . Wood, J. J. (2013). Cognitive-behavioural therapy age effects in child and adolescent anxiety: An individual patient data metaanalysis. *Depression and Anxiety, 30*(9), 829-841. doi:10.1002/da.22099

Bickman, L. (2008). A measurement feedback system (MFS) is necessary to improve mental health outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*(10), 1114-1119. doi:10.1097/CHI.0b013e3181825af8

Bickman, L., Kelley, S. D., Breda, C., de Andrade, A. R., & Riemer, M. (2011). Effects of Routine Feedback to Clinicians on Mental Health Outcomes of Youths: Results of a Randomized Trial. *Psychiatric Services, 62*(12), 1423-1429.

Bilenberg, N. (2003). Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA)--results of a Danish field trial. *European Child and Adolescent Psychiatry, 12*(6), 298-302. doi:10.1007/s00787-003-0343-1

Bilsker, D., & Goldner, E. M. (2002). Routine outcome measurement by mental health-care providers: is it worth doing? *Lancet, 360*(9346), 1689-1690. doi:10.1016/S0140-6736(02)11610-2

Bird, H. R., Canino, G., Rubiostipec, M., Gould, M. S., Ribera, J., Sesman, M., . . . Moscoso, M. (1988). Estimates of the prevalence of childhood maladjustment in a community survey in Puerto-Rico - The use of combined measures. *Archives of General Psychiatry, 45*(12), 1120-1126.

Bird, H. R., Yager, T. J., Staghezza, B., Gould, M. S., Canino, G., & Rubiostipec, M. (1990). Impairment in the epidemiologic measurement of childhood psychopathology in the community. *Journal of the American Academy of Child and Adolescent Psychiatry, 29*(5), 796-803. doi:10.1097/00004583-199009000-00020

Birmaher, B., Arbelaez, C., & Brent, D. (2002). Course and outcome of child and adolescent major depressive disorder. *Child and Adolescent Psychiatric Clinics of North America, 11*(3), 619-+.

Brann, P., & Coleman, G. (2010). On the meaning of change in a clinician's routine measure of outcome: HoNOSCA. *Australian & New Zealand Journal of Psychiatry, 44*(12), 1097-1104. doi:10.3109/00048674.2010.513037

Brann, P., Coleman, G., & Luk, E. (2001). Routine outcome measurement in a child and adolescent mental health service: an evaluation of HoNOSCA. The Health of the Nation Outcome Scales for Children and Adolescents. *Australian & New Zealand Journal of Psychiatry, 35*(3), 370-376. doi:anp890 [pii]

Bridge, J. A., Iyengar, S., Salary, C. B., Barbe, R. P., Birmaher, B., Pincus, H. A., . . . Brent, D. A. (2007). Clinical response and risk for reported suicidal ideation and suicide attempts in pediatric antidepressant treatment: a meta-analysis of randomized controlled trials. Journal Of The American Medical Association*, 297*(15), 1683-1696. doi:297/15/1683 pii]10.1001/jama.297.15.1683

Brofoss, K. E., Larsen, F., Friis, S., & Norges forskningsråd. (2009). E*valuering av Opptrappingsplanen for psykisk helse (2001-2009) : sluttrapport - syntese og analyse av evalueringens delprosjekter*.

Brookman-Frazee, L., Haine, R., & Garland, A. (2006). Measuring outcomes of real-world youth psychotherapy: Whom to ask and what to ask? P*sychiatric Services, 57(*10), 1373-1375. doi:10.1176/appi.ps.57.10.1373

Brøndbo, P. H., Mathiassen, B., Martinussen, M., Heiervang, E., Eriksen, M., Moe, T. F., . . . Kvernmo, S. (2011). The strengths and difficulties questionnaire as a screening instrument for norwegian child and adolescent mental health services, application of UK scoring algorithms. C*hild and Adolescent Psychiatry and Mental Health, 5,* 32. doi:10.1186/1753-2000-5-32

Campbell, T. C. (2005). An introduction to clinical significance: An alternative index of intervention effect for group experimental designs. J*ournal of Early Intervention, 27(*3), 210-227. doi:10.1177/105381510502700307

Cartwright-Hatton, S., McNicol, K., & Doubleday, E. (2006). Anxiety in a neglected population: prevalence of anxiety disorders in pre-adolescent children. C*linical Psychology Review, 26(*7), 817-833. doi:10.1016/j.cpr.2005.12.002

Cartwright-Hatton, S., Roberts, C., Chitsabesan, P., Fothergill, C., & Harrington, R. (2004). Systematic review of the efficacy of cognitive behaviour therapies for childhood and adolescent anxiety disorders. British Journal of Clinical Psychology*, 43(*Pt 4), 421-436. doi:10.1348/0144665042388928

Chu, B. C., & Harrison, T. L. (2007). Disorder-specific effects of CBT for anxious and depressed youth: A meta-analysis of candidate mediators of change. C*linical Child and Family Psychology Review, 10,* 352-372. doi:10.1007/s10567-007-0028-2

Chu, B. C., Skriner, L. C., & Zandberg, L. J. (2013). Shape of Change in Cognitive Behavioral Therapy for Youth Anxiety: Symptom Trajectory and Predictors of Change. J*ournal of Consulting and Clinical Psychology, 81(*4), 573-587. doi:10.1037/a0033390

Clark, D. M., Fairburn, C. G., & Wessely, S. (2008). Psychological treatment outcomes in routine NHS services: a commentary on Stiles et al. (2007). P*sychological Medicine, 38(*5), 629-634. doi:10.1017/s0033291707001869

Cohen, J. (1988). S*tatistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, N. J.: Laurence Erlbaum.

