Danish Centre for Health Economics
- DaCHE

Health economic perspectives on early treatment of mental health problems in children

PhD thesis 2020 | Rasmus Trap Wolf

Title

Health economic perspectives on early treatment of mental health problems in children

Danish title

Sundhedsøkonomiske perspektiver på tidlig behandling af mentale helbredsproblemer hos børn

Author

Rasmus Trap Wolf, MSc

Academic advisors

Dorte Gyrd-Hansen, PhD, Professor

Danish Centre for Health Economics – DaCHE, Department of Public Health,

University of Southern Denmark

Pia Jeppesen, PhD, Associate Professor

Mental Health Services, The Capital Region of Denmark

The Child and Adolescent Mental Health Center Research Unit

Assessment committee

Pernille Tanggaard Andersen, PhD, Professor (Chair)

The Unit for Health Promotion Research, Department of Public Health,

University of Southern Denmark

Kristian Bolin, PhD, Professor

Department of Economics, University of Gothenburg

Søren Dalsgaard, PhD, Professor

Department of Economics and Business Economics, Aarhus University

Frontpage illustration: Kat Jayne, Pexels

This thesis has been submitted to the Graduate School of Health Sciences, University of Southern Denmark

Acknowledgements

This PhD thesis was carried out at the Danish Centre for Health Economics (DaCHE), Department of Public Health, University of Southern Denmark, and at the Child and Adolescent Mental Health Center, Mental Health Services, Capital Region of Denmark. My PhD fellowship was funded by the Capital Region of Denmark on a grant from Trygfonden. The thesis was written from August 2017 to July 2020, where I have had my daily work at both DaCHE and at the Research Unit of the Child and Adolescent Mental Health Center.

First and foremost, I would like to express my sincere gratitude to all the children, parents and families who, despite their challenges with mental health problems, have participated in the different research projects which made this thesis possible. My greatest hope for the research presented in this thesis is that it will contribute to making the present as well as the future a bit better and brighter for children who suffer from mental health problems.

I am very thankful to my supervisor, Professor Dorte Gyrd-Hansen, for the important contributions to this thesis, for guiding me in the world of academia and for introducing me to other health economists, both nationally and internationally. Thank you to Postdoc Anne Sophie Oxholm for the valuable contribution to this thesis and all the guidance in the Danish registers. Also, thank you to all my other colleagues at DaCHE for the feedback on my work, for the different collaborations and for the many cakes. Thank you to Professor Julie Ratcliffe for the contribution to the thesis and for letting me be part of the health economic research environment at Flinders University in Adelaide, Australia.

A special thanks to my supervisor, Associate Professor Pia Jeppesen, for the indispensable contribution to this thesis and for taking a chance on me by letting me be part of the Mind My Mind research group. I am very grateful for getting the opportunity to conduct a PhD and to contribute with research in such an important and interesting area. I am thankful to Pia for trusting me with the task of investigating important aspects of the Mind My Mind intervention, while at the same time expressing interest and enthusiasm about my own research proposals. Thank you to my colleagues in the Mind My Mind research group - Postdoc Louise Berg Puggaard, Research Assistant Mette Maria Agner Pedersen and Research Assistant Amanda Schulz Rottbøll, to the authors of the Mind My Mind manual, all the personnel from the municipalities and regions involved in the Mind My Mind trial and to Birthe Wielandt Houe and the Psykiatrifonden for all the hard work that made the Mind My Mind trial happen. I would also like to thank the whole advisory board of the Mind My Mind trial for all their feedback to my work and their important contributions to this thesis.

Also, I would like to thank everyone who has contributed to the Copenhagen Child Cohort 2000. I am grateful for getting the opportunity to contribute with research based on the huge amount of work that have been done establishing, maintaining and developing the data of the cohort. A special thanks to Martin Rimvall for the introduction to the cohort and the collaboration we have had.

Finally, I would like to thank my wife Ravinder for her endless patience with me and for the unconditional support - also in my decision to pursue a PhD and thereby, embracing the move to Denmark with open arms.

Rasmus Trap Wolf

Copenhagen, July 2020

Table of contents

Papers in the thesis

Paper I: Wolf RT.[1,2], Jeppesen P.[1,3], Gyrd-Hansen D.[2], The CCC2000 Study Group[17], Oxholm AS.[2]. Evaluation of a screening algorithm using the Strengths and Difficulties Questionnaire to identify children with mental health problems: A five-year register-based follow-up on school performance and healthcare use. Fonseca-Pedrero E, editor. PLOS ONE. 2019. 23;14(10). DOI: 10.1371/journal.pone.0223314

Paper II: Wolf RT.[1,3], Puggaard LB.[1], Pedersen MMA.[1], Pagsberg AK.[1,3], Silverman WK.[12], Correll CU.[13,14,15,16], Plessen KJ.[1,4], Neumer S.[10,11], Gyrd-Hansen D.[3], Thastum M.[8], Bilenberg N.[5,6], Thomsen PH.[7,8], Jeppesen P.[1,3]. Systematic identification and stratification of help-seeking school-aged youth with mental health problems - A novel approach to stage-based stepped-care. *Submitted for publication.*

Paper III: Wolf RT.[1,2], Ratcliffe J.[18], Chen G.[19], Jeppesen P.[1,3]. The longitudinal validity of the CHU9D – Evidence from a mental health trial. *Submitted for publication.*

Paper IV: Wolf RT.[1,2], Jeppesen P.[1,3], Pedersen MMA[1], Puggaard LB.[1], Thastum M.[4], Bilenberg N.[5,6], Thomsen PH. [7,8], Silverman WK.[9], Plessen KJ.[1,10], Neumer S.[11,12], Correll CU.[13,14,15,16], Pagsberg AK.[1,3], Gyrd-Hansen D.[2]. Cost-effectiveness of a transdiagnostic psychotherapy program for youth with common mental health problems. *Prepared for submission.*

[1]Child and Adolescent Mental Health Centre, Mental Health Services Capital Region of Denmark, Denmark.

[2]University of Southern Denmark, Department of Public Health, Danish Centre for Health Economics, Denmark.

[3]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark.

[4]Division of Child and Adolescent Psychiatry, Department of Psychiatry, Lausanne University Hospital CHUV, Switzerland.

[5]Department for Child and Adolescent Psychiatry, Mental Health Services in the Region of Southern Denmark, Denmark.

[6]University of Southern Denmark, Denmark.

[7]Research Center at Department for Child- and Adolescent Psychiatry, Aarhus University Hospital, Denmark.

[8]Institute of Clinical Medicine, Faculty of Health, Aarhus University, Denmark.

[9]Department of Psychology and Behavioural Sciences, Aarhus BSS, Aarhus University, Denmark.

[10]Centre for Child and Adolescent Mental Health, Oslo, Norway.

[11]Centre for Child and Youth Mental Health and Child Welfare, The Arctic University of Norway, North Norway (RKBU North), Norway.

[12]Anxiety and Mood Disorders Program, Yale Child Study Center, Yale University, School of Medicine, New Haven, CT, USA.

[13]Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Department of Psychiatry and Molecular Medicine, New York, USA.

[14]The Zucker Hillside Hospital, Department of Psychiatry, Glen Oaks, New York, USA.

[15]Feinstein Institute for Medical Research, Center for Psychiatric Neuroscience, Manhasset, New York, USA.

[16]Charité Universitätsmedizin, Department of Child and Adolescent Psychiatry, Germany.

[17]Centre for Clinical Research and Prevention, Capital Region of Denmark, Denmark.

[18]College of Nursing and Health Sciences, Flinders University, Australia.

[19]Centre for Health Economics, Monash University, Australia.

Abbreviations

CAMHS	Child and Adolescent Mental Health Services
CBA	Cost-Benefit Analysis
CBCL	The Child Behavior Checklist
CBT	Cognitive and Behavioral Therapies
CEA	Cost-Effectiveness Analysis
CHU9D	The Child Health Utility 9 Dimensions
CI	Confidence Interval
CUA	Cost-Utility Analysis
DAWBA	The Developmental And Well-Being Assessment
DCE	Discrete Choice Experiment
GPA	Grade Point Average
HRQoL	Health-Related Quality of Life
HSUV	Health State Utility Value
HTA	Health Technology Assessment
ICER	Incremental Cost-Effectiveness Ratio
MFQ	The Mood and Feelings Questionnaire
MMM	Mind My Mind
NICE	The National Institute for Health and Care Excellence
PPR	The Educational and Psychological Counselling Services
QALY	Quality-Adjusted Life Years
RCT	Randomized Controlled Trial
SCAS	Spence Children's Anxiety Scale
SDQ	The Strength and Difficulties Questionnaire
SG	Standard Gamble
TTO	Time Trade Off
WHO	World Health Organization
WTP	Willingness-To-Pay

Summary

All over the world, a large proportion of children are suffering from mental health problems. Both diagnosable mental disorders and subthreshold conditions have a significant negative impact on the life of the children, their families and the society. Despite the promising body of evidence supporting the use of cognitive and behavioral therapies (CBT) for treatment of mental health problems in children, there is very limited access to such treatments. Part of the challenge has been that the CBT-programs focus on single disorders. Therefore, wide scale access would require several programs to be implemented in order to meet the need for treatment of common mental health problems and the diverse set of symptoms children normally present with. This approach is likely to endure substantial costs due to the high costs of training and supporting infrastructure when such costs are generated individually for each program. Transdiagnostic treatment programs that address more than one type of disorder may therefore constitute an approach that is not only more effective in terms of treating the diverse mental health problems but may also be less costly. The Mind My Mind (MMM) transdiagnostic CBT-program was recently developed with the aim of providing evidence-based early treatment for school-aged children with emotional and/or behavioral problems below the threshold for psychiatric referral. The overall aim of this thesis was to contribute to the evidence base of early treatment of children with common mental health problems. This was done through four papers that evaluated: 1) new innovative ways of identifying the children in need of early treatment (paper I-II), 2) the use of a child-specific instrument for estimating Quality-Adjusted Life Years (QALY) (paper III), and 3) the cost-effectiveness of the MMM intervention (paper IV).

Paper I evaluated the potential of a new screening algorithm to identify children with sub-threshold and undiagnosed mental disorders in the need of an early treatment. A combination of historical birth cohort data and data from national registers was used in a longitudinal study design to investigate the potential for a screening of children aged 11-12 years. The screening was based on parent-reported answers to the Strength and Difficulties Questionnaire (SDQ) for children aged 11-12 years and an algorithm with a cut-off based on the distribution of scores in population surveys. The children's later school performance at end of compulsory schooling was analyzed as an outcome in addition to the children's health care service utilization during a five-year follow-up. No feedback was given to parents, teachers or health care professionals at the time of the SDQ survey. The evaluation of the screening algorithm showed that the children who screened positive had significantly poorer school outcomes when compared to the children who screened negative. The majority of the children who screened positive did not have a mental health consultation in the health care system during the five-year follow-up period. The children who screened positive and did have a mental health consultation, on average, had the first consultation after 2.5 years. Thus, the results demonstrate that there is a potential for identifying children in need of an early treatment for their mental health problems.

Paper II investigated whether a novel visitation model for school-aged children with mental health problems facilitated a systematic identification and stratification process with equal access to care. The novel stage-based stepped-care model aimed to identify children with mental health problems requiring an intervention, and to stratify the children into three groups (stages) with increasing severity of problems. This was accomplished using a self-referral two-phase stratification process involving a web-based assessment, that utilized the algorithm presented in paper I, and a semi-structured psychopathological interview of the child and parents. To assess problems with inequity in access, individual-level socioeconomic data were obtained from national registers for the participating population and for the background population. A total of 573 children

and their parents took part in the evaluation of the visitation model. The model succeeded in defining three groups of children with increasing severity of mental health problems. The assessment of the children revealed that the majority of the children had long-standing and impacting mental health problems and met the diagnostic criteria of at least one mental disorder diagnosis or were assessed as having indications of severe mental disorders. The evaluation revealed important problems in reaching children of lower mothers with low level of education, children with foreign-born parents and older school-aged children.

The objectives of **Paper III** were to examine the construct validity and the longitudinal validity of the preference-based health-related quality-of-life measure, Child Health Utility 9D (CHU9D), in children with mental health problems, using two sets of utility weights derived from an adult and an adolescent population. Data from 396 children participating in a randomized controlled trial was used in the validation. Discriminant validity and convergent validity were examined using the mental health specific SDQ and the generic KIDSCREEN-27. Longitudinal validity was assessed by examining the floor-ceiling effects, the magnitude of change over time, and the ability of the instrument to differentiate between improvements in mental health and no improvement. The findings demonstrated that CHU9D is an appropriate preference-based health-related quality of life measure for use in children with mental health problems. Furthermore, the examination of the two different sets of preference weights showed that the choice of preference weights could have a significant impact on the results of cost-utility analyses concerning mental health interventions.

In **Paper IV,** the cost-effectiveness of the transdiagnostic CBT-program MMM for children with common mental health problems was evaluated using a cost-utility analysis framework and data from a pragmatic randomized controlled trial of MMM versus management-as-usual involving 396 children. The sensitivity of the incremental QALYs to the choice of informant in reporting of health-related quality of life (child or parent) and choice of informant in deriving preference weights (adult or adolescents) were analyzed. The results showed that the MMM had positive effects on the children's health-related quality-of-life, but also that the MMM had higher costs when compared with management-as-usual. Analyses of extrapolation scenarios revealed that the cost-effectiveness of MMM intervention is highly dependent on the duration of the treatment effect, and that there is a reasonable likelihood of the MMM intervention being cost-effective if effects last beyond the observation period. Cost effectiveness was, however, shown to be highly sensitive to both the choice of informant on quality of life, as well as informant on preferences.

The four studies all have limitations, which require consideration. It is especially important that the potential problems with inequity in access are addressed. Nevertheless, the studies provide empirical evidence that supports the introduction of systematic identification, visitation and early treatment of children with mental health problems to meet a substantial unmet need. Early treatment of mental health problems in children remain a developing area in health care. The evidence presented in this thesis suggests that it is time for a change of focus from randomized controlled trials to execution and implementation of early treatment programs. The results also highlight that the estimation of QALY is sensitive to who reports on health-related quality of life as well as who's preferences the preference weights are based on. This methodological issue should in the future be considered carefully when conducting and evaluating cost-utility analyses of health care initiatives targeting children's mental health problems.

Dansk resume (Danish summary)

Mange børn i hele verden lider af mentale helbredsproblemer. Både diagnosticerbare psykisksygdomme og mentale helbredsproblemer, der ligger under tærskelgrænsen for diagnose, har signifikante negative konsekvenser for børn, deres familier og samfundet som helhed. På trods af den lovende evidens, der støtter brugen af kognitiv- og adfærdsterapi (CBT) til behandling af mentale helbredsproblemer hos børn, er der stadig meget begrænset adgang til sådanne behandlinger i praksis. En del af udfordringen består i, at CBT-programmer typisk fokuserer på en specifik type mentale helbredsproblemer. For at imødekomme behovet for behandling af almindelige mentale helbredsproblemer vil det således kræve, at flere forskellige programmer vil skulle implementeres. Denne tilgang vil formodentlig medføre betragtelige omkostninger, idet ressourcebehov relateret til uddannelse og infrastruktur vil skulle dækkes for hvert enkelt program. Transdiagnostiske behandlingsprogrammer, der adresserer mere end én type mentale helbredsproblemer, kan derfor potentielt udgøre en tilgang, der både er mere effektiv, ved at muliggøre behandling af børn med forskelligartede mentale helbredsproblemer i et samlet program, og som samtidig potentielt er mindre omkostningsfuld. Mind My Mind (MMM) er et nyt transdiagnostisk CBT-program udviklet til at levere evidensbaseret tidlig behandling til børn i skolealderen, der lider af emotionelle og/eller adfærdsmæssige mentale helbredsproblemer.

Det overordnede mål med denne afhandling var at bidrage til evidensbasen for tidlig behandling af børn med almindelige mentale helbredsproblemer. Dette blev gjort ved udarbejdelsen af fire artikler.

Artikel I evaluerede potentialet for en ny screeningsalgoritme til at identificere børn med ikke-diagnosticeret psykisksygdom eller mentale helbredsproblemer under tærskelgrænsen for at få en diagnose, der har behov for en tidlig behandling. En kombination af historisk kohorte data og data fra nationale registre blev brugt i et longitudinelt studiedesign til at undersøge potentialet for en screening af børn i alderen 11-12 år. Screeningen var baseret på forældrebesvarelser af the Strength and Difficulties Questionnaire (SDQ) og en algoritme med en tærskelværdi baseret på fordelingen af score i populationsundersøgelser. Børnenes præstation ved 9. klasses afgangseksamen blev analyseret som effektmål sammen med børnenes forbrug af offentlige sundhedsydelser i en opfølgningsperiode på fem år. Der blev ikke givet nogen feedback på SDQ-undersøgelsen til forældre, lærer eller sundhedsprofessionelle. Evalueringen af screeningsalgoritmen viste, at de børn, der testede positivt, klarede sig signifikant dårlige ved 9. klasses afgangseksamen sammenlignet med de børn, der testede negativt. Størstedelen af de børn, der testede positivt, havde i løbet af femårs opfølgningen ikke en konsultation i børne- og ungdomspsykiatrien eller hos en offentlig betalt psykolog eller psykiater i sundhedsvæsenet. De børn, der testede positivt og havde kontakt til offentlig finansieret psykolog, psykiater eller børne- og ungdomspsykiatrien, havde i gennemsnit deres første konsultation 2,5 år efter forældrene besvarede SDQ. Resultaterne demonstrerede således, at der er et potentiale for at identificere børn, som har et behov for at få en tidlig behandling for deres mentale helbredsproblemer.

Artikel II undersøgte, hvorvidt en ny visitationsmodel for børn i skolealderen med mentale helbredsproblemer baseret på en stage-based stepped-care tilgang kunne facilitere en systematisk identifikation og stratificeringsproces med lige adgang til behandling. Formålet med visitationsmodellen var at identificere børn med mentale helbredsproblemer, som havde behov for en intervention, samt stratificere børnene med problemer i tre grupper med stigende alvorlighed. Dette blev gjort ved hjælp af selvhenvisning og en stratificeringsproces i to faser. Stratificeringsprocessen indebar en web-baseret vurdering, der benyttede algoritmen

præsenteret i artikel I, og et semistruktureret psykopatologisk interview af børn og forældre. For at vurdere, hvorvidt der var problemer med ulighed i adgang, blev der benyttet socioøkonomiske data fra nationale registre på individniveau til at sammenligne børnene, der tog del i visitationsmodellen, med jævnaldrende børn i baggrundspopulationen. I alt tog 573 børn og deres forældre del i visitationsmodellen. Modellen lykkedes med at definere tre grupper med mentale helbredsproblemer af stigende alvorlighed. Vurderingen af børnene viste, at størstedelen af børnene havde langvarige mentale helbredsproblemer. Problemerne var af en sådan grad, at de fleste børn levede op til kriterierne for mindst en diagnose, eller det blev vurderet, at der var tegn på svær psykisksygdom. Evalueringen viste, at der var betydende problemer med at nå børn af mødre med lavt uddannelsesniveau, børn med anden etnisk baggrund samt ældre skolebørn.

Formålet med **artikel III** var at undersøge begrebsvaliditet og følsomhed for det præferencebaserede sundhedsrelaterede livskvalitetsmål Child Health Utility 9D (CHU9D) blandt børn med mentale helbredsproblemer. To sæt af præferencevægte til estimeringen af QALY blev undersøgt. De var baseret på henholdsvis en voksen populations præferencer og en ung populations præferencer. Data vedrørende 396 børn fra et randomiseret kontrolleret forsøg blev brugt i valideringen. CHU9Ds evne til at skelne mellem forskellige kendte grupper og konvergerende validitet blev undersøgt ved hjælp af det generiske livkvalitetsmål KIDSCREEN-27 og SDQ, der er specifikt designet til at undersøge mentale helbredsproblemer. Følsomhed blev vurderet ved at undersøge 'floor-ceiling' effekt, omfanget af ændringer over tid og CHU9Ds evne til at skelne mellem forbedring i mentalt helbred og ingen forbedring. Resultaterne viste, at CHU9D er et passende præferencebaseret sundhedsrelateret livskvalitetsmål til brug i forsøg relateret til mental sundhed. Undersøgelsen af de to sæt præferenceværdier viste desuden, at valget af præferencevægte kan have en signifikant betydning for resultaterne af cost-utility analyser, der omhandler mental sundhed.

I **artikel IV** blev omkostningseffektiviteten af det transdiagnostiske psykoterapiprogram MMM evalueret ved en cost-utility analyse, der brugte data fra et randomiseret kontrolleret forsøg af MMM sammenlignet med sædvanlig behandling. Forsøget involverede 396 børn. Det blev desuden undersøgt, hvilken betydning valget af respondent til vurdering af livskvalitet og valget af præferencevægte havde på resultatet af cost-utility analysen. Resultaterne viste, at MMM havde en positiv effekt på børnene sundhedsrelaterede livskvalitet, men også at MMM medførte højere omkostninger end sædvanlig behandling. Analyser af ekstrapolationsscenarier viste, at MMM har potentialet til at blive vurderet omkostningseffektiv, men at omkostningseffektiviteten er afhængig af varigheden af behandlingseffekten. Analyserne viste desuden, at omkostningseffektiviteten er følsom overfor, om det er forældre eller barn, der rapporterer sundhedsrelateret livskvalitet, samt om præferencevægtene er er baseret på voksnes eller unges præferencer.

De fire studier har nogle begrænsninger. Det er især vigtigt, at de potentielle problemer med ulighed i adgang til behandling bliver adresseret. Studierne bidrager dog med vigtig empirisk evidens, der støtter implementering af systematisk identificering, visitation og tidlig behandling af børn med mentale helbredsproblemer. Tidlig behandling af mentale helbredsproblemer er fortsat et udviklingsområde i sundhedsvæsenet. Resultaterne, der bliver præsenteret i denne afhandling, peger således på, at tiden er moden til et skift i fokus fra randomiserede kontrollerede forsøg til eksekvering og implementering af evidensbaserede behandlingsprogrammer. Resultaterne understreger desuden, at valget af respondent til afrapportering af sundhedsrelateret livskvalitet samt valget af præferencevægte har afgørende betydning ved estimering af QALYs. Disse metodiske problemstillinger bør i fremtiden overvejes nøje både i forbindelse med udførsel og evaluering af cost-utility analyser, der omhandler børn og unge.

Preface

This thesis consists of four papers and a synopsis.

Paper IV reports a health economic evaluation of the Mind My Mind intervention based on data from a randomized controlled trial of the intervention. The paper that includes the main reporting of the clinical results from the trial, which I co-author, has yet to be published. Submission of paper IV is withheld by the publishing of the clinical evaluation.

The synopsis consists of five sections. First, the background for conducting the research is presented. This is followed by the aims of the research. Hereafter follows a summary of the four papers. The papers are then discussed with regard to the findings, the strengths and limitations, and the future implications and directions. The synopsis ends with a conclusion on the conducted research.

There is lack of consistency in the literature when it comes to describing mental health problems. To clarify, the following definitions will be used consistently throughout this thesis: **Mental disorders** will refer to *diagnosed* mental disorders, while **mental health problems** refers to all mental health conditions independent of severity, and independent of diagnosis status, unless otherwise stated.

1. Background

Mental health problems affect a large proportion of children and adolescents (herein referred to as children) world-wide, and have been named the leading cause of years lost to disability among children by The World Health Organization (WHO) (1). A review investigating mental disorders among children found a pooled point prevalence of 13% for diagnosable mental disorders in the general population of children when looking at the evidence from epidemiological studies conducted in 27 countries around the world (2). The problem is also present in Denmark, where a recent study using comprehensive nation-wide register data from the last 25 years found that 15% of Danish children had been clinically diagnosed with a mental disorder before the age of 18 (3). In addition to the large proportion of children suffering from mental disorders, there is an unknown, but potentially large, proportion of children who suffer from mental health problems below diagnostic thresholds. Regardless of the diagnostic threshold, these common mental health problems can potentially incur a significant negative impact on the children's lives (4,5).

1.1 The costs of children's mental health problems

Several studies have been carried out to estimate the economic impact of mental disorders in children. A WHO review found eight studies from Europe on the economic impact of mental disorders among children (6). Seven of the studies were carried out in the United Kingdom. The studies all focus on specific diagnoses or specific subgroups e.g. children who have committed criminal offenses. The studies apply varying perspectives in their costing analysis. All include costs from the health care system and the educational sector, and some include costs from other sectors as well as from the criminal justice system and social services. Two studies include estimates of the lost productivity of the career, and in both studies, these costs were higher than the other sector costs including the health care sector (7,8). Based on the differences in methods, diagnoses, subgroups, and included sectors, it is no surprise that the estimated annual costs of mental disorders amongst children in the review differ by a factor seven and range from €7,376 to €52,725. Another review identified studies from the whole world and had similar findings of studies focusing on specific diagnosis, mainly autism and attention deficit hyperactivity disorder but also anxiety, depression, and conduct disorder (9). Only a smaller part of the studies identified in this review included costs from outside the health care sector. However, those that did, also identified costs from other sectors, especially the educational sector, to be the main component of the total societal costs. The authors found no clear conclusion for which type of mental disorder had the highest cost for society (9).

The majority of the cost estimates are based on data collected via postal questionnaires, interviews, or a combination of both. This method has the embedded risk of information bias, particularly recall bias. Another thing they have in common is that they all find that health care sector costs constitute only of a minor part of the total costs, while the majority of the costs comes from the education system, the criminal justice system and the lost productivity of the careers, depending on the disorder, the specific group of patients and the included costs (6,9).

A more recent study from Brazil investigated the costs of not only children with mental disorders but also children with mental health problems below the threshold of diagnosis (10). They found that although children with mental disorders carried the highest costs, sub-threshold conditions were also associated with considerably higher costs than children without mental health problems. In this study, they found that the

majority of the costs were due to health care services and costs associated to the educational sector for both the children with mental disorders and the children with sub-threshold conditions.

With one exception, the studies mentioned above focuses on a short time horizon by only estimating cost retrospectively or over a limited time period in childhood. Besides the direct economic impact of mental disorders in children, there is strong empirical evidence that mental disorders and sub-threshold conditions during childhood are associated with a profoundly higher risk of mental illnesses later in life (4,11–16). Hence, the median age of onset of the common mental disorders - anxiety, mood and behavioral disorders - is observed to occur before the age of 14 years (17). This link between childhood and adulthood mental health suggests that the total cost of children's mental health problems should include the costs related to the increased risk of mental disorders in adulthood.

Mental disorders among adults are recognized as some of the most costly diseases for society due to both health care cost and lost productivity (18). Specifically, a report from the Danish Health Authority examined 21 diseases with some of the highest burden of disease and found anxiety, depression, and schizophrenia to be among the most expensive diseases for the Danish society (19). The high societal costs of adult mental disorders and the strong association with childhood mental health problems points towards a potentially large saving if effective preventive interventions can be introduced, as such interventions could lead to lower incidence of mental health problems in children and thus, also in adulthood.

1.2 The unmet need

The number of children diagnosed with mental disorders have risen doing the last decades. The increase could be due to a number of reasons e.g. improved awareness, better detection, better access to diagnostics, actual higher incidence rates, or most likely, a combination of all these factors. No matter the causes, the situation has led to an increased pressure on Child and Adolescent Mental Health Services (CAMHS) to provide effective treatments. Studies from the United States suggest that not all children in need of mental health care services actually receive it (20,21). In Denmark, the number of children with contact to CAMHS has increased steadily over the last decade, with 49.4% more children having contact to CAMHS in 2018 compared to 2009 (22). Despite the increase in contacts, there is still a large proportion of children who do not have access to appropriate care. In England, it was estimated that only 25% of children with a diagnosable condition are actually seen by CAMHS (23). In the Capital Region of Denmark, 27% (N=1,566) of all referrals to CAHMS were rejected in 2018 (personal communication). Part of the explanation for the large number of rejections is likely to be that the referred children suffer from milder degrees of anxiety, mood, or behavioral disorders, which are not deemed severe enough for treatment at CAHMS. It is expected that milder cases are first treated in school and community settings. In Denmark, there are 98 municipalities that vary considerably in size with populations from above 500,000 to less than 10,000. There does not exist an overview of the mental health interventions offered in the different municipalities. The types and extent of interventions are, however, likely to vary considerably, and no municipality offers systematic identification of children with mental health problems or have systematic and broad access to evidence-based interventions for children with mental health problems.

1.3 Early treatment of mental health problems

Preventive interventions for children with mental disorders can be categorized in different ways. A general way is to distinguish between universal prevention, selective prevention, and indicated prevention (24). Universal prevention targets a whole population and aims to prevent development of one or more disorders. School-based programs such as the anti-bullying campaign is an example hereof. Selective prevention targets subgroups that are at a particularly high risk of developing one or more mental disorders. For example, there are programs that target children of parents with severe mental disorders such as schizophrenia and bipolar disorder, as these children are known to have a higher risk of developing mental disorders themselves. Finally, indicated prevention targets children with detectable clinical manifestations of mental disorders, who still do not meet diagnostic criteria of severe mental disorder. Indicated prevention can also be categorized as early treatment, as it aims to both treat the present symptoms that may cause impairment here and now and aims to prevent the future development of severe mental disorders.

There is a large body of evidence supporting the use of Cognitive and Behavioral Therapies (herein referred to as CBT) for treatment of mental health problems in children. CBT has been found effective for children with anxiety (25–31), depressive symptoms (26,29,32–41) and for children with disruptive behavior disorders (42,43). One of the limitations of the evidence on CBT is the lack of long-term follow-up beyond one or two years. The typical comparator in trials of CBT is the waiting list condition. It is often used to maintain participants in the trial based on the expectation that the active intervention is effective and hence, ethically correct to offer the controls after the first treatment period. The consequence of this research design is that the control group is less "contaminated" with other active treatments when compared to the experimental condition. However, the design has been criticized because being on a waiting list means that the natural course of healing may also be on hold. Also, the following treatment of the control group prevents the study of long-term effects of the experimental intervention. The few controlled studies of CBT-programs with access to mid-long term follow-up data have found beneficial and statistically significant effects up to 12 months after end of treatment (42–49). There exists several well-documented early treatment programs using CBT, for example Coping Cat (50) and Cool Kids (51) for the treatment of anxiety.

Despite the promising evidence base, there has been relatively little success in implementing these evidence-based programs on a wide scale. A major challenge in the implementation is the single-disorder focus of the programs. With the broad spectrum of such single-disorder programs, this would call for many programs with varying content to be implemented and run simultaneously leading to high costs of training and infrastructure as such costs are generated individually for each program. Transdiagnostic treatment programs that address more than one type of disorder may, therefore, constitute an approach that is not only more effective in terms of treating the diverse mental health problems (52,53), but may also be a less costly option.

Transdiagnostic programs can be categorized into two types: Programs with a unified approach that targets underlying problems across different disorder types, and programs with a modular approach that allows for individualization of treatment by the dosing and sequencing of modules to fit the individual child's mental health problems (52). The development and evaluation of modular transdiagnostic early treatment programs is still relatively new, but has shown promising results (54,55).

Evidence on the cost-effectiveness of CBT (56–62) suggests that only some programs may be cost-effective. Currently, existing evidence only includes the evaluation of programs offering universal prevention or

selective prevention, and programs which concern only patients already in contact with CAMHS. There is, thus, a lack of evidence concerning the cost-effectiveness of indicated prevention, i.e. early treatment CBT-programs aiming to meet the needs of children with common mental health problems (63). For implementation of new innovative transdiagnostic programs to be justifiable, they should not only be effective, but the costs of implementing the programs should also be considered reasonable. Evidence on cost-effectiveness is, therefore, a key argument if funding for establishment is to be secured.

1.4 Stage-based stepped-care models in children's mental health

Stage-based stepped-care models in children's mental health can be understood as a development of stepped-care. In a stepped-care approach, interventions, in increasing intensity levels, are delivered, so that an effective intervention is delivered to the patient at the lowest level of required resources. The idea is to offer low intensity interventions as the first step, and subsequently interventions with higher intensity if required, so that the costs of the interventions and the time invested in effective interventions by the child and parents are kept as low as possible (64). The stepped-care approach has been critiqued for potentially delaying access to effective treatment among children with more severe mental health problems as these children may have to go through inadequate interventions before being offered an appropriate intervention (65). A related critique involves the practicalities in how to assess the needs of the children and make the right decision about which interventions should be given to whom and when. Children typically present themselves with a diverse set of symptoms and a poorly defined disorder (66), making it difficult to fit them into stepped-care models that are often designed to target a single type of disorder (65). The stepped-care models also often lack a clear description of when which steps should be taken (67). Stage-based stepped-care models have been developed to meet these different issues while building on the general idea from stepped-care in delivering effective interventions on the lowest cost-level (65,68). The staging part focuses on the matching of an intervention with appropriate intensity to match the current needs of the child based on assessments of the mental health problems and the impact of problems across the diagnostic boundaries, i.e. using a transdiagnostic approach to improve the understanding and treatment of the developing mental health problems. The stages are defined by the possible interventions. In theory, the stage-based stepped-care models allows for the complexity of mental health problems by not necessarily focusing on specific diagnosis or diagnostic thresholds but instead, the potentials of effective interventions (68).

The implementation of stage-based stepped-care models is challenged by the organization of mental health services as there is a lack of a coherency across preventive interventions, early treatment and treatment at CAMHS. Prevention and early treatment is, in many countries, including Denmark, located outside the health care system in for example, schools and the local government (69,70), while the specialist diagnostic assessment and treatments are located in the tertiary health care system. Stage-based stepped-care models would allow for a successful integration of the sectors by providing children with a systematic assessment of needs to ensure that they receive timely intervention of appropriate intensity in the appropriate setting. In the context of early treatment, a structured and standardized approach is paramount to ensuring the timely and effective assessment to identify the children that would benefit from an early treatment. It is important that resources are not wasted by providing an intensified early treatment to children with mild mental health problems who would benefit from less demanding and costly interventions. At the same time, it is important that children identified with potentially severe mental disorders are assessed without delay and treated in specialist services like the CAMHS.

1.5 The Mind My Mind Intervention

The Mind My Mind (MMM) program was recently developed with the aim of providing evidence-based early treatment for school-aged children with emotional and/or behavioral problems below the threshold for psychiatric referral. The intervention is a transdiagnostic, manualized, modular CBT-program, with a treatment manual that contains evidence-based CBT methods and techniques. The program comprises 9-13 weekly individual sessions targeting anxiety, depression and/or behavioral problems. The weekly therapy was followed by a booster session 6–8 weeks after end of treatment. The CBT-methods in MMM are organized into modules, thereby enabling the intervention to be tailored for the youth by dosing and sequencing of the modules. Parents are engaged to support the youth in therapy, and parent management training is offered as the main part of the treatment for youth with behavioral problems. All treatment is designed to be delivered by trained local psychologist from the educational-psychological advisory service, with supervision from trained psychologist from the regional CAMHS. The development of the program was driven by the Danish foundation, Psykiatrifonden, and sponsored by the Danish foundation, Trygfonden.

A feasibility randomized controlled trial (RCT) proved sufficient acceptability of the implementation and evaluation of the program in Danish Municipalities, and it provided data for the sample size estimation for the design of a definitive RCT. The present PhD project was conducted as part of the next pragmatic, open-label, analysts-masked, parallel, two-arm, randomized superiority trial of MMM versus management as usual (MAU) for youth with common emotional and behavioral problems in four Danish municipalities. A separate goal was to evaluate the cost-effectiveness of the MMM program compared to MAU while contributing to advancing of the methodology of measuring Quality-Adjusted Life Years (QALY) in the youth population. The research was conducted by the MMM research group in the Child and Adolescent Mental Health Center, Mental Health Services - Capital Region of Denmark.

The pragmatic RCT of MMM versus MAU was carried out in four Danish municipalities (Holstebro, Vordingborg, Næstved, Helsingør). A total of 396 children were included and randomly allocated on an individual level 1:1 via independent, blinded, centralized, computer-generated allocation sequences with variable and unknown block sizes, stratified by geographical region (two strata), age-group (6–10 years or 11–16 years) and their top-problem (anxiety, depressive symptoms or behavioral problems). The clinical evaluation found the MMM intervention to be superior to MAU (71).

To be able to identify the children for the MMM RCT, a novel systematic identification and stratification model building on the principals of stage-based stepped-care was developed and implemented in the four municipalities. One of the aims of the current thesis was to evaluate whether this model could facilitate a systematic identification and stratification process with equal access to care. This is further described in section 2.

1.6 Introduction to health economic evaluation

Resource scarcity is a basic condition of health care systems due to limited resources being available and an unlimited potential use for them. With the complexity of the health care system and the constant research and development in the area of health care, there will always be potentially effective interventions and policies to fund. At the same time, resources are always limited. In a universal health care system, the budget is clearly limited by government budgets. This causes the need for prioritization: decisions need to be made on

which interventions and policies to fund, and which not to fund. Cost-of-illness studies such as the ones presented in section 1.1 can be effective in raising awareness of a problem, but has limited use when it comes to prioritization, as those kind of studies do not offer recommendations on possible solutions (63). Health economic evaluation, on the other hand, is designed to help decision-makers with prioritization by informing them of both the health effects and the economic consequences of introducing an intervention (72).

Full economic evaluations looking at both cost and output can be classified into two categories: cost-benefit analysis (CBA) and cost-effectiveness analysis (CEA), based on the choice of output, i.e. monetary or health effects, respectively. In simplified terms, the choice of CBA and CEA can be seen as the choice between a welfarist and an extra-welfarist approach. Within the welfarism paradigm, focus is on optimizing societal welfare, which in this paradigm can be defined as the sum of the individual's utilities. The utility gain can be measured in terms of willingness-to-pay, as this indicates the willingness to forego consumption of other goods. In contrast, the extra-welfarist approach focuses on maximizing health and allows for societal evaluations of health outcomes rather than evaluations of the sum of individuals' utilities (73). In a health care system prioritization context where the objective is to maximize health for a population within a restricted budget, this makes CEA the preferable type of analysis to inform decisions (72).

In a CEA, the goal is to estimate the consequences, in the form of costs and health outcomes, of introducing a new intervention by comparing it to one or more alternative interventions. This is done by calculating the ratio between the difference in costs and the difference in effects, which is termed the incremental cost-effectiveness ratio (ICER):

$$\mathrm{ICER} = \frac{\Delta\mathrm{Cost}}{\Delta\mathrm{Effect}} = \frac{\mathrm{Cost}_{\mathrm{New\ intervention}} - \mathrm{Cost}_{\mathrm{Comparator}}}{\mathrm{Effect}_{\mathrm{New\ intervention}} - \mathrm{Effect}_{\mathrm{Comparator}}}$$

For CEAs with an ICER estimate that shows the new interventions to be less costly and to have better effect, the new intervention dominates the old intervention, and the new intervention is clearly preferred. On the other hand, if the new interventions are more costly and have worse effect, the old intervention is clearly preferred. However, situations arise where the new intervention is less costly but has a worse effect than the comparator, or where, as is often the case, the new intervention is more costly but also more effective than the comparator. In these cases, it is necessary to have knowledge of the opportunity costs (the shadow price) or a defined willingness-to-pay per health effect to be able to conclude whether the new intervention can be deemed cost-effective.