Compton, S. N., Peris, T. S., Almirall, D., Birmaher, B., Sherrill, J., Kendall, P. C., . . . Albano, A. M. (2014). Predictors and Moderators of Treatment Response in Childhood Anxiety Disorders: Results From the CAMS Trial. J*ournal of Consulting and Clinical Psychology, 82(*2), 212-224. doi:10.1037/a0035458

Costello, E. J., Egger, H. L., & Angold, A. (2005). The developmental epidemiology of anxiety disorders: Phenomenology, prevalence, and comorbidity. C*hild and Adolescent Psychiatric Clinics of North America, 14(*4), 631-+. doi:10.1016/j.chc.2005.06.003

Costello, E. J., Erkanli, A., & Angold, A. (2006). Is there an epidemic of child or adolescent depression? J*ournal of Child Psychology and Psychiatry, 47(*12), 1263-1271. doi:10.1111/j.1469-7610.2006.01682.x

Costello, E. J., Mustillo, S., Erkanli, A., Keeler, G., & Angold, A. (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. Archives of General Psychiatry*, 60(*8), 837-844. doi:10.1001/archpsyc.60.8.837

Cox, G. R., Callahan, P., Churchill, R., Hunot, V., Merry, S. N., Parker, A. G., & Hetrick, S. E. (2014). Psychological therapies versus antidepressant medication, alone and in combination for depression in children and adolescents. C*ochrane Database Syst Rev, 11,* CD008324. doi:10.1002/14651858.CD008324.pub3

Cummings, C. M., Caporino, N. E., & Kendall, P. C. (2014). Comorbidity of anxiety and depression in children and adolescents: 20 years after. P*sychological Bulletin, 140(*3), 816-845. doi:10.1037/a0034733

Daleiden, E. L., Chorpita, B. F., Donkervoet, C., Arensdorf, A. M., & Brogan, M. (2006). Getting better at getting them better: Health outcomes and evidence-based practice within a system of care. J*ournal of the American Academy of Child and Adolescent Psychiatry, 45(*6), 749-756. doi:DOI 10.1097/01.chi.0000215154.07142.63

Day, C., & Davis, H. (2006). The effectiveness and quality of routine child and adolescent mental health care outreach clinics. British Journal of Clinical Psychology*, 45(*Pt 4), 439-452. doi:10.1348/014466505X79986

Deighton, J., Argent, R., De Francesco, D., Edbrooke-Childs, J., Jacob, J., Fleming, I., . . . Wolpert, M. (2016). Associations between evidence-based practice and mental health outcomes in child and adolescent mental health services. C*linical Child Psychology and Psychiatry, 21(*2), 287-296. doi:10.1177/1359104515589637

Dolle, K., & Schulte-Körne, G. (2013). The treatment of depressive disorders in children and adolescents. *Deutsches Ärzteblatt International, 110(*50), 854-860. doi:10.3238/arztebl.2013.0854

Edbrooke-Childs, J., Jacob, J., Law, D., Deighton, J., & Wolpert, M. (2015). Interpreting standardized and idiographic outcome measures in CAMHS: what does change mean and how does it relate to functioning and experience? C*hild and Adolescent Mental Health, 20(*3), 142-148. doi:10.1111/camh.12107

Emslie, G. J., Kennard, B. D., & Mayes, T. L. (2011). Predictors of treatment response in adolescent depression. P*ediatric annals, 40(*6), 300-306. doi:doi: 10.3928/00904481-20110512-05

Emslie, G. J., Mayes, T. L., Laptook, R. S., & Batt, M. (2003). Predictors of response to treatment in children and adolescents with mood disorders. P*sychiatric Clinics of North America, 26(*2), 435-+. doi:10.1016/s0193-953x(02)00110-7

Esbjorn, B. H., Hoeyer, M., Dyrborg, J., Leth, I., & Kendall, P. C. (2010). Prevalence and co-morbidity among anxiety disorders in a national cohort of psychiatrically referred children and adolescents. J*ournal of Anxiety Disorders, 24(*8), 866-872. doi:10.1016/j.janxdis.2010.06.009

Ewing, D. L., Monsen, J. J., Thompson, E. J., Cartwright-Hatton, S., & Field, A. (2015). A Meta-Analysis of Transdiagnostic Cognitive Behavioural Therapy in the Treatment of Child and Young Person Anxiety Disorders. Behavioural and cognitive psychotherapy*, 43(*5), 562-577. doi:10.1017/S1352465813001094

Fitzmaurice, G. M., Laird, N. M., Ware, J. H. (2004). A*pplied longitudinal analysis.* New Jersey: John Wiley & Sons, Inc.

Fleming, I., Jones, M., Bradley, J., & Wolpert, M. (2016). Learning from a Learning Collaboration: The CORC Approach to Combining Research, Evaluation and Practice in Child Mental Health. A*dministration and Policy in Mental Health and Mental Health Services Research, 43(*3), 297-301. doi:10.1007/s10488-014-0592-y

Forti-Buratti, M. A., Saikia, R., Wilkinson, E. L., & Ramchandani, P. G. (2016). Psychological treatments for depression in pre-adolescent children (12 years and younger): systematic review and meta-analysis of randomised controlled trials. European Child and Adolescent Psychiatry doi:10.1007/s00787-016-0834-5

Fugard, A., Stapley, E., Ford, T., Law, D., Wolpert, M., & York, A. (2015). Analysing and reporting UK CAMHS outcomes: an application of funnel plots. C*hild and Adolescent Mental Health, 20(*3), 155-162. doi:10.1111/camh.12086

Fuggle, P. (2015). Are clinician ratings useful in evaluating outcomes in Child and Adolescent Mental Health Services (CAMHS)? A study of a continuous series of 1446 cases from an inner city CAMHS. J*ournal of Evaluation in Clinical Practice, 21(*4), 626-632. doi:10.1111/jep.12353

Garland, Bickman, L., & Chorpita, B. F. (2010). Change What? Identifying Quality Improvement Targets by Investigating Usual Mental Health Care. *Administration and Policy in Mental Health and Mental Health Services Research, 37(*1-2), 15-26. doi:10.1007/s10488-010-0279-y

Garland, Haine-Schlagel, R., Brookman-Frazee, L., Baker-Ericzen, M., Trask, E., & Fawley-King, K. (2013). Improving Community-Based Mental Health Care for Children: Translating Knowledge into Action. A*dministration and Policy in Mental Health and Mental Health Services Research, 40(*1), 6-22. doi:10.1007/s10488-012-0450-8

Garoff, F. F., Heinonen, K., Pesonen, A. K., & Almqvist, F. (2012). Depressed youth: treatment outcome and changes in family functioning in individual and family therapy. J*ournal of Family Therapy, 34(*1), 4-23. doi:10.1111/j.1467-6427.2011.00541.x

Garralda, M. E., Yates, P., & Higginson, I. (2000). Child and adolescent mental health service use. HoNOSCA as an outcome measure. B*ritish Journal of Psychiatry, 177,* 52-58.

Gondek, D., Edbrooke-Childs, J., Fink, E., Deighton, J., & Wolpert, M. (2016). Feedback from Outcome Measures and Treatment Effectiveness, Treatment Efficiency, and Collaborative Practice: A Systematic Review. *Administration and Policy in Mental Health and Mental Health Services Research, 43(*3), 325-343. doi:10.1007/s10488-015-0710-5

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. J*ournal of Child Psychology and Psychiatry and Allied Disciplines, 38(*5), 581-586. doi:10.1111/j.1469-7610.1997.tb01545.x

Gordon, M., Antshel, K. M., & Lewandowski, L. (2012). Predictors of treatment outcome in a child and adolescent psychiatry clinic: A naturalistic exploration. C*hildren and Youth Services Review, 34(*1), 213-217. doi:10.1016/j.childyouth.2011.10.004

Gowers, S., Levine, W., Bailey-Rogers, S., Shore, A., & Burhouse, E. (2002). Use of a routine, self-report outcome measure (HoNOSCA-SR) in two adolescent mental health services. Health of the Nation Outcome Scale for Children and Adolescents. British Journal of Psychiatry*, 180,* 266-269.

Gowers, S. G., Harrington, R. C., Whitton, A., Lelliott, P., Beevor, A., Wing, J., & Jezzard, R. (1999). Brief scale for measuring the outcomes of emotional and behavioural disorders in children. Health of the Nation Outcome Scales for children and Adolescents (HoNOSCA). British Journal of Psychiatry*, 174,* 413-416.

Gunlicks, M. L., & Weissman, M. M. (2008). Change in child psychopathology with improvement in parental depression: A systematic review. J*ournal of the American Academy of Child and Adolescent Psychiatry, 47(*4), 379-389. doi:10.1097/CHI.0b013e3181640805

Gwet, K. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. B*ritish Journal of Mathematical & Statistical Psychology, 61,* 29-48. doi:10.1348/000711006X126600

Hall, C. L., Moldavsky, M., Baldwin, L., Marriott, M., Newell, K., Taylor, J., . . . Hollis, C. (2013). The use of routine outcome measures in two child and adolescent mental health services: a completed audit cycle. B*MC Psychiatry, 13,* 270. doi:10.1186/1471-244X-13-270

Hall, C. L., Moldavsky, M., Taylor, J., Sayal, K., Marriott, M., Batty, M. J., . . . Hollis, C. (2014). Implementation of routine outcome measurement in child and adolescent mental health services in the United Kingdom: a critical perspective. E*uropean Child & Adolescent Psychiatry, 23(*4), 239-242. doi:10.1007/s00787-013-0454-2

Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. T*utor Quantitative Methods of Psychology, 8(*1), 23-34.