In CEA, the effects are denominated as natural health units. This could, for example, be life-years saved, avoided readmissions or changes in severity of symptoms. The choice of health unit can be dictated by the health condition and the type of intervention, so it reflects the most crucial outcome for the particular decision (74). The use of such natural health unit is sufficient if the decision context is very narrow, e.g. only focusing on deciding between interventions for very similar health conditions and within a budget earmarked for that area. However, the prioritization context is typically broader and concerns making decision across different health conditions and interventions types. For this kind of prioritization, standard CEA comes up short, as it is practically impossible to define comparable opportunity costs or willingness-to-pay thresholds for all kinds of different health units. Instead, cost-utility analysis (CUA) is the preferred type of analysis for informing decision-makers. CUA can be considered a subset of CEA where the outcome is defined as a generic measure of health capturing both mortality and morbidity. This allows for comparison of the cost-

effectiveness across all types of health conditions and types of interventions (72). The most widely used outcome in CUA is the QALY, that is also increasingly being adopted in Danish health care prioritization (75).

QALYs are calculated by multiplying a patient's health state utility value (HSUV) with the time they spend in that health state. The health state utilities are weighted so that 1 represents perfect health and 0 represents death. Thus, one year in perfect health represents 1 QALY. The calculation of HSUV is typically done in two steps. In the first step, the patient of interest describes their health state based on their health-related quality-of-life (HRQoL) using a generic descriptive instrument, also called preference-based HRQoL instruments. In the second step, each health state is given a utility value based on a set of preference weights. For estimation of QALYs in the adult population, the EQ-5D-3L is a widely used preference based HRQoL instrument and is the standard of health technology assessment (HTA) institutes in many countries including England and the Nordic countries (76). When it comes to estimation of QALYs in child populations, there is both less experience and less clear guidance from the HTA institutions (77,78). Furthermore, the often used EQ-5D-3L has been found to be inadequate in estimating QALYs in specific populations of adults with severe mental disorder (79,80). This makes the choice of preference-based HRQoL instrument for estimating QALYs in a child mental health setting especially challenging.

1.7 Preference based HRQoL instruments for use in child populations

There exists several generic preference-based HRQoL instruments. A review have found that nine different instruments have been used in child populations (81). The instruments vary in both the dimensions of health, the number of dimensions, and in the description of the dimensions and the possible response categories. The instruments can be categorized into two. Those that are meant to be used in both children and adult populations and those that are adaptions from adult preference-based HRQoL instruments. The EQ-5D-Young is an example of the latter, as it is an adaption of the widely used EQ-5D-3D designed for adult populations (82,83). Unlike the eight other instruments, the Child Health Utility 9D (CHU9D) is designed exclusively from its inception for application with children. This potentially gives CHU9D an advantage over the other instruments, as the dimensions, the type of response categories, and all wordings are designed to capture the HRQoL of children based on qualitative work and testing in child populations (84).

CHU9D has nine items with five levels of severity representing nine dimensions of HRQoL: Worried, Sad, Pain, Tired, Annoyed, Schoolwork/homework, Sleep, Daily routine and Activities. Several studies have investigated the content validity, face validity, and construct validity of the CHU9D in children populations with good results (85–91). This includes both a Danish validation of the translated version of CHU9D (90). The validation studies have mainly focused on self-reported CHU9D in the age-group 7-17. An exemption is an Australian validation that examined the face validity, practicality, internal consistency and convergent validity of CHU9D in a population of children aged 5-17 years with continuous contact to CAMHS (91). They found that the parent-reported CHU9D showed good results for use in this population of children with mental disorders when compared with the Strengths and Difficulties Questionnaire (SDQ). All validations of CHU9D are, however, based on cross-sectional data. This means that the longitudinal validity of CHU9D has not yet been investigated as it requires longitudinal data. Longitudinal validity can be defined as the ability of the instrument to detect clinically important changes over time (92).

When using a preference based HRQoL instrument in a child population, another issue is whether the respondent should be the child or a proxy e.g. a parent. While the preferred source of information normally is

the patient or in this case the child (86), it is necessary to consider the cognitive ability of younger children to understand the questions and their ability to answer them adequately. This is especially an issue in a mental health population, where the perception of 'normal' might be influenced by the mental health problems themselves, but where some internalizing mental health problems, at the same time, may go unnoticed by a parent (93).

A recent review identified 35 studies with data on both children- and proxy-reported HSUVs (94). The agreement between children and parents were stronger for the more objective physical dimensions of health compared to more subjective dimensions e.g. emotions and pain. They found no consistent pattern in the directions of the difference between respondents. The authors concluded that HSUV derived from children and parents as proxies are not interchangeable and that CUAs potentially will have significantly different results depending on the respondent (94). These findings highlight the question of which respondent should be used in a CUA involving a mental health school-aged population such as the one in the MMM trial.

1.8 Preference weights for estimating QALYs in child populations

Another issue that has recently been discussed in the literature concerns whose preferences should be used when estimating the preference weights that are used to calculate QALYs in child populations (77,78). There are several possibilities of whose preference to use e.g. the general populations', health professionals' or patients'. When adults are concerned, HTA institutions typically, e.g. in England and in the Nordic countries, require that preferences of the general population are used. An argument for the use of preferences of the general adult population is that they represent the taxpayers and thus, the insurers, who ultimately fund the health care system, and whose ex ante preferences should be reflected in the design of the insurance policy. It has, however, been suggested that when children are concerned, the preferences of the general child population should be used instead of those of the general adult (87). Studies have found that the two populations do have different preferences (87,95). There are certain methodological considerations that should be taken into account when deciding whose preferences to use.

Children, and especially younger children, can have difficulties understanding the elicitation procedure required to generate utilities. There are different methods for eliciting preferences. Traditionally, the standard-gamble (SG) and the time-trade-off (TTO) have been the preferred methods (72). Both methods can be used to generate utilities on the scale where 1 is perfect health and 0 is death. The methods do, however, require that the health state of death is being considered by the subjects. The concept of death can be difficult for children to understand, and it can be argued that it is unethical to ask children to consider death, especially in the case of younger children. An alternative to SG and TTO is discrete choice experiments (DCE), where ordinal techniques are used to rank the different dimensions and health states. This procedure requires less abstract thinking for the subject, which makes it better suited for children. If used without the consideration of death, this technique does, however, require an anchoring procedure to generate utilities on the 1-0 perfect health – death scale. The anchoring can be done using different techniques, but all requires a form of elicitation which include death as a possible event (77).

If adult preferences are used, it requires an elicitation procedure where the adults are asked about preferences of others besides themselves. When eliciting preferences for preference weights used in the adult population, the adults are typically asked to imagine themselves being in the different health states. This is not in the same way possible when dealing with children's health states. For example, questions about

children's daily routines are not likely to be relevant for adults. Here, the adult would have to be asked either to imagine themselves as a child having the health state, imagine the health state for their child or grandchild, or some other scenario. While the first approach is likely to be subjected to recall-bias, the approach of the adults giving their preferences of others' health state has the problematic issue that the preferences might be based on not only the health states, but also the adult's view on the particular group being in this health state (77).

Rowen and colleagues (77) have reviewed the methods utilized to generate preference weights for existing preference-based HRQoL instruments and found that there is a considerable variation. They concluded that even though the choice of whose preferences to use is ultimately a political choice, further empirical evidence on the consequences is warranted to inform the choice (77).

For CHU9D, the original set of preference weights were derived from a sample of 300 adults representing the general adult population in the United Kingdom. SG methods was used to generate utility scores ranging between 0.3261 for the worst health state (pit-state) to 1 for perfect health (89). Another set of preference weights were generated using best-worst scaling DCE and an anchoring procedure using TTO. A representative sample of 1,982 Australian children aged 11-17 from the general population took part in the DCE, while the anchoring procedure was based on a TTO experiment with a sample of 152 young adults. The utility scores range between -0.1059 for the pit-state to 1 for perfect health (96). In a comparison of Australian preferences, using the best-worst scaling DCE and the health states defined by CHU9D, Ratcliffe and colleagues (87) found that there are differences between the values adult and children attach to the different health states. This was especially true for the dimensions, worried, sad and annoyed, for which children place a higher weight (87). With these dimensions being especially important for mental health, this could mean that the choice of preference weights is especially impactful when CHU9D is used in a mental health population.

1.9 Perspective and costs

When conducting a health economic evaluation, the perspective of the analysis is important. The aim of the analysis is to inform decision makers or stakeholders so that the limited resources can be distributed optimally, with the benefits of the implemented intervention being greater than the opportunity costs. To be able to meet this aim it is important that the perspective of the analysis is clear, as it defines which cost that should be included in the analysis. No perspective is universally more correct than others, it depends on who the analysis is meant to inform (72). Something that is regarded a cost in one perspective might not be in another. For example, in a health care sector perspective, an intervention involving an extra schoolteacher in a classroom due to a child suffering from a mental disorder would not be included as a cost, while in a societal perspective, it would be included. A societal perspective should be taken if the aim is to inform how to achieve the optimal societal welfare. In this perspective all relevant societal costs should be included in the analysis, as the aim is to make the decision that across sectors will lead to benefits greater than the opportunity costs. Often HTA institutions focus on a narrower health care sector perspective, as they aim to achieve the optimal health outcome within their limited budget. For example, NICE has a perspective that only includes costs for the National Health Service and Personal Social Services (97).

A limiting factor in including all relevant costs is, however, often the availability of data. For health economic evaluations of interventions for children, there are special challenges when it comes to including all the

relevant costs regardless of the perspective (98). Interventions for children often have a preventive perspective or an element of chronic disease which makes the relevant time horizon for cost to consider potentially lifelong. This is especially challenging when concerning children. As the children grow up, there are likely to be changes in treatment settings even for the same conditions. For example, a child suffering from mental disorder might in childhood primarily incur costs in the educational sector and in the CAMHS, while costs in the adult Mental Health Services and productivity costs might be the most relevant in adulthood. Hence, the data requirements get correspondingly larger than when dealing with adults that remain in a more homogenous treatment setting compared to children (98).

Another special concern with children is their parents. Depending on the perspective of the analysis some costs concerning the parents might need to be taken into consideration. A child's health will impact on parents' lives. This can be in the form of informal care, transport, participating in interventions, or even being the subject of the intervention e.g. in interventions meant to teach the parents how to help their child deal with mental health problems. This time could potentially have been spent working and can thus, be included as lost productivity costs (72).

As the results of a health economic evaluations is not impacted by costs that are the same between the compared groups, a pragmatic way of assessing whether it is necessary to include specific costs is to consider whether the interventions are likely to affect the cost in question within the relevant time horizon. For example, in the case of an intervention for milder mental health problems in a country with low juvenile crime rates, it might not be necessary to consider costs in the criminal justice system despite having a societal perspective.

Data limitations often makes it difficult to consider all relevant costs in health economic evaluations involving children. Especially the estimation of the future costs in adulthood have in general not been considered in studies concerning cost-effectiveness and children's mental health (63). The obvious and inevitable problems with data availability this are, however, still an important limitation.

Another concern is how to estimate the costs identified as relevant for the analysis. This is normally done in two steps. Firstly, the quantities of resources need to be measured and secondly, they need to be assigned a price or unit cost (72).

In Denmark, there is comprehensive national register data on the use of services in the health care sector (99). For health services provided in the municipalities, this is not the case. This makes it challenging to measure the quantities of the services provided in this domain. For children's mental health, this is especially relevant because some of the treatment is expected to take place in the municipalities as described in section 1.2. Different standardized questionnaires have been developed in an attempt to capture and quantify services provided for children by self- or parent-reporting (98). It is, however, difficult for these standardized questionnaires to capture the relevant aspects across the different types of diseases and different countries. Therefore, if this approach is utilized, it will often be necessary to design the questionnaires to the specific setting.

Theoretically, the correct price for a unit is the opportunity cost, which is defined as the value of the foregone benefits that the unit could have led to in an alternative use (72). In a perfectly competitive market without distortions, market prices will reflect opportunity costs. However, such a market is seldom present in the

context of health care services. Therefore, in practice, the unit costs are typically estimated more pragmatically using available non-market prices or tariffs. For health care services, this often means the use of tariffs and average prices, for instance, the use of diagnosis-related group (DRG) based tariffs that represent a unit cost for hospital treatment of a specific diagnosis. In practice, the different patients will have different treatment pathways, and the actual cost for a patient will inevitably differ. The same is true for tariffs for other health care costs, for example, visits at general practitioners. An alternative to the use of tariffs is a micro costing approach, where the actual use of time, equipment etc. are measured. This is, however, a very time consuming and labor demanding approach. Thus, the use of tariffs is a limited approach but often the most feasible.

2. The aims of this thesis

Every year, a large proportion of children get diagnosed with mental disorders which, to some extent potentially, could have been prevented (24), and an unknown number of children suffer from mental health problems with no access to appropriate care. The economic case for prevention and early treatment seems clear with both considerable cost of childhood mental health problems and high cost of the linked adulthood mental disorders. Despite the increasing evidence base of the effectiveness of CBT-programs to treat mental health problems and potentially prevent mental disorders, there is a lack of implementation and dissemination of evidence-based early treatment for children with common mental health problems. Besides not meeting the needs of the children, this can also cause inequity in access to care, as resourceful families are more likely to be able to get self-financed support and hence more timely care.

This all calls for initiatives to introduce a systematic approach to identify children with mental health problems early and provide them with appropriate evidence-based treatment. When introducing interventions to a population that currently does not receive care, the new interventions require investments. It is, therefore, crucial that the cost-effectiveness of a new intervention is evaluated to make sure that decisions of introducing the intervention can be taken on the best possible informed basis.

The overall aim of this thesis was to contribute to the evidence base of early treatment in children with common mental health problems. This included the contribution of empirical evidence on new innovative ways of identifying and stratifying the target population for providing early treatment of mental health problems (paper I & II) and on the cost-effectiveness of the early treatment MMM (paper IV). It also included methodological work in the form of a validation of the use of the Danish version of CHU9D (paper III) and empirical evidence on the impact of the choices taken when estimating QALYs in child populations (paper IV).

2.1 The specific aims

The aim of **paper I** was to evaluate whether parents' responses to SDQ and an algorithm could be used in the general population to identify children with mental health problems in need of an early treatment.

The aims of **paper II** were to describe the help-seeking population of children with mental health problems and a novel stage-based stepped-care model, as well as evaluate whether the model could facilitate a systematic identification and stratification process with equal access to care in a naturalistic setting.

The aims of **paper III** were to examine the construct validity and longitudinal validity of CHU9D in a mental health setting, and to examine differences in this validation between preference weights derived from an adult and an adolescent population.

The aims of **paper IV** were to evaluate the cost-effectiveness of the Mind My Mind intervention versus management as usual using a cost-utility analysis framework, and to analyze the impact of choice of informant with respect to both quality-of-life reporting and preference weights on the results.

3. Summary of papers

3.1 *Evaluation of a screening algorithm using the Strengths and Difficulties Questionnaire to identify children with mental health problems: A five-year register-based follow-up on school performance and healthcare use*

3.1.1 Objectives

The objective was to evaluate the potential for using parent-reported SDQ and a screening algorithm to identify children with mental health problems in need of an early intervention. To do this, a longitudinal study was conducted using data from a historical birth cohort to follow children for five years and analyzed their school performances and health care utilization based on their screening status at baseline.

3.1.2 Methods

The study used data from the Copenhagen Child Cohort (CCC2000) that included all 6,090 children born in the year 2000 in the former Copenhagen county. The parents of the children were invited to complete the SDQ when the children were aged 11-12 years.

The SDQ is well-validated in both clinical and population samples for identifying and assessing mental health problems (100–104). The first part of the SDQ covers several subscales in the 25 items: pro-social behavior, emotional problems, peer problems, behavioural problems and hyperactivity. The four latter subscales can be combined into a total difficulties score (105). The functional impairment score is calculated from five items that address whether the difficulties upsets or distresses the child and how much the difficulties interfere with home life, friendships, classroom learning and leisure activities (100,106).

The study population was composed of 2,015 children with parent-reported SDQ and no previous contact to mental health care. The children were defined as screening positive for mental health problems or screening negative based on a cut-off delineated by an SDQ scoring algorithm. The cut-off algorithm was based on population surveys (106–108), so that screening positive implied that the child was in the top 10 percentile with the poorest mental health.

Data on school performance and health care utilization were obtained from national registers. Graduating from secondary school and grade point average (GPA) were analyzed as separate outcomes. School performance was used as a long-term outcome and as an indicator of the potential life-long outcome of mental health problems. Mental health problems have previously been linked with negative school outcomes (109–111), and school performance is a predictor for several adverse events later in adult life (112–114). Health care costs for the five-year period were analyzed both at an aggregated level and per type of health care. A range of potential confounders were included in the analyses in the attempt to isolate the effects of mental health problems. These included an IQ-proxy score from a test conducted face-to-face with the child, and several register-based variables on perinatal factors, parental health and socio-economic characteristics of the family.

Logit regression models and Ordinary Least Squares regression models were used to analyze school outcomes, while Generalized Linear Models and two-part models were employed in the analysis of health care

utilization to account for the heavily right-skewed data and the many cases of zero costs. Several tests were carried out for each analysis to determine which model type, link function and distribution had the best fit.

3.1.3 Results

A total of 194 (10%) children with no previous contact to mental health services screened positive for mental health problems at the age of 11-12 years. When analyzed as one group, these children did not have a higher odds ratio of not graduating than the children screening negative, but they did get a significantly lower GPA. They also had significantly higher health care costs in the five-year follow-up period. In total, the health care costs of the five-year period were estimated to be €6,044 higher on average per child compared to the children that screened negative. The costs were mainly due to visits at CAMHS. Forty-four of the 194 children that screened positive had a visit to a psychologist, psychiatrist and/or visited CAMHS during the five-year follow-up period. The remaining 150 children did not have any visits relating to mental health care.

When analyzed separately, we found that the 44 children that screened positive and did have mental health care visits were the source of increased health care utilization. These 44 children had a significantly higher risk of not completing secondary school, and those who did complete obtained a significant lower GPA. On average, the 44 children had their first mental health care consultation over 2.5 years after their parents completed the SDQ. The 150 children that screened positive but did not have a mental health care visit during the follow-up period, did not differ in their level of health care utilization from the children that screened negative. Also, they did not have a higher risk of not completing secondary school but when they graduated, it was with a significantly lower GPA.

3.1.4 Summary

The results showed that it was possible to identify a group of children with parent-reported mental health problems that had significantly poorer school outcomes compared to their peers. The findings showed that the majority of children that screened positive did not have any mental health care consultations during the five-year follow-up period. These children could potentially benefit from an early intervention. The children that screened positive and had a mental health care consultation, on average, had the first consultation after 2.5 years. This finding showed a potential “window-of-opportunity” to provide these children with an early treatment before the mental health problems progress into severe disorders.

3.2 *Systematic identification and stratification of help-seeking school-aged youth with mental health problems - A novel approach to stage-based stepped-care*

3.2.1 Objectives

The objectives were to describe the psychopathology of help-seeking school-aged children with mental health problems, and to evaluate whether a new visitation model was able to facilitate a systematic identification and stratification process with equal access to care.

3.2.2 Methods

The visitation model was based on a stage-based stepped-care approach and developed in the context of the MMM trial to identify children with mental health problems qualifying for early treatment. The model was implemented in the setting of Educational and Psychological Counselling Services (PPR) in the four municipalities that participated in the trial. The aim of the model was to stratify the children with mental health problems into three groups: A group of children in need of a less intense intervention than the MMM (stage group 1), children in need of an early treatment such as MMM (stage group 2) and a group with indication of severe mental disorder who were in need of an assessment at CAMHS or other specialized treatment (stage group 3).

There exists a variety of well validated questionnaire instruments for use in measuring mental health in children. Some focus on specific domains of mental health problems such as The Mood and Feelings Questionnaire (MFQ), which focuses on symptoms of depression (115), Spence Children's Anxiety Scale (SCAS), which focuses on symptoms of anxiety (116) and Eyberg Child Behavior Inventory (ECBI), which focuses on disruptive child behavior (117). All three instruments, together with additional instruments, were included in the MMM trial. As children typically present themselves with a diverse set of symptoms, (66) MFQ, SCAS and ECBI are, on an individual basis, not ideal for describing a child's psychopathology when the child is first assessed due to the likely possibility of the child having symptoms in more than one domain. Other instruments have a more generic approach. Of these, the Child Behavior Checklist (CBCL) (118) and the Strengths and Difficulties Questionnaire (SDQ) (100) are most often used. The two are highly correlated (119), but an important difference is that while CBCL only focuses on symptoms, SDQ also focuses on the impact of the mental health problems on the child's life. In psychometric measurement, where the difference between normal and problematic levels of symptoms is difficult to define clearly and typically relies on population norms and other more or less arbitrary thresholds, the information on how the child's everyday life is affected gives a better foundation to both assess the need for an intervention and the impact of the intervention.

Information about the MMM trial and the visitation procedure was advertised in various relevant places. Help-seeking parents had to contact the local PPR administration for the child to get enrolled into the visitation. The visitation model consisted of two phases. In phase 1, all parents were invited to report SDQ in an online format. The SDQ screening algorithm presented in paper 1 was used to define stage group 1 by defining the children that scored below the cut-off as not having severe enough mental health problems to qualify for an early treatment program. For the children with parent-reported SDQ scores above the algorithm cut-off, both parents and children were invited to report SCAS and MFQ before they were invited to phase 2 of the visitation process. Phase 2 consisted of a face-to-face meeting with a PPR psychologist. Based on the results from the online questionnaires and a brief semi-structured psychopathological interview with the child and the parents, the psychologist assessed if the child had indications of severe mental disorders and

needed assessment at CAMHS. This assessment was used to define stage group 3. The children who were identified with need for early treatment constituted the target group for the MMM trial defined as stage group 2. The time lag between phase 1 and 2 was examined to evaluate potential problems with delay in access to care.

Parent-reported SDQ was used as the primary instrument to describe the psychopathology of the children in the three groups. The parents of the children that were enrolled in the MMM trial answered the Developmental And Well-Being Assessment (DAWBA) at the baseline of the trial. DAWBA is a questionnaire and rating technique that can generate psychiatric diagnoses in children according to the Diagnostic and Statistical Manual of Mental Disorders, 4th and 5th Editions (DSM-IV/-5) (120,121). The DAWBA diagnoses of stage group 2 were analyzed to provide a more in-depth understanding of the severity of the mental health problems.

To evaluate whether the visitation model provided equal access, we compared the age, sex and socioeconomic characteristics of the children that took part in the visitation model with all children aged 6-16 in the four participating municipalities. All data on both the participating children and the background child population were obtained from national registers on an individual level. Differences between the population that participated in the visitation program and the background population could be due to differences in the prevalence of mental health problems based on age, gender and socioeconomic characteristics. To qualify issues of inequity in access ,we therefore interpreted all findings in light of the differences in prevalence of mental health problems identified in paper 1, and differences in the prevalence of diagnosed mental disorders identified in nationwide Danish cohorts (122–124).

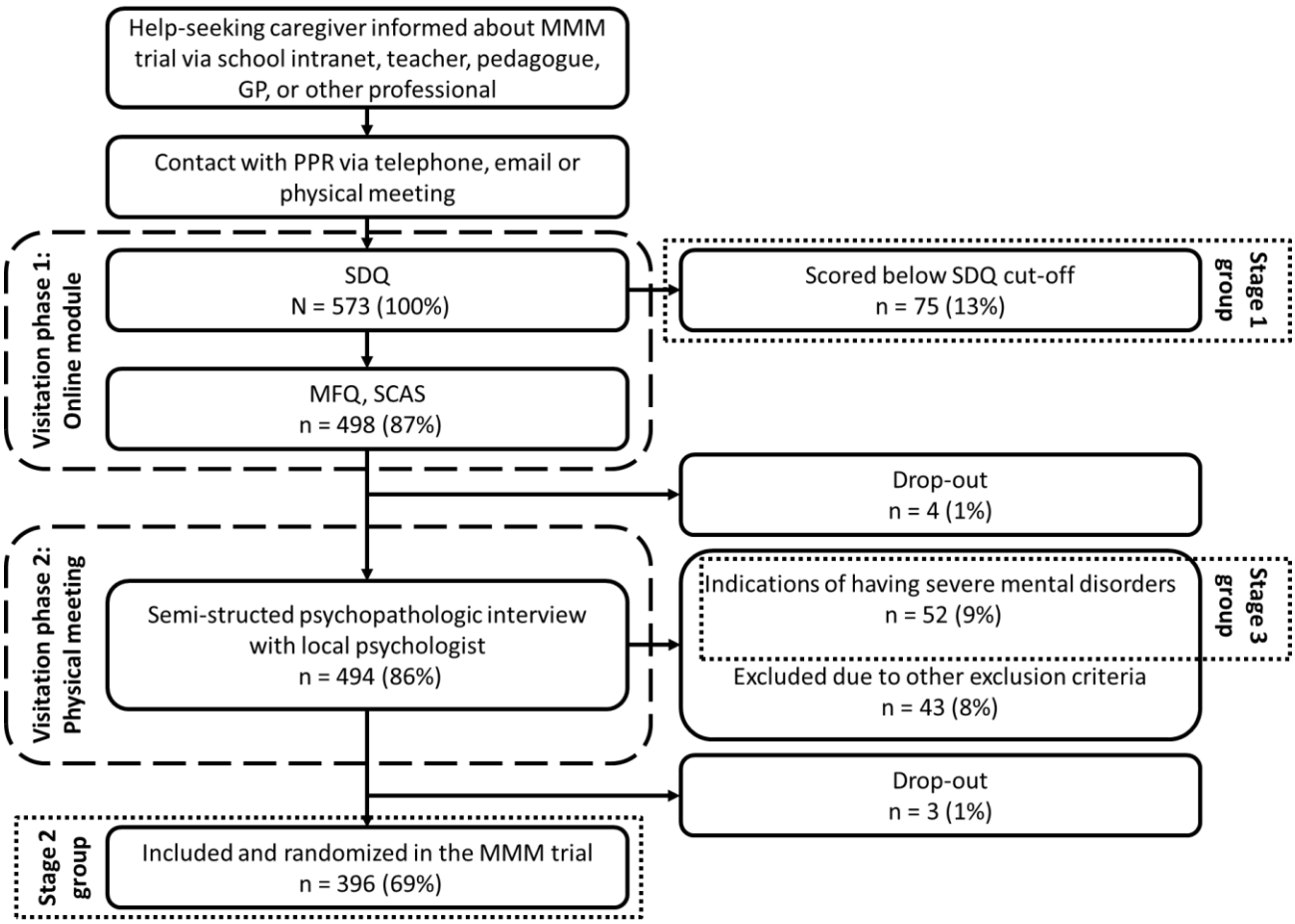
3.2.3 Results

573 children with mental health problems and help-seeking parents were enrolled into the visitation model. As many as 89% of the 573 parents reported that the child had suffered from the mental health problems for more than five months.

The flow of the visitation process is shown in Figure 1. The parent-reported SDQs of 75 children (13%) scored below the algorithm-based cut-off and were therefore assessed as not being in need of an early intervention of the intensity that was offered in the MMM treatment course. In phase 2, the psychologist assessed that 52 children (9% of the full sample) had indications of severe mental disorders. These children were therefore advised to seek assessment at CAMHS. The remaining 396 children were assessed to be in need of early treatment and were therefore enrolled in the MMM trial. Seven (1%) dropped out during the visitation process for unknown reasons, and 43 children (8%) were excluded due to a range of exclusion criteria; “parents not interested in participating in a trial” being the most common.

The mean time period from phase 1 to the face-to-face interview was 18.7 days. The median time was 15 days and within 36 days of reporting the online psychopathological questionnaires, 95% of the children had the face-to-face interview.

Figure 1: Flowchart of visitation process



Analyses of the SDQ data revealed that all three stage groups had statistically significant ($p<0.05$) higher SDQ total difficulties score and SDQ impact score than the Danish norms (107). For both scores, there were also statistically significant differences between each of the three groups, with increasing scores from stage 1 group to stage 2 group, and from stage 2 group to stage 3 group. The results can be seen in Table 1.

Table 1: Psychopathology of the children at visitation

	Stage 1 group n = 75	Stage 2 group n = 396	Stage 3 group n = 52	Danish norms[1] N = 3,146
Strength and Difficulties Questionnaire, mean (standard deviation)				
Emotional problems	3.5 (2.0)[*]	7.0 (2.4)	6.4 (2.4)	2.3 (2.3)[$]
Behavioural problems	1.4 (1.3)[*]	2.8 (2.0)	3.3 (1.8)	0.9 (1.3)[$]
Hyperactivity	3.1 (2.1)[*]	5.0 (2.8)[+]	6.3 (2.3)	2.5 (2.5)[$]
Peer problems	1.6 (1.6)[*]	2.8 (2.1)[+]	3.7 (2.6)	1.3 (1.7)[£]
Pro-social behaviour	8.4 (1.9)[*]	7.6 (2.1)	7.0 (2.3)	8.7 (1.5)[£]
Total difficulties score	9.6 (4.1)[*]	17.5 (5.1)[+]	19.7 (5.1)	7.1 (5.8) [$]
Impact score	1.1 (1.6)[*]	4.2 (1.9)[+]	5.3 (2.2)	0.6 (1.6) [$]

Note: [1]Danish norms of children aged 6-17 from Arnfred et al. (107). [*]Group 1 is statistically significantly different from group 2 and 3 on a p-level <0.05 tested with both t-test and Mann-Whitney U test. [+]Group 2 is statistically significantly different from group 3 on a p-level <0.05 tested with both t-test and Mann-Whitney U test. [$]Danish norms are statistically significantly different from group 1, 2, and 3 on a p-level <0.05 tested with t-test. [£]Danish norms are statistically significantly different from group 2 and 3 on a p-level <0.05 tested with t-test.

The DAWBA diagnoses from stage group 2 revealed that 80% of the 396 children suffered from a diagnosable mental disorder, and that 26% suffered from more than one main category of mental disorder, i.e. anxiety and depression.

Three potential problems with the visitation model were identified when comparing the participating population with the background population. Firstly, the participating population had a clinically and statistically significant ($p<0.05$) lower age than the background population. Only 10% of the participating population were aged 14-16 years compared to 29% in the background population. Secondly, there were less than five children with an immigrant background in the participating population compared to 10% in the background population. Thirdly, the proportion of children with mother's highest education being lower secondary school was significantly ($p<0.05$) lower in the participating population (8% versus 15%). When interpreting these findings in light of the external population studies, the observed differences in age, ethnicity and education should not be present if participation in the visitation model was representative of the prevalence of mental health problems.

3.2.4 Summary

We found that there is a large group of children with long-standing mental health problems who do not receive help despite their parents being help-seeking. The majority of the children that self-referred to the visitation model met the diagnostic criteria of at least one mental disorder diagnosis or were assessed as

having indications of severe mental disorders. The results showed that the visitation model provided a stage-based stepped care framework that succeeded in stratifying the children into three groups with increasing severity of mental health problems eligible for three different types of interventions. The process of the visitation was done in less than a month for most of the children. Problems with inequity in access were identified in the study, suggesting that older children, children of mothers with a low level of education and children of parents born outside Denmark are not as likely to access interventions provided based on the visitation model. It is essential that these problems are addressed if the visitation model is to be implemented on a wider scale.

3.3 *The longitudinal validity of the CHU9D – Evidence from a mental health trial*

3.3.1 Objectives

The objectives of this study were to examine the construct and longitudinal validity of the CHU9D in a mental health setting, and to examine how the choice of preference weights affects the validity.

3.3.2 Methods

We used blinded data from the MMM trial for the validation. 396 children aged 6-16 with mental health problems were enrolled in the trial following the visitation procedure described in paper 2. The validation used parent-reported data on CHU9D, SDQ and KIDSCREEN-27 collected at baseline and at end-of-treatment (week 18). The HSUV based on CHU9D were calculated in two ways using two different sets of preference weights. One based on preferences of adults and one based on preferences of adolescents. The total difficulties score and the impact score of SDQ were used separately in the validation. KIDSCREEN-27 measures HRQoL and wellbeing in five domains. All domains were used in the validation, but the domain psychological well-being was the only one used in the examination of discriminant validity and responsiveness. The different instruments do not measure the same thing, but they do have conceptual overlaps, as shown in Table 2. This was taken into account when formulating the hypothesis and when interpreting the results.

Table 2: Simplified conceptual overlap between measures by quality of life dimensions

	Quality of life dimensions						
	Mental health	Physical health	Sleep	Parental relation	Social relations	School	Activities and daily routines
CHU9D	**X**	**X**	**X**			**X**	**X**
SDQ Total difficulties score	**X**				**X**		
SDQ Impact Score	**X**			**X**	**X**	**X**	**X**
KIDSCREEN Physical Well-being		**X**					**X**
KIDSCREEN Psychological Well-being	**X**						**X**
KIDSCREEN Autonomy & Parent Relation				**X**			
KIDSCREEN Social Support & Peers					**X**		
KIDSCREEN School Environment						**X**	

Construct validity was assessed using the baseline data and by examining discriminant and convergent validity. In order to test for discriminant validity of CHU9D, the children were divided into three groups for each of the psychometric instruments: 1) the 25% with the lowest score, 2) the 25% with the highest score, and 3) the 50% with the scores in between. To account for the non-normality of the utility distributions, Kruskall-Wallis test was used to test for statistical differences between the groups. Spearman rank correlations were used to examine the convergent validity.

Longitudinal validity was assessed using baseline data and data from end-of-treatment (week 18) by examining the floor/ceiling effects, the magnitude of change over time, and the ability of CHU9D to discriminate between children that had an improvement and those who did not (defined by the other psychometric instruments). Standard Response Mean statistics were used in the assessment of the magnitude of change over time. The group differences in mean changes between the children who improved and the children who did not improve was tested using Mann-Whitney test due to the non-normality of the utility distributions.

A priori hypotheses were defined for all assessments.

3.3.3 Results

The discriminant validity showed that CHU9D performed well in distinguishing between the three groups with low, intermediate and high levels of mental health problems on both SDQ scores and KIDSCREEN psychological well-being ($p<0.001$). In the examination of the convergent validity, a strong correlation was found between KIDSCREEN psychological well-being, a moderate correlation was found for SDQ impact score, and a weak correlation was found for SDQ total difficulties. Moderate correlations were hypothesized for all three correlations a priori. The assessment of the longitudinal validity showed that CHU9D was sensitive to change over time. A statistically significant higher mean change in HSUV was found for the children that improved their mental health independent of which of the three other instruments that were used to measure mental health. These results can be seen in Table 3.

Table 3: Mean change in CHU9D utility scores for those who improved, or did not improve, post intervention and the mean difference between these two groups

Change in:	Change in CHU9D adult weights, mean (n)	Difference, mean (95% CI)	Change in CHU9D adolescent weights, mean (n)	Difference, mean (95% CI)
SDQ Total difficulties score				
≥0 n=105	0.017 (105)	0.055***	0.030 (105)	0.105***
<0 (improved score) n=233	0.072 (233)	(0.028;0.083)	0.135 (233)	(0.055;0.155)
SDQ Impact Score				
≥0 n=105	0.015 (105)	0.059***	0.026 (105)	0.111***
<0 (improved score) n=233	0.074 (233)	(0.032;0.086)	0.137 (233)	(0.061;0.160)
KIDSCREEN Physical Well-being				
<0 n=86	-0.018 (86)	0.098***	-0.023 (86)	0.168***
≥0 (improved score) n=253	0.080 (253)	(0.070;0.126)	0.144 (253)	(0.117;0.219)
***p-value >0.001				

A comparison of CHU9D's discriminant validity when HSUV was derived from adult and adolescent populations, respectively, showed that the magnitude of the mean differences between the comparison groups at baseline were higher when using the adolescent weights. The same was found when assessing the mean change in HSUV for the children who improved their mental health over time, as seen in Table 3.

3.3.4 Summary

The results of this study showed that CHU9D is an appropriate instrument to include in mental health trials to enable the estimation of QALYs. Furthermore, the study provided the first assessment of the longitudinal validity of CHU9D. Combined with the comprehensive evidence of validity from cross-sectional studies, the positive assessment of the longitudinal validity in this study suggests that CHU9D is in general an appropriate preference based HRQoL instrument for use in child populations.

In the examination of the differences in HSUV generated on the basis of the two sets of preference weights derived from an adult population and from an adolescent population, respectively, we found that the magnitude of the mean difference between groups were larger when using the adolescent weights. This was true across groups with different severity of mental health problems at baseline, and also true across groups that improved and did not improve over time based on the SDQ and KIDSCREEN. The findings suggest that the choice of preference weights will have a significant impact on the results of CUAs conducted in mental health settings.

3.4 *Cost-effectiveness of a transdiagnostic psychotherapy program for youth with common mental health problems*

3.4.1 Objectives

The objective was to evaluate the cost-effectiveness of the transdiagnostic psychotherapy program MMM designed to help children with common mental health problems using a cost-utility analysis framework. Furthermore, we investigated how choices underlying the estimation of QALYs impacted the results. The choice between self-reported and parent-reported HRQoL was investigated, as well as the choice between preference weights derived from an adult population and preference weights derived from an adolescent population.

3.4.2 Methods

The study uses data from the MMM trial. 396 children aged 6-16 participated in the pragmatic RCT that was carried out in four Danish municipalities. The participating children were identified through the visitation procedure described in paper 2. They all had mental health problems concerning anxiety, depression, and/or behavioral problems. 197 children were randomized to MMM and 199 children to management as usual (MAU).

The MMM intervention is a manualized, modular CBT-program that comprise 9-13 individual sessions plus a booster-session. MMM was carried out by trained local psychologists under supervision of trained psychologists from CAMHS. The MMM group was not eligible for other psychotherapeutic interventions in the municipality doing the trial period. The MAU group was eligible for all municipality interventions. In order to enhance the MAU intervention and hence the motivation to stay in the trial , this group was offered two care-coordination visits in their municipality.

The cost-utility analysis was carried out using an extended health care sector perspective. In the base case, a time horizon of 26 weeks was applied in alignment with the length of the trial. Four one-year extrapolation scenarios were carried out making the time horizon 1.5 years for these analyses. QALYs were estimated using the CHU9D reported by the parents (as proxies) in the base case, but an analysis using self-reported CHU9D was also carried out. CHU9D was collected at baseline, at end-of-treatment (week 18) and at follow-up (week 26). In the base case, QALYs were estimated using preference weights based on the preferences of an adolescent population, but an analysis using preferences of an adult population was also carried out. Costs of MMM and the care-coordination visits for MAU were estimated using registered trial data. Cost of other MAU interventions as well as the cost of private psychologists and health care visits in both groups were estimated using parent-reported data at week 18 and week 26 and public available costs of the services.