Hanssen-Bauer, K., Aalen, O. O., Ruud, T., & Heyerdahl, S. (2007). Inter-rater reliability of clinician-rated outcome measures in child and adolescent mental health services. Administration and Policy in Mental Health and Mental Health services Research*, 34(*6), 504-512. doi:10.1007/s10488-007-0134-y

Hanssen-Bauer, K., Gowers, S., Aalen, O. O., Bilenberg, N., Brann, P., Garralda, E., . . . Heyerdahl, S. (2007). Cross-national reliability of clinician-rated outcome measures in child and adolescent mental health services. Administration and Policy in Mental Health and Mental Health services Research*, 34(*6), 513-518. doi:10.1007/s10488-007-0135-x

Heiervang, E., Goodman, A., & Goodman, R. (2008). The Nordic advantage in child mental health: separating health differences from reporting style in a cross-cultural comparison of psychopathology. J*ournal of Child Psychology and Psychiatry, 49(*6), 678-685. doi:10.1111/j.1469-7610.2008.01882.x

Statens Helsetilsyn (2000). Psykiatriske poliklinikker: En evaluering av arbeidsformer og produktivitet innen voksen- og barne- og ungdomspykiatrien. Statens helsetilsyn.

Helsetilsyn, S., & World Health Organization. (1999). I*CD-10 : psykiske lidelser og atferdsforstyrrelser : kliniske beskrivelser og diagnostiske retningslinjer.* Oslo: Universitetsforl.

Higa-McMillan, C., Francis, S., Rith-Najarian, L., & Chorpita, B. (2016). Evidence Base Update: 50 Years of Research on Treatment for Child and Adolescent Anxiety. J*ournal of Clinical Child and Adolescent Psychology, 45(*2), 91-113.

Holzer, L., Tchemadjeu, I. K., Plancherel, B., Bolognini, M., Rossier, V., Chinet, L., & Halfon, O. (2006). Adolescent Drug Abuse Diagnosis (ADAD) vs. Health of Nation Outcome Scale for Children and Adolescents (HoNOSCA) in clinical outcome measurement. Journal Of Evaluation in Clinical Practice*, 12(*5), 482-490. doi:10.1111/j.1365-2753.2006.00654.x

Hudson, J. L., Keers, R., Roberts, S., Coleman, J. R. I., Breen, G., Arendt, K., . . . Eley, T. C. (2015). Clinical Predictors of Response to Cognitive-Behavioral Therapy in Pediatric Anxiety Disorders: The Genes for Treatment (GxT) Study. J*ournal of the American Academy of Child and Adolescent Psychiatry, 54(*6), 454-463. doi:10.1016/j.jaac.2015.03.018

Hudson, J. L., Rapee, R. M., Lyneham, H. J., McLellan, L. F., Wuthrich, V. M., & Schniering, C. A. (2015). Comparing outcomes for children with different anxiety disorders following cognitive behavioural therapy. B*ehaviour Research and Therapy, 72,* 30-37. doi:10.1016/j.brat.2015.06.007

Iachina, M., & Bilenberg, N. (2012). Measuring reliable change of emotional and behavioural problems in children. P*sychiatry Res, 200(*2-3), 867-871. doi:10.1016/j.psychres.2012.06.023

Ingul, J. M., Aune, T., & Nordahl, H. M. (2014). A randomized controlled trial of individual cognitive therapy, group cognitive behaviour therapy and attentional placebo for adolescent social phobia. P*sychother Psychosom, 83(*1), 54-61. doi:10.1159/000354672

Jacobson, N. S., & Truax, P. (1991). Clinical-significance- a statistical approach to defining meaningful change in psychotherapy research. J*ournal of Consulting and Clinical Psychology, 59(*1), 12-19. doi:10.1037//0022-006x.59.1.12

James, A. C., James, G., Cowdrey, F. A., Soler, A., & Choke, A. (2015). Cognitive behavioural therapy for anxiety disorders in children and adolescents. C*ochrane Database Systematic Review, 2,* CD004690. doi:10.1002/14651858.CD004690.pub4

Kahn, J. (2011). Multilevel Modeling: Overview and Applications to Research in Counseling Psychology. J*ournal of Counseling Psychology, 58(*2), 257-271. doi:10.1037/a0022680

Kane, R. L. (2006). U*nderstanding Health Care Outcomes Research* (2nd ed.). Sudbury, Massachusetts: Jones and Bartlett Publishers.

Karlsson, L., Kiviruusu, O., Miettunen, J., Heila, H., Holi, M., Ruuttu, T., . . . Marttunen, M. (2008). One-year course and predictors of outcome of adolescent depression: A case-control study in Finland. J*ournal of Clinical Psychiatry, 69(*5), 844-853.

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., . . . Ryan, N. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children Present and Lifetime version (K-SADS-PL): Initial reliability and validity data. J*ournal of the American Academy of Child and Adolescent Psychiatry, 36(*7), 980-988. doi:10.1097/00004583-199707000-00021

Kendall, P. C., Compton, S. N., Walkup, J. T., Birmaher, B., Albano, A. M., Sherrill, J., . . . Piacentini, J. (2010). Clinical characteristics of anxiety disordered youth. *Journal of Anxiety Disorders, 24(*3), 360-365. doi:10.1016/j.janxdis.2010.01.009

Kendler, K. S. (1996). Major depression and generalised anxiety disorder same genes, (Partly) different environments--Revisited. B*ritish Journal of Psychiatry, 168(*Suppl 30), 68-75.

Kerns, C. M., Read, K. L., Klugman, J., & Kendall, P. C. (2013). Cognitive behavioral therapy for youth with social anxiety: Differential short and long-term treatment outcomes. J*ournal of Anxiety Disorders, 27(*2), 210-215. doi:10.1016/j.janxdis.2013.01.009

Kisely, S., Campbell, L. A., Cartwright, J., Cox, M., & Campbell, J. (2010). Do the Health of the Nation Outcome Scales measure outcome? Canadian Journal of Psychiatry*, 55(*7), 431-439.