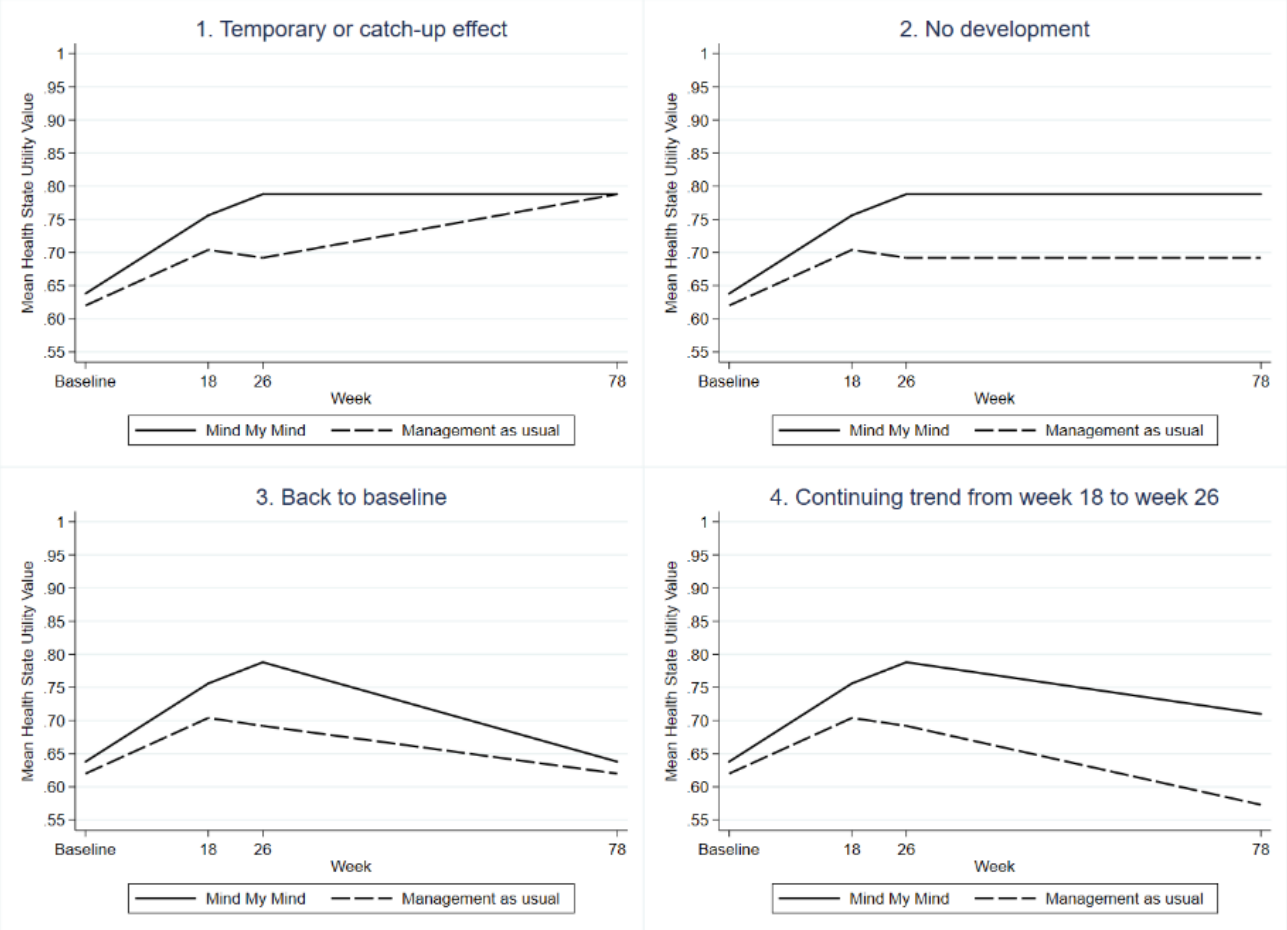
A total of 74 children (19%) had a missing HSUV or cost estimate at any point. The data was treated as missing at random within each group based on analysis of the observed variables. Missing data was handled using multiple-imputation with chained-equations and predictive mean matching performed on utility score level and cost-component level. In all analyses of incremental QALYs, results were adjusted for baseline utilities. Sample uncertainty was examined using non-parametric bootstrapping with samples being drawn from each of the multiple imputed datasets. In total, 10,000 iterations were carried out.

3.4.3 Results

The MMM group had incremental costs of €2,981 (95% CI: 2,731-3,251) during the 26 weeks of the trial. The main driver of the costs was the MMM intervention. Care-coordination visits were utilized in the MAU group, but the extent of psychological treatment provided by the municipalities in the MAU group was limited. The MAU group had slightly higher costs relating to other health care visits, due to more visits at private psychologists.

Both groups experienced an increase in HRQoL from baseline until end-of-treatment (week 18), but while the MMM HRQoL was consistently increasing during follow-up (up to week 26), the opposite happened for MAU. After adjusting for baseline utility, the incremental QALY was estimated at 0.017 (95% CI: 0.006-0.029) in the base case. This resulted in an ICER of €170,465 per QALY gained. In Figure 2, the mean HSUVs for the two groups can be seen at the observed time points (baseline, week 18 and 26) and at the four one-year extrapolation time points (week 78).

Figure 2: Mean health state utility values for the four scenario extrapolations



Using the data presented in Figure 2 resulted in a higher number of incremental QALYs and an ICER ranging from €23,653 to €50,480 per QALY gained.

When using the self-reported CHU9D to estimate QALYs, the incremental QALYs fell to 0.014 (95% CI: 0.002-0.027) in the base case. This corresponds to an ICER of €210,757 per QALY gained which is 24% higher than

when using the parent-reported CHU9D. Choosing the preference weights derived from an adult population in the estimation of QALYs lead to an incremental QALYs of 0.010 (95% CI: 0.004-0.016), and an ICER of €291,916 per QALY gained. This ICER is 71% higher than the base case ICER. The differences in the QALY estimates can be seen in Table 4 where they are compared with the base case.

Table 4: Comparison of the QALY estimates from the base case and the sensitivity analyses

	Parent-reported CHU9D Tariffs based on preferences of an adolescent population (base case)		Self-reported CHU9D Tariffs based on preferences of an adolescent population		Parent-reported CHU9D Tariffs based on preferences of an adult population	
	MMM (n=197)	MAU (n=199)	MMM (n=197)	MAU (n=199)	MMM (n=197)	MAU (n=199)
Health-related quality of life						
Baseline	0.638 ±0.015	0.620 ±0.015	0.642 ±0.016	0.631 ±0.016	0.806 ±0.009	0.802 ±0.008
End-of-treatment (week 18)	0.756 ±0.015	0.704 ±0.017	0.728 ±0.016	0.684 ±0.018	0.873 ±0.008	0.844 ±0.009
Follow-up (week 26)	0.788 ±0.014	0.692 ±0.018	0.769 ±0.015	0.708 ±0.186	0.886 ±0.008	0.839 ±0.009
QALY	0.360 ±0.006	0.337 ±0.007	0.352 ±0.007	0.335 ±0.007	0.426 ±0.003	0.414 ±0.004
Incremental QALYs						
Unadjusted	0.024 ±0.009		0.017 ±0.009		0.012 ±0.005	
Adjusted for baseline	0.017 (95% CI: 0.006-0.029)		0.014 (95% CI: 0.002-0.027)		0.010 (95% CI: 0.004-0.016)	
ICER	€170,465 /QALY gained		€210,757 /QALY gained		€291,916 /QALY gained	

Note: Values presented are mean ±standard error of the intention-to-treat population from 20 imputed datasets. Confidence intervals (CI) are derived from 10,000 bootstrap replications from 20 imputed datasets using the standard normal method.

3.4.4 Summary

The MMM intervention has positive effects on the HRQoL of children with mental health problems compared to MAU, but it comes at a cost. Whether MMM can be considered cost-effective is a normative question, but considering the general range of the ICERs in the analyses of the extrapolated scenarios, it seems likely that MMM can be cost-effective based on the direct treatment effects, even when not taking into account possible preventive effects on mental disorders later in life. It is, however, dependent on the actual duration of effects as the ICER in the restricted base case can be considered very high.

The results of the analyses investigating the effects of the choices between respondent and source of preferences for the estimation of QALYs showed that both choices have significant impact on the ICER. Given that HTA institutes rely on the comparability of results between CUAs, it is important that guidelines and standards are developed on how to handle these choices.

4. Discussion

This thesis addresses a number of challenges in the delivery of early treatment for children with mental health problems. The findings show that it is possible to identify the children in need of an early intervention, and that a new innovative visitation model has the potential to facilitate a systematic assessment and staging of the children that can make sure they get offered appropriate, timely intervention. The thesis also contributes important empirical evidence on the estimation of QALYs in child populations. The results support the use of CHU9D to estimate QALYs in child with mental health problems and highlights the importance of making conscious choices in respect to both the choice of HRQoL informant and the choice of whose preference to use in the estimation of QALYs. Finally, the cost-effectiveness of the new transdiagnostic early treatment program MMM was evaluated with promising results.

The results presented in this thesis have important policy and methodological implications. These implications will, together with the strengths and weaknesses of the studies, be discussed in this section.

4.1 The unmet need for early treatment of children with mental health problems

In paper 1, we identified 10% of children in the birth cohort, with no previous specialized health care visits relating to mental health problems, as having mental health problems and being in need of an early treatment. This unmet need constitutes a massive issue for the children, their families and society as a whole, due to the potential serious impact of mental health problems. At the same time, the findings represent a unique opportunity to provide this population with effective interventions and thereby, potentially avoiding some of the negative impact for the children, the families and society. Our study does, however, have limitations that need to be considered.

To assess whether the children had mental health problems, we used school performance as an indicator of the long-term outcome. School performance is likely to be the result of several factors, whereby mental health problems is just one of them. Despite the rich data and our efforts to isolate the effects of mental health problems on school performance, there is still a risk that our findings were confounded by unobserved variables. Hence, we are not able to conclude that our findings are causal. Another limitation is that we did not have any data on interventions that may have been initiated in the municipalities. The findings in paper 4, of very limited targeted psychological interventions being delivered in the MAU group, does however suggest that this limitation does not affect our conclusion drawn in paper 1 on the potential for early treatment. It is likely that more universal interventions, e.g. promotion of mental health, have been conducted in the municipalities doing the follow-up, but these should be seen as a complement to early treatments rather than interchangeable.

Using the same screening algorithm as in paper 1, we found in paper 2 that a help-seeking population above the cut-off suffered from severe impairment due to their mental health problems, and 80% fulfilled the DSM-IV/5 diagnostic criteria of at least one mental disorder. These are children that in the current organization of the health care system, despite their diagnosable mental disorder, do not qualify for treatment at CAMHS. They are currently not being helped systematically, and the majority are not offered any evidence-based treatment in the health care system despite their parents looking for help. This finding supports not only the causality of the associations found in paper 1, but also further underlines that there is a window of

opportunity for helping these children, as we find that they are both massively in need of early treatment and are actively looking for it.

The findings in both paper 1 and paper 2 may be limited in terms of their generalizability due to imperfect representativity. We find that the children who participate in the studies are more likely to come from more resourceful families where the parents are more educated and have a higher income. We do, however, suggest that our findings are likely to be as relevant for the children from less resourceful parents, as these families could have more comprehensive issues in coping with the mental health problems, whereby these children may benefit even more from effective early treatment. The finding of imperfect representativity is not only a limitation for the results, but also an interesting finding itself. The lack of representativity suggests that less resourceful parents and parents born outside Denmark may be more difficult to reach. Hence, a more proactive strategy for reaching these parents and their children may be warranted, to ensure improvement in equity in health.

4.2 Usage of the screening algorithm

Measuring mental health problems is complicated due to the lack of general biomarkers making psychometric questionnaires, interviews and observations the only way to assess the presence of mental health problems. Furthermore, mental health problems cover a broad range of diverse problems. This makes it difficult to design a perfect screening tool to identify mental health problems.

The evaluation of the screening algorithm based on SDQ presented in paper 1 relies on group averages of later school performance to assess the success of the screening algorithm. The longitudinal design of the study combined with the unbiased assessment of the school performance outcomes strengthens the findings of the study. The study does, however, also has some limitations. Screening tests are typically evaluated based on their sensitivity and specificity, meaning how big a fraction of the true positives are being detected, and how big a fraction of the true negatives are actually screened negative (125). In the analyses presented in paper 1, we decided not to examine these fractions as we do not have a professional interview-based assessment of the children's need for care that could be used as an unbiased gold standard comparison. We could have described the sensitivity and specificity of the SDQ-screening with regard to the DAWBA-based probability of a mental disorder, that was assessed at the same age, but this comparison would have been biased by use of parents as the same informant (126). As the screening test relies on parent-reported SDQ, it is likely that there will be some reporting bias. How parents see and interpret their children's behavior is likely to vary to some degree. Furthermore, some internalizing mental health problems might go undetected by the parents (93). Hence, the screening procedure using the parent-reported SDQ-scores and the simple algorithm (presented in paper 1) might not be suitable for mass screening, but more useful as targeted screening among help-seekers and used as the starting point of a more thorough assessment.

With the screening algorithm, we aimed to identify children that would benefit from an early intervention. To screen positive, it is required that the parent report that the child has "quite a lot" or "a great deal" of impairment due to mental health problems. It could, however, be argued that children with high levels of mental health symptoms, who do not yet experience impairment, may likely benefit from an early treatment that could prevent them from developing impairment. The same could be the case for children with milder symptoms that already are experiencing impairment. Four and six percent of the children who screened

negative had a SDQ total difficulties score above the threshold and a SDQ impact score above the threshold, respectively.

In paper 1, we primarily focused on evaluating the screening based on children aged 11-12 years. Supplementary analyses of partly the same population aged 5-7 years showed that the association found between screening positive and later school performance was not present when analyzing the children aged 5-7 years. This suggests that the screening algorithm might not be equally appropriate across age, and that it should not be used in preschool populations. A possible explanation might be that part of the mental health problems reported by the parents at age 5-7 years, i.e. children in pre-school age, are related to less stable or more treatable developmental problems (3). Hence, an explanation for the lack of association between screening positive at age 5-7 years and school performance at end of secondary school might be that these children, to some degree, matured or received appropriate help at an early stage and therefore, were less affected.

4.3 Implementation of the visitation model

The evaluation of the visitation model demonstrated a real-world usage of the screening algorithm. The visitation model succeeded in providing a fast stratification of the children based on the principles of stage-based stepped-care. A strength of the study is that the evaluation of the visitation model was done in a pragmatic setting in four diverse municipalities and the procedures were carried out by local employees. The positive findings of the study in this diverse setting seem promising for a more widescale national implementation. We did identify problems with recruiting the older school children, children of mothers with low level of education and children with foreign born parents. Careful consideration about how to also make the visitation accessible for these groups needs to be done before further implementation to avoid developing problems with equity, and to achieve the full potential of the visitation model.

The visitation model succeeded in identifying children in need of an early treatment, and the treatment was delivered fast in the form of participation in the MMM trial. To gain the full advantage of the visitation model, all three stage groups must receive the appropriate intervention fast and without unnecessary delay. It was outside the scope of paper 2 to look into the end product of the visitation model, which involves a timely and equitable delivery of effective interventions at all three stages. If the children do not receive an intervention after participating in the visitation, the procedure may have an adverse effect. For instance, if the children who have indications of severe mental disorders are rejected by CAMHS, it is likely to cause frustration among both children and parents, given that they are told by health professionals that their children are too heavily impacted by their mental health problems to receive an early treatment, but are told by other health professionals that their problems are not severe enough to be assessed or treated for severe mental disorder. For the visitation model to succeed in providing better mental health care, it is necessary to havea well-functioning cooperation between the relevant departments in the municipalities and CAMHS. In the same way, it is important that the children with help-seeking parents who are not offered an early intervention due to the mental health problems being deemed too mild are instead appropriately informed and are offered low intensity interventions such as general guidance and self-help interventions. It is also important that appropriate information is provided about the possible early warning signs indicating that the child should be assessed again, so that children who start developing problems can get timely help.

As the visitation model was implemented in the context of the MMM trial, there was not much emphasis on the interventions for the two groups assessed as having too mild and too severe mental health problems. Instead, we relied on the municipalities to have appropriate interventions in place for these groups. With the municipalities normally being responsible for prevention and referral to CAMHS, this viewpoint could be deemed appropriate. It could, however, be argued that with the systematic approach of identifying children in the need of both low intensity intervention and fast referral to CAMHS, there is likely to be an increase in the number of children being identified. Hence, it could be beneficial for the municipalities to have a set of defined low intensity interventions in stock ready for the help-seeking parents, and for the municipalities to have a clearly defined cooperation with the CAMHS about children with indications of a severe mental or developmental disorder in order to ensure fast access to specialized assessment in the mental health services. .

Another important limitation of the evaluation presented in paper 2 is the lack of an economic evaluation. The current data does not allow for a full health economic evaluation of the implementation of the visitation model, as this would require both data on the benefits of the post visitation interventions and data on how comparable children would have been treated, or not treated, without the visitation model. If the visitation model is fully implemented in a municipality, it is possible that this could be done in a meaningful way using register data and comparing changes over time in a quasi-experimental difference-in-difference design, where the population of the municipality is compared with a population from a comparable municipality that has not introduced the visitation model.

As the evaluation presented in paper 2 found that the visitation model identified a large group of children with a need for early treatment or assessment at CAMHS, it could be argued that it is unethical to await years of research before implementing the visitation model on a larger scale. Instead of waiting for a full health economic evaluation, the budget impact for the municipalities could instead be evaluated to assess whether the costs are low enough for decision makers to accept the uncertainty around the cost-effectiveness. The budget impact assessment would need to take into consideration the cost of a one-day training in psychopathology that was provided to the local psychologist, the time spent by the psychologists, the time spent cooperating with CAMHS, as well as the cost of maintaining the IT system.

4.4 Cost-utility analysis using CHU9D

The results presented in paper 3 showed that CHU9D is an appropriate preference-based HRQoL measure for use in mental health trials. The results showed that CHU9D was both able to distinguish between groups with different severity of mental health problems, and able to capture changes in severity of mental health problems. These are key features in making CHU9D a relevant instrument for estimating QALYs for CUAs.

There are, however, some limitations to our findings. We used SDQ and KIDSCREEN as the standard that CHU9D is compared against. The optimal standard to compare against in the examination of validity is the 'gold' standard for the assessment of the issue in question. Within psychopathology, the 'gold' standard could arguably be interviews and observations by specialist rather than parent-reported questionnaires. Hence, a validation where this kind of data is available could strengthen the evidence of CHU9D's validity in mental health settings. Another limitation is that our study only examines the parent-reported CHU9D and only against parent-reported SDQ and KIDSCREEN. As presented in section 1.8, the answers from children and parents are not likely to be interchangeable and an examination of the longitudinal validity of self-reported

CHU9D would therefore further strengthen the evidence base. In our data, this examination was hindered by the lack of self-reported SDQ data for children younger than 11 years, as the instrument is not validated for younger children. Hence, other standards should be used in such a study.