Kisely, S., Campbell, L. A., Crossman, D., Gleich, S., & Campbell, J. (2007). Are the Health of the Nation Outcome Scales a valid and practical instrument to measure outcomes in North America? A three-site evaluation across Nova Scotia. C*ommunity Mental Health Journal, 43(*2), 91-107. doi:10.1007/s10597-006-9067-2

Kley, H., Heinrichs, N., Bender, C., & Tuschen-Caffier, B. (2012). Predictors of outcome in a cognitive-behavioral group program for children and adolescents with social anxiety disorder. J*ournal of Anxiety Disorders, 26(*1), 79-87. doi:10.1016/j.janxdis.2011.09.002

Kornør, H., & Heyerdahl, S. (2013). Psychometric properties of the Norwegian version of Strengths and Difficulties Questionnaire, self-report (SDQ-S). Retrieved from http://www.r-bup.no/CMS/ptb.nsf/pages/sdq-s

Kornør, H. S., G. (2016). Måleegenskaper ved den norske versjonen av Schedule for Affective Disorders and Schizophrenia for School-Aged Children: Present and Lifetime Version (6-18) (Kiddie-SADS-PL). Retrieved from http://psyktestbarn.no/CMS/ptb.nsf/pages/k-sads-pl

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. A*rchives of General Psychiatry, 59(*10), 877-883.

Krogh, F. (2016). A*ktivitetsdata for psykisk helsevern for barn og unge 2015.* Retrieved from https://helsedirektoratet.no/Lists/Publikasjoner/Attachments/1168/BUP-rapport%202015.pdf:

La Greca, A. M., Silverman, W. K., & Lochman, J. E. (2009). Moving Beyond Efficacy and Effectiveness in Child and Adolescent Intervention Research. J*ournal of Consulting and Clinical Psychology, 77(*3), 373-382. doi:10.1037/a0015954

Lambert, M. J., & Shimokawa, K. (2011). Collecting Client Feedback. P*sychotherapy, 48(*1), 72-79. doi:10.1037/a0022238

Lambert, W., & Bickman, L. (2004). Child & adolescent psychiatry: the "clock-setting" cure: how children's symptoms might improve after ineffective treatment. P*sychiatric Services, 55(*4), 381-382.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. B*iometrics, 33(*1), 159-174.

Larsson, B., Ingul, J., Jozefiak, T., Leikanger, E., & Sund, A. M. (2016). Prevalence, stability, 1-year incidence and predictors of depressive symptoms among Norwegian adolescents in the general population as measured by the Short Mood and Feelings Questionnaire. Nordic Journal of Psychiatry*, 70(*4), 290-296. doi:10.3109/08039488.2015.1109137

Lee, W., Jones, L., Goodman, R., & Heyman, I. (2005). Broad Outcome Measures May Underestimate Effectiveness: An Instrument Comparison Study. C*hild and Adolescent Mental Health, 1*$_{01}$*0(*3), 143-144. doi:DOI: 10.1111/j.1475-3588.2005.00350.x

Lewinsohn, P. M., Rohde, P., & Seeley, J. R. (1998). Major depressive disorder in older adolescents: Prevalence, risk factors, and clinical implications. *Clinical Psychology Review, 18(*7), 765-794.

Lidal, E., Krogh, F., Mogseth, E., & Karstensen, A. (2007). A*ktivitetstall for psykisk helsevern for barn og unge (BUP) 2006.* Retrieved from https://helsedirektoratet.no/Lists/Publikasjoner/Attachments/514/Aktivitetsdata-for-psykisk-helsevern-for-barn-og-unge-2006-IS-1448.pdf:

Lundh, A., Forsman, M., Serlachius, E., Lichtenstein, P., & Landen, M. (2013). Outcomes of child psychiatric treatment. A*cta Psychiatrica Scandinavica, 128(*1), 34-44. doi:10.1111/acps.12043

Lundh, A., Kowalski, J., Sundberg, C. J., Gumpert, C., & Landen, M. (2010). Children's Global Assessment Scale (CGAS) in a naturalistic clinical setting: Inter-rater reliability and comparison with expert ratings. P*sychiatry Research, 177(*1-2), 206-210. doi:10.1016/j.psychres.2010.02.006

Lundkvist-Houndoumadi, I., Hougaard, E., & Thastum, M. (2014). Pre-treatment child and family characteristics as predictors of outcome in cognitive behavioural therapy for youth anxiety disorders. Nordic Journal of Psychiatry*, 68(*8), 524-535. doi:10.3109/08039488.2014.903295

Manderson, J., & McCune, N. (2003). The use of HoNOSCA in a child and adolescent mental health service. *Irish Journal of Psychological Medicine, 20(*2), 52-55.

Manteuffel, B., Stephens, R. L., Sondheimer, D. L., & Fisher, S. K. (2008). Characteristics, service experiences, and outcomes of transition-aged youth in systems of care: programmatic and policy implications. The Journal of Behavioral Health Service and Research*, 35(*4), 469-487. doi:10.1007/s11414-008-9130-6

March, J., Silva, S., Petrycki, S., Curry, J., Wells, K., Fairbank, J., . . . Team, T. (2005). The treatment for adolescents with depression study (TADS): Demographic and clinical characteristics. J*ournal of the American Academy of Child and Adolescent Psychiatry, 44(*1), 28-40. doi:10.1097/01.chi.0000145807.09027.82

March, J., Silva, S., Petrycki, S., Curry, J., Wells, K., Fairbank, J., . . . Team, T. (2004). Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression - Treatment for adolescents with depression study (TADS) randomized controlled trial. J*ama-Journal of the American Medical Association, 292(*7), 807-820.

Mathai, J., Anderson, P., & Bourne, A. (2003). Use of the Strengths and Difficulties Questionnaire as an outcome measure in a child and adolescent mental health service. A*ustralasian Psychiatry, 11(*3), 334-337. doi:doi:10.1046/j.1440-1665.2003.00544.x

McMakin, D., Olino, T., Porta, G., Dietz, L., Emslie, G., Clarke, G., . . . Brent, D. (2012). Anhedonia Predicts Poorer Recovery Among Youth With Selective Serotonin Reuptake Inhibitor Treatment-Resistant Depression. J*ournal of the American Academy of Child and Adolescent Psychiatry, 51(*4), 404-411. doi:10.1016/j.jaac.2012.01.011

Melton, T. H., Croarkin, P. E., Strawn, J. R., & McClintock, S. M. (2016). Comorbid Anxiety and Depressive Symptoms in Children and Adolescents: A Systematic Review and Analysis. Journal of *Psychiatric Practice, 22(*2), 84-98. doi:10.1097/PRA.0000000000000132

Mohr, C., & Schneider, S. (2013). Anxiety disorders. E*uropean Child & Adolescent Psychiatry, 22,* S17-S22. doi:10.1007/s00787-012-0356-8

Mufson, L., Dorta, K. P., Wickramaratne, P., Nomura, Y., Olfson, M., & Weissman, M. M. (2004). A randomized effectiveness trial of interpersonal psychotherapy for depressed adolescents. A*rchives of General Psychiatry, 61(*6), 577-584.

Ng, M. Y., & Weisz, J. R. (2016). Annual Research Review: Building a science of personalized intervention for youth mental health. J*ournal of Child Psychology and Psychiatry, 57(*3), 216-236. doi:10.1111/jcpp.12470

O'Neil, K. A., & Kendall, P. C. (2012). Role of Comorbid Depression and Co-occurring Depressive Symptoms in Outcomes for Anxiety-Disordered Youth Treated with Cognitive-Behavioral Therapy. *Child & Family Behavior Therapy, 34(*3), 197-209. doi:10.1080/07317107.2012.707086

Obel, C., Heiervang, E., Rodriguez, A., Heyerdahl, S., Smedje, H., Sourander, A., . . . Olsen, J. (2004). The Strengths and Difficulties Questionnaire in the Nordic countries. *European Child & Adolescent Psychiatry, 13 Suppl 2,* II32-39.

Ogles, B. M., Carlson, B., Hatfield, D., & Karpenko, V. (2008). Models of case mix adjustment for Ohio Mental Health Consumer Outcomes among children and adolescents. *Adm Policy Ment Health, 35(*4), 295-304. doi:10.1007/s10488-008-0171-1

Patterson, P., Matthey, S., & Baker, M. (2006). Using mental health outcome measures in everyday clinical practice. *Australasian Psychiatry, 14(*2), 133-136. doi:http://dx.doi.org/10.1111/j.1440-1665.2006.02266.x

Peris, T. S., Compton, S. N., Kendall, P. C., Birmaher, B., Sherrill, J., March, J., . . . Piacentini, J. (2015). Trajectories of Change in Youth Anxiety During Cognitive-Behavior Therapy. *Journal of Consulting and Clinical Psychology, 83(*2), 239-252. doi:10.1037/a0038402

Restifo, K., & Bogels, S. (2009). Family processes in the development of youth depression: Translating the evidence to treatment. *Clinical Psychology Review, 29(*4), 294-316. doi:10.1016/j.cpr.2009.02.005

Reynolds, S., Wilson, C., Austin, J., & Hooper, L. (2012). Effects of psychotherapy for anxiety in children and adolescents: a meta-analytic review. *Clinical Psychology Review, 32(*4), 251-262. doi:10.1016/j.cpr.2012.01.005

Rohde, P., Waldron, H. B., Turner, C. W., Brody, J., & Jorgensen, J. (2014). Sequenced Versus Coordinated Treatment for Adolescents With Comorbid Depressive and Substance Use Disorders. *Journal of Consulting and Clinical Psychology, 82(*2), 342-348. doi:10.1037/a0035808

Schorre, B. E., & Vandvik, I. H. (2004). Global assessment of psychosocial functioning in child and adolescent psychiatry. A review of three unidimensional scales (CGAS, GAF, GAPD). *European Child and Adolescent Psychiatry, 13(*5), 273-286. doi:10.1007/s00787-004-0390-2

Shaffer, D., Gould, M., Brasic, J., Ambrosini, P., Fisher, P., Bird, H., & Aluwhalia, S. (1983). A childrens´ global assessment scale (CGAS). *Archives of General Psychiatry, 40(*11), 1228-1231.

Shrout, P. E. (1998). Measurement reliability and agreement in psychiatry. *Statistical Methods in Medical Research, 7(*3), 301-317.