The EQ-5D-3L from the EuroQol Group is among the most used preference-based HRQoL instruments for adults. Many HTA institutions, including The National Institute for Health and Care Excellence (NICE) in England (97), prefers that submitted CUAs uses the EQ-5D-3L to estimate QALYs. This ensures comparability of the results across applications, unlike if various instruments are used. There is still no clear preferences from HTA institutions, including NICE, on which instruments are preferred when estimating QALYs in child population (127). It could be suggested that the EQ-5D-Y is likely to become the preferred instrument by HTA institutions as it is a direct adoption of the EQ-5D-3L where only the wordings have been changed to be more suitable for children. Hence, there is the same number of health states which might make CUA results more comparable across child and adult populations. On the other hand, there are, as presented in section 1.9, different considerations that need to be taken when deciding on whose preferences the preference weights should be based on when estimating QALYs. Hence, it is argued that the EQ-5D-3L preference weights cannot be used for the EQ-5D-Y, and a valuation protocol for EQ-5D-Y has therefore recently been developed (128). If different preference weights are applied, the argument of EQ-5D-Y being preferred due to comparability across childhood and adulthood is no longer valid (77). Instead, it may be argued that CHU9D is a superior instrument as it is the only instrument developed from its inception to measure the most important dimensions of HRQoL for children (81). Alternatively, instruments like the HUI-2 and HUI-3 which can be used across child and adult populations could be preferred (77). In a comparison of CHU9D and EQ-5D-Y where HSUVs were estimated with preference weights derived from adult populations, the two measures showed a good level of agreement with an intraclass correlation coefficient of 0.8 (85).

In paper 3, we found that the HSUV generated by using the preferences derived from the adolescent population had a better construct validity and responsiveness than the HSUV generated using adult preferences when comparing the magnitude of the mean differences between the defined groups. Besides being derived from population groups of different ages, the two sets of preference weights also have important methodological differences. The estimated difference could, however, be explained by age-dependent difference in the preferences of mental health related states, as the younger populations put stronger weight on impairment of the CHU9D dimensions that can be considered most relevant for mental health (worried, sad, annoyed) (87). Our findings suggest that the choice of preference weights could have a considerable effect on the results of CUAs.

This hypothesis was confirmed in paper 4 where we found the ICER to be 71% higher when based on the QALYs estimated using the preference weights derived from the adult population compared to the one based on the adolescent population. This result highlights the importance of which set of preference weights that are used in CUA. If decision makers put any emphasis on the cost-effectiveness results, such a variation can easily be the difference between introducing or not introducing a new intervention. For HTA institutions that strives for comparability between different CUAs, it is essential that these are based on comparable sets of preference weights. The differences between deriving preference weights for adults' and children's health states which were presented in section 1.8 constitutes a challenge for HTA institutions, as it can be argued that preference weights derived from adult populations should not be used in child populations. On the other

hand, using these weights might make it easier to ensure consistency across different CUAs independent of the populations.

Another interesting result found in paper 4 was that the ICER was heavily impacted by who provided information on the CHU9D. We found a 24% higher ICER when QALYs were estimated with self-reported CHU9D scores compared with parent-reported. As discussed in section 1.7, it is difficult to make a clear argument for one or the other being more correct in a mental health context. Instead, it might be more appropriate to take both results into consideration and assume the cost-effectiveness to be within the interval of the two results. For HTA institutions, this finding also highlights an important issue, as it shows that QALY estimates based on self-reporting and parent-reporting are not interchangeable. Thus, there needs to be consistency in the choice of respondent across CUAs of interventions targeting children or youth to ensure comparability.

4.5 The cost-effectiveness of Mind My Mind

In paper 4, we found that the MMM intervention lead to significant incremental costs. The incremental costs were due to the costs of the MMM intervention. There were considerable costs associated with the training of the psychologists, which was partly due to the training being delivered at a hotel with full-board accommodation for five days. It could be argued that the training could be delivered in a less costly way, e.g. online training. Part of the idea behind having the training at a hotel was, however, that it would ensure that the participating psychologists were fully focused on the training without being distracted by other tasks. The hope from the organizers was that five days of focus on MMM would ensure good implementation. Thus, it is possible that the same effects would not have been found if the training was delivered in a less costly way instead.

Another concern in the cost estimation is that the full investment of the training was not realized in the trial period, as the psychologist over time could treat more children than they were able to do in the restricted trial period, which would make the cost per treated child lower. A sensitivity analysis, presented in the paper, showed that the total incremental costs decreased by 33% when assuming each psychologist would treat 20 children instead of the average of 8.2 from the trial-period. When increasing the number of treated children per psychologist, the impact was further decreased, e.g. the total incremental costs only decreased by around 7% percentage points further when assuming a rate of 100 children per psychologist instead of 20. There was a considerable turnover in PPR which indicated that the actual number of treated children per psychologist probably would be less than 100, but still closer to 20, than to the observed mean of 8.2. This would mean that we underestimated the cost-effectiveness of MMM. The impact of the caseload on the total costs was, however, small compared to the impact of the duration of the treatment effect.

The analyses of HRQoL showed an adjusted mean QALY gain of 0.17 (95% CI: 0.006-0.029) for the MMM group in the base case. When analyzing the one-year extrapolation scenarios, we found significantly higher QALY gains. This revealed that the cost-effectiveness of MMM is very dependent on the duration of the treatment effect. The time horizon of the analyses is further discussed in section 4.4.3.

We found an ICER of €170,465 in the base case analysis. The decision of whether this is to be considered cost-effective is a normative question, that ultimately depends on the decision makers' willingness-to-pay (WTP). Only a few HTA institutions have an official explicit WTP (129), and those who have it, also have varying principles that affects the WTP. A consultancy has, however, tried to estimate the WTP of a range of HTA

institutions around the world (129). Despite the inevitable uncertainty around their estimates, their findings confirm that an ICER of €170,465 must be considered considerably higher than HTA institutions' WTP. On the other hand, the ICERs in the range of €23,653 to €50,480, that we found analyzing the extrapolation scenarios would be considered cost-effective by several HTA institutions, according to the estimates from the consultancy (129). Different principles can be argued to result in higher WTP by HTA institutions. In the case of MMM, decision makers' WTP might be influenced by the fact that it is a treatment for children and that the alternative of the MMM is no evidence-based treatment as shown by the limited extent of psychological interventions in the MAU group.

There are several limitations of the analyses presented in paper 4 that can be discussed. The most fundamental concerns the choices of comparator, the perspective and the time horizon of the analyses.

4.5.1 Comparator

The choice of competing alternative intervention in a CUA is essential for the usefulness of the results. It is important that the relevant comparator or comparators are being compared with the new intervention to give a correct understanding of the cost-effectiveness (72). For example, if a new intervention is compared with an intervention that would not be considered cost-effective itself, the new intervention can wrongly be found cost-effective. An important consideration in paper 4 is whether other alternatives than MAU should have been considered when evaluating the cost-effectiveness of MMM.

The choice of MAU as comparator follows the principles of comparing the intervention with the current treatment regime. MAU was, in the trial, enhanced by the offer of two care-coordination visits at the Educational-psychological advisory service. This was done to ensure that the MAU population was definitely getting some offer after the mental health problems of the children had been assessed severe enough for an intervention. What we found was that relatively little MAU activity was going on in the trial period besides these care-coordination visits. We found that 35 (18%) reported that they from baseline to end-of-treatment (week 18) took part in individual therapy (n=25), and/or group therapy (n=8), or parental training (n=3) in the municipalities. The rest did not report any MAU psychological interventions in this period. If the care-coordination visits, which 95% participated in, are not considered, this means that MAU for the majority resembled a waitlist condition more than an active intervention, when strictly looking at psychological interventions. These finding raise the question if other alternatives should have been considered.

As presented in section 1.3, there are several evidence-based CBT-programs available for children with common mental health problems. As an example, the Cool Kids program for children with anxiety has been evaluated positively in a Danish clinical setting (130). The problem with using Cool Kids or other diagnosis specific interventions as comparators is, however, that the children in the trial population have different kinds of mental health problems within the spectrum of anxiety, depression and behavioral problems. Many even have problems in more than one of the domains. This would, in practice, mean that the relevant comparator would be not one but three different interventions that would be needed to deal with the spectrum of problems. Considering that a significant part of the costs in the delivery of MMM was due to the training of psychologists and supervisors, it does not seem feasible that multiple programs that would involve training in each of the programs would be cost-effective. Furthermore, it seems unlikely that it would be feasible to implement several new programs in the real world.

Instead, other transdiagnostic programs could be considered. The development of transdiagnostic CBT-programs is, however, relatively new, and the evidence of the effectiveness of other transdiagnostic programs is still limited with mixed results, partly depending on the study design and the implementation strategies (54,55,131). Furthermore, these programs are restricted in use, making a wide implementation in Denmark less feasible compared to MMM, that was designed with this in mind.

Based on these considerations, it seems that MAU is the correct comparator to evaluate the cost-effectiveness of MMM. The relatively low activity in the MAU group, and thereby negligible cost, also makes MAU likely to be a cost-effective intervention on its own. Thus, in regard to the comparator, the analyses presented in paper 4 seem sufficient to evaluate the cost-effectiveness of MMM.

4.5.2 Perspective

The analyses in paper 4 apply an extended health care sector perspective, meaning that costs of health care interventions are included irrespective of whether they are publicly financed or payed out-of-pocket by the parents. Limiting the perspective to that of the health care sector can be problematic as mental health problems are associated with major costs outside the health care sector (56). It could be argued that the CUA should have a societal perspective to inform decision makers in the best possible way. To assess the impact of the narrower perspective, it is important to take the time horizon of the analyses into consideration, as it is only costs that would differ across the intervention and control group during the time of observation that should be included in a CUA. Although it is possible that we would have observed differences in the cost of the educational sector and social services, it is also likely that we would require a longer time horizon to capture substantial differences in these costs. Given the long-term nature of the mental health problems, it is less plausible that for example the school would decide to react on effects from the intervention within half a year. A societal cost that could be affected within the relatively short time horizon is the lost productivity of the children's parents.

Data on self-reported time spent by the parents on MMM, MAU, and other health care visits were collected during the trial. In addition, the parents were asked to report being late at work, leaving early and days of staying home due to the mental health problems of the child. The data was not considered in paper 4 due to the perspective of the CUA. Analyses of these self-reported data did, however, show that parents in the MMM group spent nine hours on average attending MMM training with their children, while the MAU parents spent 2.7 hours on average attending MAU activities. The MAU parents did, however, take more time off work compared to the MMM parents. In total, the MMM parents spent 4.3 hours on average more during the trial period. This meant that we, by applying the health care sector perspective rather than the societal perspective, slightly underestimated the incremental costs of MMM.

4.5.3 Time horizon

An important limitation is the time horizon of the CUA. The base case analysis that was restricted to the trial period will inevitably underestimate the effect of MMM as it is not realistic, with the large observed mean difference at follow-up (week 26), that the MMM and the MAU group would have the same mean HRQoL immediately after the last observation. In the period from end-of-treatment at week 18 to the follow-up at week 26 where only the booster-session is being provided, we actually observed an increasing mean difference in HRQoL between the groups in favor of the MMM group. The lack of long-term follow-up data is an

important limitation for the results in paper 4, as it generates uncertainty about the trend in effects beyond the relatively short observation period.

The limited time horizon is, in general, a limitation seen in studies concerning cost-effectiveness and children's mental health (63). CUAs should, in general, have a time-horizon long enough to capture all important differences in both costs and effects between the compared interventions (72). It is, however, a common challenge when analyzing the cost-effectiveness of a new intervention, that the follow-up data does not capture the full effects. Decision-analytic modeling for health economic evaluation have been developed to meet this challenge by providing a mathematical framework that allows for data from multiple sources to be incorporated in cost-effectiveness analyses (132). This makes it possible to estimate the cost-effectiveness of new interventions taking into account life-long effects of new interventions by using external data, e.g. epidemiological studies, to estimate the future non-observed effects. CUAs based on decision-analytic models are, therefore, usually preferred over trial-based CUA when pharmaceutical companies submit applications to HTA institutions. An important limitation of decision-analytic models is the availability of appropriate external data. In the cost-effectiveness analysis of MMM, we assessed that there was no appropriate external data that would allow us to use decision-analytic modelling. The idea and ambition of the MMM intervention is that it will not only have a treatment effect on the current symptoms, but also have a preventive effect on the development of severe mental disorder later in life. There are epidemiological studies that supports this ambition based on the trial data, as the psychometric instruments that were used in the trial to conclude the clinical superiority of MMM over MAU, also have been shown to predict future risk of mental disorders in observational cohort studies (102). The problem with using this kind of epidemiological evidence in a cost-effectiveness model is that it has not been shown that changes in the psychometric scores caused by interventions have the same associations with later risk of mental disorders. It could potentially be a temporary change we observe in the trial period, which will not have that kind of impact on the children's future mental health.

A fundamental difference between a CUA of a not-for-profit intervention, like MMM, and a CUA of a for-profit intervention, e.g. a pharmaceutical product, for an application to an HTA institution is that there is no pricing involved in the assessment of MMM. When a pharmaceutical company negotiates the price of a new drug, the company wants to show the full effects of the drug, which are often lifetime, to be able to justify their price and ultimately, get the highest possible price for their product. As there is no such price to be negotiated in the case of MMM, it could be sufficient for decision makers to know that the cost per gained QALY is below their willingness-to-pay.

We chose the approach of analyzing four different possible one-year extrapolation scenarios. The one-year time period was chosen based on the available evidence from other studies of CBT treatments for common mental health problems among children, where there, in general, is little available long-term evidence due to lack of data. Some studies are, however, able to show statistically significant effects one year after end-of-treatment (42–49). In this way, we took a relatively conservative approach not considering any preventive effects. The choice of one-year scenarios is, however, uncertain and the sensitivity analyses show that the ICER do vary considerably if the scenarios are assumed to be shorter or longer. Future studies of the MMM trial population could be beneficial in determining the actual long-term effects of the intervention. The Danish national registers will be able to provide a unique possibility of avoiding loss to follow-up. Like previous studies, there is, however, the potential challenge with contamination of the control group.

4.6 Future directions

The findings presented in this thesis have implications for both future research and hopefully, for future implementation of procedures to help children with mental health problems.

We found that that there are many children who suffer from mental health problems, who do not receive appropriate care and who can be identified at an early stage. We found that it is possible to identify the children by self-referral by their parents, and we described a well-functioning visitation model that provides a systematic approach to stratifying the children so that they can be offered an intervention of appropriate intensity.

Future research could focus on how to ensure equal access for all children to the visitation model and ultimately, to the appropriate interventions. Also, appropriate cooperation models between the 98 Danish municipalities and the five regional CAMHS could be examined and explored to ensure that there is a coherent and fast process for children who experience severe mental health problems.

It is a huge task to implement a whole new systematic visitation on a wide national scale, and it would almost certainly pose challenges for the involved organizations. However, with the already available evidence, it seems obvious to consider the implementation of the visitation model in the Danish municipalities to provide a systematic approach to ensuring timely help to the many children suffering from mental health problems today.

The health economic evaluation of the MMM intervention did not provide a clear answer for decision makers on whether the intervention can be considered cost-effective. The results do, however, suggest that the cost per QALY-gained is not going to be far from standard willingness-to-pay levels. The considerable uncertainties about the future effects need to be considered, together with the evidence of MMM being clinically superior and that the children do not receive other evidence-based treatments today. The pragmatic design of the trial has proven that it is possible to implement MMM on a wider scale.

Future research could shed light on the long-term effects by following the children from the trial. The question is, however, whether decision makers want to wait for further evidence, that might not be more conclusive, while holding back the implementation of the intervention which has been shown to have positive effects. Another option is to begin the implementation of the MMM intervention under a protocol that allows for research. Cluster-randomized experiments could explore the potential of training the psychologists in a less costly way, thereby ensuring cost-effectiveness. Observational study designs could also be utilized to explore the long-term effects of MMM without holding back the intervention for a wider implementation.

Our findings support the usage of CHU9D in mental health settings. Future studies should investigate the validity of CHU9D in other patient populations including other specific mental disorders, e.g. common mental disorders in children including the neurodevelopmental disorders such as ADHD and autism. Our findings further underline the need for clear HTA guidelines for the estimation of QALYs in child populations. HTA institutions need to consider guidelines for both the informant of HRQoL and the source of preferences for the preference weights. More research providing empirical evidence on how the QALY estimation is impacted by these choices in other populations and possibly using other preference-based instruments could further help guide these decisions. Another interesting related future research area, that in this thesis was not

considered in depth, is how to model HRQoL and estimate QALYs when populations are followed from childhood into adulthood. This would require choices about both preference-based HRQoL instruments, and informants and preference-weights, and whether these choices should be consistent or changed over the course of the modelled period.

5. Conclusion

The evaluation of the screening algorithm based on parental SDQ responses showed that the children who screened positive had significantly poorer school outcomes when compared to the children who screened negative. The majority of the children who screened positive did not have a mental health consultation in the health care system during the five-year follow-up period and did not have higher health care costs. The children who screened positive and did have a mental health consultation, on average, had the first consultation after 2.5 years. These findings showed that there is a potential for identifying children in need of an early treatment for their mental health problems.

In paper 2 we found that the majority of the self-referred children had long-term mental health problems and met the diagnostic criteria of at least one mental disorder diagnosis or were assessed as having indications of severe mental disorders. The evaluation of the visitation model showed that it succeeded in stratifying the children into three groups with increasing severity of mental health problems eligible for three different types of interventions. The evaluation also revealed that older children, children of mothers with low education level and children with foreign-born parents were not as likely to be self-referred. It is critical that solutions to meet these problems are developed prior to wide scale implementation to ensure equal access.

The validity of CHU9D for use in a mental health setting was positively evaluated. Both the construct validity and the longitudinal validity showed that CHU9D is an appropriate instrument to include in mental health trials to enable the estimation of QALYs. Combined with the comprehensive evidence of validity from cross-sectional studies, the positive assessment suggests that CHU9D is, in general, an appropriate preference based HRQOL instrument for use in child populations. The examination of the two different sets of preference weights showed that the choice of preference weights can have a significant impact on the results of CUAs conducted in mental health settings. This was found to be true for the CUA of the MMM intervention where the results showed that both the choice of preference weights and the choice of respondent to CHU9D had major impact on the results. It is crucial that this impact is taken into consideration when conducting and evaluating CUAs.

The MMM intervention was found to have positive effects on the HRQoL of children with mental health problems compared to MAU, but the intervention generated added costs. Whether MMM can be considered cost-effective is a normative question. The analyses of extrapolation scenarios showed that the cost-effectiveness is dependent on the duration of the treatment effect. All four scenarios of what could happen during the year after the last observation did result in ICERs below € 51,000 per QALY gained. Hence, it seems plausible that MMM would be considered cost-effective, even when not taking into account possible preventive effects on mental disorders later in life.

Early treatment of mental health problems in children remain a developing area in health care. The negative consequences are huge for the children, their families and society. This thesis has provided empirical evidence that supports the introduction of a systematic identification, visitation and early treatment of children's mental health problems. Despite the limitations and unknown factors, the presented evidence suggests that it is time for a change of focus from randomized controlled trials to execution and implementation of evidence-based early treatment.

6. References

1. Gore FM, Bloem PJ, Patton GC, Ferguson J, Joseph V, Coffey C, et al. Global burden of disease in young people aged 10–24 years: a systematic analysis. The Lancet. 2011 Jun;377(9783):2093–102.

2. Polanczyk GV, Salum GA, Sugaya LS, Caye A, Rohde LA. Annual Research Review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. Journal of Child Psychology and Psychiatry. 2015 Mar;56(3):345–65.

3. Dalsgaard S, Thorsteinsson E, Trabjerg BB, Schullehner J, Plana-Ripoll O, Brikell I, et al. Incidence Rates and Cumulative Incidences of the Full Spectrum of Diagnosed Mental Disorders in Childhood and Adolescence. JAMA Psychiatry. 2020 Feb 1;77(2):155.

4. Copeland WE, Wolke D, Shanahan L, Costello EJ. Adult Functional Outcomes of Common Childhood Psychiatric Problems: A Prospective, Longitudinal Study. JAMA Psychiatry. 2015 Sep 1;72(9):892.

5. Balázs J, Miklósi M, Keresztény Á, Hoven CW, Carli V, Wasserman C, et al. Adolescent subthreshold-depression and anxiety: psychopathology, functional impairment and increased suicide risk: Adolescent subthreshold-depression and anxiety. Journal of Child Psychology and Psychiatry. 2013 Jun;54(6):670–7.

6. Suhrcke M, Pillas D, Selai C. Economic aspects of mental health in children and adolescents In: Social cohesion for mental well-being among adolescents. Copenhagen: World Health Organization, Regional Office for Europe; 2008.

7. Romeo R, Knapp M, Scott S. Economic cost of severe antisocial behaviour in children - and who pays it. British Journal of Psychiatry. 2006 Jun;188(6):547–53.

8. Järbrink K, Fombonne E, Knapp M. Measuring the Parental, Service and Cost Impacts of Children with Autistic Spectrum Disorder: A Pilot Study. Journal of Autism and Developmental Disorders. 2003;33(4):10.

9. Beecham J. Annual Research Review: Child and adolescent mental health interventions: a review of progress in economic studies across different disorders. Journal of Child Psychology and Psychiatry. 2014 Jun;55(6):714–32.

10. Fatori D, Salum G, Itria A, Pan P, Alvarenga P, Rohde LA, et al. The economic impact of subthreshold and clinical childhood mental disorders. Journal of Mental Health. 2018 Apr 28;1–7.

11. Costello EJ, Copeland W, Angold A. Trends in psychopathology across the adolescent years: What changes when children become adolescents, and when adolescents become adults?: Trends in psychopathology across the adolescent years. Journal of Child Psychology and Psychiatry. 2011 Oct;52(10):1015–25.

12. Copeland WE, Adair CE, Smetanin P, Stiff D, Briante C, Colman I, et al. Diagnostic transitions from childhood to adolescence to early adulthood. Journal of Child Psychology and Psychiatry. 2013 Jul;54(7):791–9.

13. Reef J, van Meurs I, Verhulst FC. Children's Problems Predict Adults' DSM-IV Disorders Across 24 Years. Adolescent Psychiatry. 2010;49(11):8.

14. Hill RM, Pettit JW, Lewinsohn PM, Seeley JR, Klein DN. Escalation to Major Depressive Disorder among Adolescents with Subthreshold Depressive Symptoms: Evidence of Distinct Subgroups at Risk. J Affect Disord. 2014 Apr;158:133–8.

15. Shankman SA, Lewinsohn PM, Klein DN, Small JW, Seeley JR, Altman SE. Subthreshold conditions as precursors for full syndrome disorders: A 15-year longitudinal study of multiple diagnostic classes. J Child Psychol Psychiatry. 2009 Dec;50(12):1485–94.

16. Patton GC, Coffey C, Romaniuk H, Mackinnon A, Carlin JB, Degenhardt L, et al. The prognosis of common mental disorders in adolescents: a 14-year prospective cohort study. The Lancet. 2014 Apr;383(9926):1404–11.

17. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime Prevalence and Age-of-Onset Distributions of DSM-IV Disorders in the National Comorbidity Survey Replication. Archives of General Psychiatry. 2005 Jun 1;62(6):593.

18. Bloom DE, Cafiero E, Jané-Llopis E, Abrahams-Gessel S, Bloom LR, Fathima S, et al. The Global Economic Burden of Noncommunicable Diseases [Internet]. Program on the Global Demography of Aging; 2012 Jan [cited 2018 Jul 24]. (PGDA Working Papers). Report No.: 8712. Available from: https://ideas.repec.org/p/gdm/wpaper/8712.html

19. Sundhedsstyrelsen. Sygdomsbyrden i Danmark. København: Sundhedsstyrelsen; 2015.

20. Kataoka SH, Zhang L, Wells KB. Unmet Need for Mental Health Care Among U.S. Children: Variation by Ethnicity and Insurance Status. American Journal of Psychiatry. 2002 Sep;159(9):1548–55.

21. Sturm R, Ringel JS, Andreyeva T. Geographic Disparities in Children's Mental Health Care. Pediatrics. 2003 Oct 1;112(4):e308–e308.

22. Sundhedsdatastyrelsen. Udvalgte nøgletal for sygehusvæsenet og praksisområdet 2009-2018 [Internet]. København; 2019. Available from: https://sundhedsdatastyrelsen.dk/da/tal-og-analyser/analyser-og-rapporter/sundhedsvaesenet/noegletal-om-sundhedsvaesenet

23. National Audit Office. Improving children and young people's mental health services [Internet]. Department of Health & Social Care, NHS England and Health Education England; 2018. Available from: https://www.nao.org.uk/wp-content/uploads/2018/10/Improving-children-and-young-peoples-mental-health-services.pdf

24. Arango C, Díaz-Caneja CM, McGorry PD, Rapoport J, Sommer IE, Vorstman JA, et al. Preventive strategies for mental health. The Lancet Psychiatry. 2018 Jul;5(7):591–604.

25. Davis III TE, May A, Whiting SE. Evidence-based treatment of anxiety and phobia in children and adolescents: Current status and effects on the emotional response. Clinical Psychology Review. 2011;31(4):592–602.

26. Compton SN, March JS, Brent D, Albano AM, Weersing VR, Curry J. Cognitive-behavioral psychotherapy for anxiety and depressive disorders in children and adolescents: An evidence-based medicine review. Journal of the American Academy of Child and Adolescent Psychiatry. 2004;43(8):930–59.

27. Higa-McMillan CK, Francis SE, Rith-Najarian L, Chorpita BF. Evidence Base Update: 50 Years of Research on Treatment for Child and Adolescent Anxiety. Journal of Clinical Child and Adolescent Psychology. 2016;45(2):91–113.

28. Reynolds S, Wilson C, Austin J, Hooper L. Effects of psychotherapy for anxiety in children and adolescents: A meta-analytic review. Clinical Psychology Review. 2012;32(4):251–62.

29. Mychailyszyn MP, Brodman DM, Read KL, Kendall PC. Cognitive-Behavioral School-Based Interventions for Anxious and Depressed Youth: A Meta-Analysis of Outcomes. Clinical Psychology: Science and Practice. 2012;19(2):129–53.

30. James A, James G, Cowdrey F, Soler A, Choke A. Cognitive behavioural therapy for anxiety disorders in children and adolescents (Review). Cochrane Database of Systematic Reviews. 2018;(2).

31. Schwartz C, Barican J Lou, Yung D, Zheng Y, Waddell C. Six decades of preventing and treating childhood anxiety disorders: a systematic review and meta-analysis to inform policy and practice. Evidence Based Mental Health. 2019;22(3):103–10.

32. David-Ferdon C, Kaslow NJ. Evidence-Based Psychosocial Treatments for Child and Adolescent Depression. Journal of Clinical Child & Adolescent Psychology. 2008 Mar 3;37(1):62–104.

33. Zhou X, Hetrick SE, Cuijpers P, Qin B, Barth J, Whittington CJ, et al. Comparative efficacy and acceptability of psychotherapies for depression in children and adolescents: A systematic review and network meta-analysis. World Psychiatry. 2015 Jun;14(2):207–22.

34. Hetrick SE, Cox GR, Witt KG, Bir JJ, Merry SN. Cognitive behavioural therapy (CBT), third-wave CBT and interpersonal therapy (IPT) based interventions for preventing depression in children and adolescents. Cochrane Database of Systematic Reviews. 2016 Aug 9;(8).

35. Eckshtain D, Kuppens S, Ugueto A, Ng MY, Vaughn-Coaxum R, Corteselli K, et al. Meta-Analysis: 13-Year Follow-up of Psychotherapy Effects on Youth Depression. Journal of the American Academy of Child & Adolescent Psychiatry. 2019 Apr;

36. Stice E, Shaw H, Bohon C, Marti CN, Rohde P. A meta-analytic review of depression prevention programs for children and adolescents: Factors that predict magnitude of intervention effects. Journal of Consulting and Clinical Psychology. 2009;77(3):486–503.

37. Weisz JR, McCarty CA, Valeri SM. Effects of psychotherapy for depression in children and adolescents: A meta-analysis. Psychological Bulletin. 2006;132(1):132–49.

38. Hetrick SE, Cox GR, Merry SN. Where to go from Here? An Exploratory Meta-Analysis of the Most Promising Approaches to Depression Prevention Programs for Children and Adolescents. International Journal of Environmental Research and Public Health. 2015;12(5):4758–95.

39. Cox GR, Callahan P, Churchill R, Hunot V, Merry SN, Parker AG, et al. Psychological therapies versus antidepressant medication, alone and in combination for depression in children and adolescents. Cochrane Database of Systematic Reviews. 2014 Nov 30;2014(11. Art. No.: CD008324).

40. Arnberg A, Öst LG. CBT for Children with Depressive Symptoms: A Meta-Analysis. Cognitive Behaviour Therapy. 2014;43(4):275–88.

41. Forti-Buratti MA, Saikia R, Wilkinson EL, Ramchandani PG. 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. 2016;25(10):1045–54.

42. Waddell C, Schwartz C, Andres C, Barican JL, Yung D. Fifty years of preventing and treating childhood behaviour disorders: a systematic review to inform policy and practice. Evidence Based Mental Health. 2018 May;21(2):45–52.

43. Furlong M, McGilloway S, Bywater T, Hutchings J, Smith SM, Donnelly M. Cochrane Review: Behavioural and cognitive-behavioural group-based parenting programmes for early-onset conduct problems in children aged 3 to 12 years (Review). Evidence-Based Child Health: A Cochrane Review Journal. 2013 Mar 7;8(2):318–692.

44. Schwartz C, Barican JL, Yung D, Zheng Y, Waddell C. Six decades of preventing and treating childhood anxiety disorders: a systematic review and meta-analysis to inform policy and practice. Evidence Based Mental Health. 2019 Aug;22(3):103–10.

45. Zhou X, Zhang Y, Furukawa TA, Cuijpers P, Pu J, Weisz JR, et al. Different Types and Acceptability of Psychotherapies for Acute Anxiety Disorders in Children and Adolescents. JAMA Psychiatry. 2019 Jan;76(1):41–50.

46. Zhou X, Hetrick SE, Cuijpers P, Qin B, Barth J, Whittington CJ, et al. Comparative efficacy and acceptability of psychotherapies for depression in children and adolescents: A systematic review and network meta-analysis. World Psychiatry. 2015 Jun;14(2):207–22.

47. Hetrick SE, Cox GR, Witt KG, Bir JJ, Merry SN. Cognitive behavioural therapy (CBT), third-wave CBT and interpersonal therapy (IPT) based interventions for preventing depression in children and adolescents. Cochrane Common Mental Disorders Group, editor. Cochrane Database of Systematic Reviews [Internet]. 2016 Aug 9 [cited 2020 Apr 3]; Available from: http://doi.wiley.com/10.1002/14651858.CD003380.pub4

48. Eckshtain D, Kuppens S, Ugueto A, Ng MY, Vaughn-Coaxum R, Corteselli K, et al. Meta-Analysis: 13-Year Follow-up of Psychotherapy Effects on Youth Depression. Journal of the American Academy of Child & Adolescent Psychiatry. 2020 Jan;59(1):45–63.

49. James AC, James G, Cowdrey FA, Soler A, Choke A. Cognitive behavioural therapy for anxiety disorders in children and adolescents. Cochrane Common Mental Disorders Group, editor. Cochrane Database of Systematic Reviews [Internet]. 2015 Feb 18 [cited 2020 Apr 3]; Available from: http://doi.wiley.com/10.1002/14651858.CD004690.pub4

50. Kendall P, Hedtke K. Cognitive-behavioral therapy for anxious children: Therapist manual. 3rd Editio. Child and Adolescent Anxiety Disorders Clinic Temple University; 2006.

51. Rapee R, Lyneham H, Schniering C, Wuthrich V, Abbott M, Hudson J. The Cool Kids: Child and Adolescent Anxiety Program. Anxiety Research Unit Macquarie University; 2006.

52. McHugh RK, Murray HW, Barlow DH. Balancing fidelity and adaptation in the dissemination of empirically-supported treatments: The promise of transdiagnostic interventions. Behaviour Research and Therapy. 2009;47(11):946–53.

53. Marchette LK, Weisz JR. Practitioner Review: Empirical evolution of youth psychotherapy toward transdiagnostic approaches. Journal of Child Psychology and Psychiatry and Allied Disciplines. 2017;58(9):970–84.

54. Chorpita BF, Weisz JR, Daleiden EL, Schoenwald SK, Palinkas LA, Miranda J, et al. Long-term outcomes for the Child STEPs randomized effectiveness trial: A comparison of modular and standard treatment designs with usual care. Journal of Consulting and Clinical Psychology. 2013;81(6):999–1009.

55. Weisz JR, Chorpita BF, Palinkas LA, Schoenwald SK, Miranda J, Bearman SK, et al. Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth: A randomized effectiveness trial. Archives of General Psychiatry. 2012;69(3):274–82.

56. Hamilton MP, Hetrick SE, Mihalopoulos C, Baker D, Browne V, Chanen AM, et al. Identifying attributes of care that may improve cost-effectiveness in the youth mental health service system. Medical Journal of Australia [Internet]. 2017 Nov [cited 2020 Apr 3];207(S10). Available from: https://onlinelibrary.wiley.com/doi/abs/10.5694/mja17.00972

57. Creswell C, Violato M, Cruddace S, Gerry S, Murray L, Shafran R, et al. A randomised controlled trial of treatments of childhood anxiety disorder in the context of maternal anxiety disorder: clinical and cost-effectiveness outcomes. J Child Psychol Psychiatry. 2020 Jan;61(1):62–76.

58. Creswell C, Violato M, Fairbanks H, White E, Parkinson M, Abitabile G, et al. Clinical outcomes and cost-effectiveness of brief guided parent-delivered cognitive behavioural therapy and solution-focused brief therapy for treatment of childhood anxiety disorders: a randomised controlled trial. The Lancet Psychiatry. 2017 Jul;4(7):529–39.

59. Lynch FL, Dickerson JF, Clarke G, Beardslee W, Robin Weersing V, Gladstone TRG, et al. Cost-Effectiveness of Preventing Depression in At-Risk Youth: Post-Intervention and Two-Year Follow-up. Psychiatr Serv. 2019 Apr 1;70(4):279–86.

60. Tie H, Krebs G, Lang K, Shearer J, Turner C, Mataix-Cols D, et al. Cost-effectiveness analysis of telephone cognitive–behaviour therapy for adolescents with obsessive–compulsive disorder. BJPsych Open [Internet]. 2019 Jan 10 [cited 2020 Apr 14];5(1). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6343121/

61. Jolstedt M, Wahlund T, Lenhard F, Ljótsson B, Mataix-Cols D, Nord M, et al. Efficacy and cost-effectiveness of therapist-guided internet cognitive behavioural therapy for paediatric anxiety disorders: a single-centre, single-blind, randomised controlled trial. The Lancet Child & Adolescent Health. 2018 Nov;2(11):792–801.

62. Dickerson JF, Lynch FL, Leo MC, DeBar LL, Pearson J, Clarke GN. Cost-effectiveness of Cognitive Behavioral Therapy for Depressed Youth Declining Antidepressants. Pediatrics [Internet]. 2018 Feb [cited 2020 Apr 14];141(2). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5810604/

63. Knapp M, Wong G. Economics and mental health: the current scenario. World Psychiatry. 2020 Feb;19(1):3–14.

64. van Straten A, Hill J, Richards DA, Cuijpers P. Stepped care treatment delivery for depression: a systematic review and meta-analysis. Psychological Medicine. 2015 Jan;45(2):231–46.

65. Cross S, Hickie I. Transdiagnostic stepped care in mental health. Public Health Research & Practice. 2017;27(2).

66. McGorry P, van Os J. Redeeming diagnosis in psychiatry: timing versus specificity. The Lancet. 2013 Jan;381(9863):343–5.

67. Silverman WK, Pettit JW, Lebowitz ER. Stepping Toward Making Less More for Concerning Anxiety in Children and Adolescents. Clinical Psychology: Science and Practice. 2016 Sep;23(3):234–8.

68. Cross SPM, Hermens DF, Scott EM, Ottavio A, McGorry PD, Hickie IB. A Clinical Staging Model for Early Intervention Youth Mental Health Services. Psychiatric Services. 2014 Jul;65(7):939–43.

69. House of Commons, Committee of Public Accounts. Mental health services for children and young people [Internet]. the House of Commons; 2019 [cited 2019 Sep 19]. Available from: https://publications.parliament.uk/pa/cm201719/cmselect/cmpubacc/1593/1593.pdf

70. Braddick F, Carral V, Jenkins R, Jané-Llopis E. Child and Adolescent Mental Health in Europe: Infrastructures, Policy and Programmes. Luxembourg: European Communities; 2009.

71. Jeppesen P, Wolf RT, Nielsen SM, Christensen R, Plessen KJ, Bilenberg N, et al. Transdiagnostic Psychotherapy for Youth with Common Mental Health Problems. Submitted for publication.

72. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. Oxford University Press; 2015. 461 p.

73. Brouwer WBF, Culyer AJ, van Exel NJA, Rutten FFH. Welfarism vs. extra-welfarism. Journal of Health Economics. 2008 Mar;27(2):325–38.

74. Luyten J, Naci H, Knapp M. Economic evaluation of mental health interventions: an introduction to cost-utility analysis. Evidence Based Mental Health. 2016 May;19(2):49–53.

75. Danske Regioner - Metode i Medicinrådet ændres [Internet]. [cited 2020 May 28]. Available from: https://www.regioner.dk/services/nyheder/2019/november/metode-i-medicinraadet-aendres

76. Højgaard B, Wadmann S, Jakobsen M, Rasmussen SR, Pedersen NJM, Kjellberg J. Regulering af sygehusmedicin med udgangspunkt i omkostning og effekt. KORA; 2016.

77. Rowen D, Rivero-Arias O, Devlin N, Ratcliffe J. Review of Valuation Methods of Preference-Based Measures of Health for Economic Evaluation in Child and Adolescent Populations: Where are We Now and Where are We Going? PharmacoEconomics. 2020 Apr;38(4):325–40.

78. Hill H, Rowen D, Pennington B, Wong R, Wailoo A. A Review of the Methods Used to Generate Utility Values in NICE Technology Assessments for Children and Adolescents. Value in Health [Internet]. 2020 May [cited 2020 May 11]; Available from: https://linkinghub.elsevier.com/retrieve/pii/S1098301520301522

79. Mulhern B, Mukuria C, Barkham M, Knapp M, Byford S, Soeteman D, et al. Using generic preference-based measures in mental health: psychometric validity of the EQ-5D and SF-6D. British Journal of Psychiatry. 2014 Sep;205(03):236–43.

80. Brazier J. Is the EQ–5D fit for purpose in mental health? British Journal of Psychiatry. 2010 Nov;197(05):348–9.

81. Chen G, Ratcliffe J. A Review of the Development and Application of Generic Multi-Attribute Utility Instruments for Paediatric Populations. PharmacoEconomics. 2015 Oct;33(10):1013–28.

82. Wille N, Badia X, Bonsel G, Burström K, Cavrini G, Devlin N, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. Qual Life Res. 2010 Aug;19(6):875–86.

83. Ravens-Sieberer U, Wille N, Badia X, Bonsel G, Burström K, Cavrini G, et al. Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study. Qual Life Res. 2010 Aug;19(6):887–97.