Singer, J. D., & Willett, J. B. (2003). A*pplied longitudinal data analysis : modeling change and event occurrence.* Oxford: Oxford University Press.

Spielmans, G. I., Gatlin, E. T., & McFall, J. P. (2010). The efficacy of evidence-based psychotherapies versus usual care for youths: controlling confounds in a meta-reanalysis. Journal for the Society of P*sychotherapy Research, 20(*2), 234-246. doi:10.1080/10503300903311293

Steinhausen, H. C. (1987). Global assessment of child psychopathology. J*ournal of the American Academy of Child and Adolescent Psychiatry, 26(*2), 203-206. doi:10.1097/00004583-198703000-00014

Strauss, C. C., Last, C. G., Hersen, M., & Kazdin, A. E. (1988). Association between anxiety and depression in children and adolescents with anxiety disorders. J*ournal of Abnormal Child Psychology, 16(*1), 57-68. doi:10.1007/bf00910500

Stringaris, A., & Goodman, R. (2013). The Value of Measuring Impact Alongside Symptoms in Children and Adolescents: A Longitudinal Assessment in a Community Sample. J*ournal of Abnormal Child Psychology, 41(*7), 1109-1120. doi:10.1007/s10802-013-9744-x

Sund, A. M., Larsson, B., & Wichstrøm, L. (2011). Prevalence and characteristics of depressive disorders in early adolescents in central Norway. C*hild and Adolescent Psychiatry and Mental Health, 5,* 28. doi:10.1186/1753-2000-5-28

Tandon, M., Cardeli, E., & Luby, J. (2009). Internalizing Disorders in Early Childhood: A Review of Depressive and Anxiety Disorders. C*hild and Adolescent Psychiatric Clinics of North America, 18(*3), 593-+. doi:10.1016/j.chc.2009.03.004

Thapar, A., Collishaw, S., Pine, D. S., & Thapar, A. K. (2012). Depression in adolescence. L*ancet.* doi:S0140-6736(11)60871-4 [pii]10.1016/S0140-6736(11)60871-4

Thulin, U., Svirsky, L., Serlachius, E., Andersson, G., & Ost, L. G. (2014). The effect of parent involvement in the treatment of anxiety disorders in children: a meta-analysis. Co*gnitive Behavior Therapy, 43(3*), 185-200. doi:10.1080/16506073.2014.923928

Trask, E. V., & Garland, A. F. (2012). Are children improving? Results from outcome measurement in a large mental health system. Ad*ministration and Policy in Mental Health, 39(3*), 210-220. doi:10.1007/s10488-011-0353-0

Tuseth, A. G., Sverdrup, S., Hjort, H., & Friestad, C. (2006). Å *spørre den det gjelder. Erfaringer med bruk av Klient- og resultatstyrt terapi i psykisk helsearbeid.* Regionssenter for barn og unges psykiske helse. Helseregion Øst og Sør.

Twenge, J. M., & Nolen-Hoeksema, S. (2002). Age, gender, race, socioeconomic status, and birth cohort differences on the children's depression inventory: a meta-analysis. Journal of *Abnormal Psychology, 111(4*), 578-588.

Verduin, T. L., & Kendall, P. C. (2003). Differential occurrence of comorbidity within childhood anxiety disorders. Jo*urnal of Clinical Child and Adolescent Psychology, 32(2*), 290-295. doi:10.1207/s15374424jccp3202_15

Villabo, M. A., Cummings, C. M., Gere, M. K., Torgersen, S., & Kendall, P. C. (2013). Anxious youth in research and service clinics. Jo*urnal of Anxiety Disorders, 27(1*), 16-24. doi:10.1016/j.janxdis.2012.09.003

Vostanis, P. (2006). Strengths and Difficulties Questionnaire: research and clinical applications. *Current Opinion in Psychiatry, 19(4*), 367-372. doi:10.1097/01.yco.0000228755.72366.05 00001504-200607000-00006 [pii]

Wagner, K. D., Jonas, J., Findling, R. L., Ventura, D., & Saikali, K. (2006). A double-blind, randomized, placebo-controlled trial of escitalopram in the treatment of pediatric depression. Jo*urnal of the American Academy of Child and Adolescent Psychiatry, 45(3*), 280-288. doi:10.1097/01.chi.0000192250.38400.9e

Warren, J. S., Nelson, P. L., & Burlingame, G. M. (2009). Identifying Youth at Risk for Treatment Failure in Outpatient Community Mental Health Services. Jo*urnal of Child and Family Studies, 18(6*), 690-701. doi:10.1007/s10826-009-9275-9

Warren, J. S., Nelson, P. L., Mondragon, S. A., Baldwin, S. A., & Burlingame, G. M. (2010). Youth Psychotherapy Change Trajectories and Outcomes in Usual Care: Community Mental Health Versus Managed Care Settings. Jo*urnal of Consulting and Clinical Psychology, 78(2*), 144-155. doi:10.1037/a0018544

Watanabe, N., Hunot, V., Omori, I. M., Churchill, R., & Furukawa, T. A. (2007). Psychotherapy for depression among children and adolescents: a systematic review. Ac*ta Psychiatrica Scandinavica, 116(2*), 84-95. doi:10.1111/j.1600-0447.2007.01018.x

Weersing, V. R., Jeffreys, M., Do, M. T., Schwartz, K. T., & Bolano, C. (2016). Evidence Base Update of Psychosocial Treatments for Child and Adolescent Depression. Journal of *Clinical Child and Adolescent Psychology,* 1-33. doi:10.1080/15374416.2016.1220310

Wehry, A. M., Beesdo-Baum, K., Hennelly, M. M., Connolly, S. D., & Strawn, J. R. (2015). Assessment and Treatment of Anxiety Disorders in Children and Adolescents. Cu*rrent Psychiatry Reports, 17(7*). doi:10.1007/s11920-015-0591-z

Weisz, J. R. (2004). *Psychotherapy for children and adolescents : evidence-based treatments and case examples.* Cambridge: Cambridge University Press.

Weisz, J. R., Donenberg, G. R., Han, S. S., & Weiss, B. (1995). Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63(5*), 688-701.

Weisz, J. R., Kuppens, S., Eckshtain, D., Ugueto, A. M., Hawley, K. M., & Jensen-Doss, A. (2013). Performance of evidence-based youth psychotherapies compared with usual clinical care: a multilevel meta-analysis. *JAMA Psychiatry, 70(7*), 750-761. doi:10.1001/jamapsychiatry.2013.1176

Weisz, J. R., McCarty, C. A., & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents: A meta-analysis. *Psychological Bulletin, 132(1*), 132-149. doi:10.1037/0033-2909.132.1.132

Wergeland, G. J., Fjermestad, K. W., Marin, C. E., Bjelland, I., Haugland, B. S., Silverman, W. K., . . . Heiervang, E. R. (2016). Predictors of treatment outcome in an effectiveness trial of cognitive behavioral therapy for children with anxiety disorders. *Behavior Research and Therapy, 76,* 1-12. doi:10.1016/j.brat.2015.11.001

WHO., W. H. O. (2017). Depression Fact Sheet No 369 Updated February 2017.

Wilkinson, P. (2009). Conceptualization about internalizing problems in children and adolescents. *Ciencia & Saude Coletiva, 14(2*), 373-381.

Winters, N. C., Collett, B. R., & Myers, K. M. (2005). Ten-year review of rating functional scales, VII: Scales assessing impairment. *Journal of the American Academy of Child and Adolescent Psychiatry, 44(4*), 309-338. doi:10.1097/01.chi.0000153230.57344.cd

Wolpert, M., Jacob, J., Napoleone, E., Whale, A., Calderon, A., & Edbrooke-Childs, J. (2016). Child and parent-reported outcomes and experience from child and young people´s mental health services 2011-2015.

Wolpert, M. (2008). *Review and recommendations for national policy for England for the use of mental health outcome measures with children and young people.* http://www.ucl.ac.uk/ebpu/docs/publication_files/outcome_review_report31: Department of Children Schools and Families and Department of Health.

Wolpert, M. (2014). Uses and Abuses of Patient Reported Outcome Measures (PROMs): Potential Iatrogenic Impact of PROMs Implementation and How It Can Be Mitigated. *Administration and Policy in Mental Health and Mental Health Services Research, 41(2*), 141-145. doi:10.1007/s10488-013-0509-1

Wolpert, M., Cheng, H., & Deighton, J. (2015). Measurement Issues: Review of four patient reported outcome measures: SDQ, RCADS, C/ORS and GBO - their strengths and limitations for clinical use and service evaluation. Ch*ild and Adolescent Mental Health, 20(1*), 63-70. doi:10.1111/camh.12065

Wolpert, M., Deighton, J., De Francesco, D., Martin, P., Fonagy, P., & Ford, T. (2014a). From 'reckless' to 'mindful' in the use of outcome data to inform service-level performance management: perspectives from child mental health. Bm*j Quality & Safety, 23(4*), 272-276. doi:10.1136/bmjqs-2013-002557

Wolpert, M., Ford, T., Trustam, E., Law, D., Deighton, J., Flannery, H., & Fugard, R. J. B. (2012). Patient-reported outcomes in child and adolescent mental health services (CAMHS): Use of idiographic and standardized measures. Jo*urnal of Mental Health, 21(2*), 166-174. doi:10.3109/09638237.2012.664304

Wolpert, M., Goerzig, A., Deighton, J., Fugard, A. J. B., Newman, R., & Ford, T. (2015). Comparison of indices of clinically meaningful change in child and adolescent mental health services: difference scores, reliable change, crossing clinical thresholds and 'added value' - an exploration using parent rated scores on the SDQ. Ch*ild and Adolescent Mental Health, 20(2*), 94-101. doi:10.1111/camh.12080

Wongpakaran, N., Wongpakaran, T., Wedding, D., & Gwet, K. (2013). A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. Bm*c Medical Research Methodology, 13.* doi:10.1186/1471-2288-13-61

World Health Organization. (1992). Th*e ICD-10 classification of mental and behavioural disorders clinical descriptions and diagnostic guidelines.* In (pp. xii, 362 s.).

Zhou, X. Y., Hetrick, S. E., Cuijpers, P., Qin, B., Barth, J., Whittington, C. J., . . . Xie, P. (2015). Comparative efficacy and acceptability of psychotherapies for depression in children and adolescents: A systematic review and network meta-analysis. Wo*rld Psychiatry, 14(2*), 207-222. doi:10.1002/wps.20217

Paper I –- III