84. Stevens K. Developing a descriptive system for a new preference-based measure of health-related quality of life for children. Quality of Life Research. 2009 Oct;18(8):1105–13.

85. Chen G, Flynn T, Stevens K, Brazier J, Huynh E, Sawyer M, et al. Assessing the Health-Related Quality of Life of Australian Adolescents: An Empirical Comparison of the Child Health Utility 9D and EQ-5D-Y Instruments. Value in Health. 2015 Jun;18(4):432–8.

86. Petersen KD, Chen G, Mpundu-Kaambwa C, Stevens K, Brazier J, Ratcliffe J. Measuring Health-Related Quality of Life in Adolescent Populations: An Empirical Comparison of the CHU9D and the PedsQLTM 4.0 Short Form 15. The Patient - Patient-Centered Outcomes Research. 2018 Feb;11(1):29–37.

87. Ratcliffe J, Huynh E, Stevens K, Brazier J, Sawyer M, Flynn T. Nothing About Us Without Us? A Comparison of Adolescent and Adult Health-State Values for the Child Health Utility-9D Using Profile Case Best-Worst Scaling: ADOLESCENT VERSUS ADULT VALUES FOR THE CHU-9D. Health Economics. 2016 Apr;25(4):486–96.

88. Ratcliffe J, Stevens K, Flynn T, Brazier J, Sawyer M. An assessment of the construct validity of the CHU9D in the Australian adolescent general population. Quality of Life Research. 2012 May;21(4):717–25.

89. Stevens K, Ratcliffe J. Measuring and Valuing Health Benefits for Economic Evaluation in Adolescence: An Assessment of the Practicality and Validity of the Child Health Utility 9D in the Australian Adolescent Population. Value in Health. 2012 Dec;15(8):1092–9.

90. Petersen KD, Ratcliffe J, Chen G, Serles D, Frøsig CS, Olesen AV. The construct validity of the Child Health Utility 9D-DK instrument. Health and Quality of Life Outcomes [Internet]. 2019 Dec [cited 2020 Apr 6];17(1). Available from: https://hqlo.biomedcentral.com/articles/10.1186/s12955-019-1256-0

91. Furber G, Segal L. The validity of the Child Health Utility instrument (CHU9D) as a routine outcome measure for use in child and adolescent mental health services. Health and Quality of Life Outcomes. 2015;13(1):22.

92. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. Journal of Clinical Epidemiology. 2007 Jan;60(1):34–42.

93. Sourander A, Helstelä L, Helenius H. Parent-adolescent agreement on emotional and behavioral problems. Social Psychiatry and Psychiatric Epidemiology. 1999 Dec 31;34(12):657–63.

94. Khadka J, Kwon J, Petrou S, Lancsar E, Ratcliffe J. Mind the (inter-rater) gap. An investigation of self-reported versus proxy-reported assessments in the derivation of childhood utility values for economic evaluation: A systematic review. Social Science & Medicine. 2019 Nov 1;240:112543.

95. Mott DJ, Shah KK, Ramos-Goñi JM, Devlin NJ, Rivero-Arias O. Valuing EQ-5D-Y health states using a discrete choice experiment: do adult and adolescent preferences differ? OHE Research Paper London: Office of Health Economics [Internet]. 2019; Available from: https://www.ohe.org/publications/valuing-eq-5d-yhealth-states-using-discrete-choice-experiment-do-adult-and-adolescent

96. Ratcliffe J, Huynh E, Chen G, Stevens K, Swait J, Brazier J, et al. Valuing the Child Health Utility 9D: Using profile case best worst scaling methods to develop a new adolescent specific scoring algorithm. Social Science & Medicine. 2016 May;157:48–59.

97. NICE. The reference case | Guide to the methods of technology appraisal 2013 | Guidance | NICE [Internet]. [cited 2020 Jun 28]. Available from: https://www.nice.org.uk/process/pmg9/chapter/the-reference-case

98. Ungar WJ, Gerber A. The uniqueness of child health and challenges to measuring costs and consequences in 'Economic Evaluation in Child Health'. New York: Oxford University Press; 2010.

99. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. Scandinavian Journal of Public Health. 2011 Jul;39(7_suppl):30–3.

100. Goodman R. The Extended Version of the Strengths and Difficulties Questionnaire as a Guide to Child Psychiatric Caseness and Consequent Burden. J Child Psychol Psychiatry. 1999;40(5):791.

101. Goodman A, Goodman R. Strengths and Difficulties Questionnaire as a Dimensional Measure of Child Mental Health. Journal of the American Academy of Child & Adolescent Psychiatry. 2009 Apr;48(4):400–3.

102. Nielsen LG, Rimvall MK, Clemmensen L, Munkholm A, Elberling H, Olsen EM, et al. The predictive validity of the Strengths and Difficulties Questionnaire in preschool age to identify mental disorders in preadolescence. Hashimoto K, editor. PLOS ONE. 2019 Jun 3;14(6):e0217707.

103. Niclasen J, Teasdale TW, Andersen A-MN, Skovgaard AM, Elberling H, Obel C. Psychometric Properties of the Danish Strength and Difficulties Questionnaire: The SDQ Assessed for More than 70,000 Raters in Four Different Cohorts. Scott JG, editor. PLoS ONE. 2012 Feb 27;7(2):e32025.

104. Goodman R, Meltzer H, Bailey V. The strengths and difficulties questionnaire: A pilot study on the validity of the self-report version. Adolescent Psychiatry. 1998;7(3):7.

105. Goodman R. The Strengths and Difficulties Questionnaire: A Research Note. Journal of Child Psychology and Psychiatry. 1997;38(5):581–6.

106. Meltzer H, Gatward R, Goodman R, Ford T. The Mental Health of Children and Adolescents in Great Britain: (622732007-001) [Internet]. American Psychological Association; 2000 [cited 2018 Jul 3]. Available from: http://doi.apa.org/get-pedoi.cfm?doi=10.1037/e622732007-001

107. Arnfred J, Svendsen K, Rask C, Jeppesen P, Fensbo L, Houmann T, et al. Danish norms for the Strengths and Difficulties Questionnaire. Dan Med J. 2019;66(6).

108. Woerner W, Becker A, Rothenberger A. Normative data and scale properties of the German parent SDQ. European Child & Adolescent Psychiatry. 2004 Jul;13(S2).

109. Gustafsson J-E, Westling AM, Åkerman A, Eriksson B, Eriksson C, Fischbein L, et al. School, learning and mental health: a systematic review. Stockholm: The Health Committee, Royal Swedish Academy of Sciences; 2010.

110. Kristoffersen JHG, Obel C, Smith N. Gender differences in behavioral problems and school outcomes. Journal of Economic Behavior & Organization. 2015 Jul;115:75–93.

111. Keilow M, Sievertsen HH, Niclasen J, Obel C. The Strengths and Difficulties Questionnaire and standardized academic tests: Reliability across respondent type and age. Goertz M, editor. PLOS ONE. 2019 Jul 25;14(7):e0220193.

112. French MT, Homer JF, Robins PK. What You Do in High School Matters: Eastern Economic Journal. 2015;41(3):370–86.

113. Gevrek D, Gevrek ZE, Guven C. Benefits of Education at the Intensive Margin: Childhood Academic Performance and Adult Outcomes among American Immigrants. Eastern Economic Journal. 2015 Jun;41(3):298–328.

114. Oreopoulos P, Salvanes KG. Priceless: The Nonpecuniary Benefits of Schooling. Journal of Economic Perspectives. 2011 Feb;25(1):159–84.

115. Daviss WB, Birmaher B, Melhem NA, Axelson DA, Michaels SM, Brent DA. Criterion validity of the Mood and Feelings Questionnaire for depressive episodes in clinic and non-clinic subjects. Journal of Child Psychology and Psychiatry. 2006;47(9):927–34.

116. Arendt K, Hougaard E, Thastum M. Psychometric properties of the child and parent versions of Spence Children's Anxiety Scale in a Danish community and clinical sample. Journal of Anxiety Disorders. 2014 Dec;28(8):947–56.

117. Eyberg SM, Pincus D. ECBI Eyberg Child Behavior Inventory and SESBI-R Sutter – Eyberg Student Behavior Inventory – Revised, Professional manual. Odessa, FL: Psychological Assessment Resources, Inc.; 1999.

118. Butlington V. Manual for the Child Behavior Checklist/ 4-18 and 1991 Profile. University of Vermont, Department of Psychiatry; 1991.

119. Goodman R, Scott S. Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: Is Small Beautiful? Journal of Abnormal Child Psychology. 1999;27(1):17–24.

120. Goodman A, Heiervang E, Collishaw S, Goodman R. The 'DAWBA bands' as an ordered-categorical measure of child mental health: description and validation in British and Norwegian samples. Social Psychiatry and Psychiatric Epidemiology. 2011 Jun;46(6):521–32.

121. Goodman R, Ford T, Richards H, Gatward R, Meltzer H. The Development and Well-Being Assessment: Description and Initial Validation of an Integrated Assessment of Child and Adolescent Psychopathology. J Child Psychol Psychiatry. 2000;41(5):645–55.

122. Dalsgaard S, Thorsteinsson E, Trabjerg BB, Schullehner J, Plana-Ripoll O, Brikell I, et al. Incidence Rates and Cumulative Incidences of the Full Spectrum of Diagnosed Mental Disorders in Childhood and Adolescence. JAMA Psychiatry [Internet]. 2019 Nov 20 [cited 2020 Jan 21]; Available from: https://jamanetwork.com/journals/jamapsychiatry/fullarticle/2755318

123. Wolf RT, Jeppesen P, Gyrd-Hansen D, The CCC2000 Study Group, Oxholm AS. Evaluation of a screening algorithm using the Strengths and Difficulties Questionnaire to identify children with mental health problems: A five-year register-based follow-up on school performance and healthcare use. Fonseca-Pedrero E, editor. PLOS ONE [Internet]. 2019 Oct 23 [cited 2019 Oct 24];14(10). Available from: http://dx.plos.org/10.1371/journal.pone.0223314

124. Steinhausen H-C, Jakobsen H. Incidence Rates of Treated Mental Disorders in Childhood and Adolescence in a Complete Nationwide Birth Cohort. The Journal of Clinical Psychiatry [Internet]. 2019 Apr 23 [cited 2019 Oct 9];80(3). Available from: https://www.psychiatrist.com/JCP/article/Pages/2019/v80/17m12012.aspx

125. Ades AE. Evaluating screening tests and screening programmes. Arch Dis Child. 1990 Jul;65(7):792–5.

126. Rimvall MK, Jeppesen P, Verhulst F. The Problem of Same-Rater Bias. Journal of the American Academy of Child & Adolescent Psychiatry. 2018 Sep;57(9):700–1.

127. Hill H, Rowen D, Pennington B, Wong R, Wailoo A. A review of the methods used to estimate and model utility values in NICE technology appraisals for paediatric populations. NICE DSU Report; 2019.

128. Ramos-Goñi JM, Oppe M, Stolk E, Shah K, Kreimeier S, Rivero-Arias O, et al. International Valuation Protocol for the EQ-5D-Y-3L. PharmacoEconomics. 2020 Jul;38(7):653–63.

129. Zhang K, Garau M. International cost-effectiveness thresholds and modifiers for HTA decision making [Internet]. Office of Health Economics: OHE Consulting Report, London; 2020. Available from: https://www.ohe.org/publications/international-cost-effectiveness-thresholds-and-modifiers-hta-decision-making

130. Arendt K, Thastum M, Hougaard E. Efficacy of a Danish version of the Cool Kids program: a randomized wait-list controlled trial. Acta Psychiatrica Scandinavica. 2016 Feb;133(2):109–21.

131. Weisz JR, Bearman SK, Ugueto AM, Herren JA, Evans SC, Cheron DM, et al. Testing Robustness of Child STEPs Effects with Children and Adolescents: A Randomized Controlled Effectiveness Trial. J Clin Child Adolesc Psychol. 2019 Sep 13;1–14.

132. Briggs A, Claxton K, Sculpher M. Decision Modelling for Health Economic Evaluation. New York: Oxford University Press; 2011.

7. Paper I-IV

Paper I

Evaluation of a screening algorithm using the Strengths and Difficulties Questionnaire to identify children with mental health problems: A five-year register-based follow-up on school performance and healthcare use.

PLOS ONE. 2019. 23;14(10)

PLOS ONE

RESEARCH ARTICLE

Evaluation of a screening algorithm using the Strengths and Difficulties Questionnaire to identify children with mental health problems: A five-year register-based follow-up on school performance and healthcare use

Rasmus Trap Wolf[1,2]*, **Pia Jeppesen**[2,3], **Dorte Gyrd-Hansen**[1], **The CCC2000 Study Group**[4¶], **Anne Sophie Oxholm**[1]

1 University of Southern Denmark, Department of Public Health, Danish Centre for Health Economics, Odense, Denmark, **2** Child and Adolescent Mental Health Centre, Mental Health Services Capital Region of Denmark, Copenhagen, Denmark, **3** Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, **4** Centre for Clinical Research and Prevention, Capital Region of Denmark, Copenhagen, Denmark

¶ Membership of the CCC2000 Study Group is provided in the Acknowledgments.
* rasmus.trap.wolf@regionh.dk

Citation: Wolf RT, Jeppesen P, Gyrd-Hansen D, The CCC2000 Study Group, Oxholm AS (2019) Evaluation of a screening algorithm using the Strengths and Difficulties Questionnaire to identify children with mental health problems: A five-year register-based follow-up on school performance and healthcare use. PLoS ONE 14(10): e0223314. https://doi.org/10.1371/journal.pone.0223314

Editor: Eduardo Fonseca-Pedrero, University of La Rioja, SPAIN

Received: August 1, 2019

Accepted: September 18, 2019

Published: October 23, 2019

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pone.0223314

Data Availability Statement: The data utilized in the current study are defined as sensitive personal

Abstract

Background

Treatment of mental health problems (MHP) is often delayed or absent due to the lack of systematic detection and early intervention. This study evaluates the potential of a new screening algorithm to identify children with MHP.

Methods

The study population comprises 2,015 children from the Copenhagen Child Cohort 2000 whose mental health was assessed at age 11–12 years and who had no prior use of specialised mental health services. A new algorithm based on the Strengths and Difficulties Questionnaire (SDQ) is utilised to identify MHP by combining parent-reported scores of emotional and behavioural problems and functional impairments. The screening is done on historical data, implying that neither parents, teachers nor health care professionals received any feedback on the screening status. The screening status and results of an IQ-test were linked to individual-level data from national registries. These national registers include records of each child's school performance at the end of compulsory schooling, their health care utilisation, as well as their parents' socio-economic status and health care utilisation.

Results

10% of the children screen positive for MHP. The children with MHP achieve a significantly lower Grade Point Average on their exams, independently of their IQ-score, perinatal factors and parental characteristics. On average, the children with MHP also carry higher health

data, and cannot be shared publicly due to existing data protection laws in Denmark, and imposed by the Danish Data Protection Agency. The collected data is uploaded and linked to data from the Danish National registers and analysed through Statistics Denmark (https://www.dst.dk/en). Interested researchers can access the data used in this study free of cost through a study description and research agreement with the Copenhagen Child Cohort 2000 steering committee, by way of contact Pia Jeppesen at pia.jeppesen@regionh.dk or Allan Linneberg at allan.linneberg@regionh.dk.

Funding: PJ and RTW received funding from Commercial Company: Trygfonden, https://www.trygfonden.dk. AMS (The CCC2000 Study Group) received funding from Commercial Company: Lundbeckfonden, https://www.lundbeckfonden.com. AMS (The CCC2000 Study Group) received funding from Commercial Company: Trygfonden, https://www.trygfonden.dk. The funders have no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

care costs over a five-year follow-up period. The higher health care costs are only attributed to 23% of these children, while the remaining children with MHP also show poorer school performance but receive no additional health care.

Conclusions

The results demonstrate that children with MHP and a poor prognosis can be identified by the use of the brief standardised questionnaire SDQ combined with a screening algorithm.

Introduction

Today more than 13% of children and adolescents worldwide are affected by mental disorders at any given time [1] with a significantly negative impact on their quality of life [2]. Several studies show that children's mental health problems lead to a higher risk of mental disorders later in life [3–9]. The median age of onset of anxiety, mood and behavioural disorders occur before age 14 years [10]. Despite the progress in the development and positive evaluation results of early treatment programs for mental health problems in childhood [11,12], such programs are not offered systematically and on a wide scale in Denmark or, to the best of our knowledge, in any other countries.

There is a growing awareness of the potentials of preventive interventions to improve children's mental health [11]. Specifically, on the potential of implementing them in 'stepped-care models' that offer interventions on increasing intensity levels ranging from general counseling to specialised mental health care, and with indicated prevention and early treatments as an intermediate step to children with common emotional and behavioral problems [13]. A systematic and wide-scale implementation of a stepped care model for service delivery needs to consider the complexities of childhood mental health problems. Children usually do not present with a recent-onset and well-defined single disorder. More commonly, the children have a long history of several problems, distress and impairments below or above diagnostic thresholds [13]. Thus, the 'matching' of the intervention to the children's fluctuating and complex problems is not straight forward. A systematic approach to identifying the children in need of interventions that are beyond prevention and counseling is therefore a necessary part of a successful stepped-care model.

For more than a decade it has been discussed how to identify mental health problems at an early stage among school children [14]. As outlined in a recent review there exists several studies evaluating different screening programs [15]. However, the quality of these evaluations is poor. The majority of the evaluations are based on cross-sectional study designs, and the few existing cohort studies include short follow-up periods and/or are dependent on self-reported outcomes [15]. Furthermore, the existing evaluations focus on diagnosis or access to mental health care. This could be problematic outcomes, as not all children in need of mental health care are likely to receive it [16–18]. Moreover, children with problems below the threshold for diagnosis may still require mental health care to avoid major difficulties later in life [9,19].

This study evaluates the potential of a new screening algorithm to identify children with sub-threshold and undiagnosed mental disorders, in this study referred to as mental health problems (MHP), in a general population. A unique combination of historical birth cohort data and high-quality data from national registers is used to mimic the performance of screening by analysing the later school performance and health care utilisation of the screened children. As neither parents, teachers nor health care professionals receive any feedback on the

screening this allows for analyses of how the children who screen positive perform in school up to five years later and how they utilise the health care system in a naturalistic follow-up. The rich data allows analyses that controls for potential confounders, i.e. the children's IQ, perinatal factors and their parents' health care utilisation and socio-economic characteristics. The methods section provides the precise definitions of the potential confounders.

The screening is based on parent-reported answers to the Strength and Difficulties Questionnaire (SDQ) and an algorithm, that combines scores of emotional and behavioural problems with functional impairments and uses cut-off based on population studies. The children's later school performance is investigated as a primary indicator of the long-term outcome of the screening for MHP in preadolescence. As previous studies find an association between children's MHP and their school performance [20–22], and school performance and schooling are known to be comprehensive predictors for the rest of a child's life with respect to both wealth, health, and happiness [23–25], this indicator is well-suited for the evaluation of the screening. The children's health care service utilisation during the five-year follow-up period is analysed for two purposes. First, to investigate whether and when the children that screen positive for MHP get access to specialised mental health care. Thereby the potential for early treatment can be assessed. Second, to investigate whether a different health care utilisation pattern is observed for the children that screen positive in the health care system.

Materials and methods

Ethics approval and consent to participate

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/ or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The study was approved by The Danish Data Protection Agency (J.nr. 2010-41-4438) and by the Capital Region of Denmark (J.nr. 2007-58-0015). The National Committee on Health Research Ethics was consulted (J.nr. H-C-FSP-2010) in accordance with national guidelines. Participation was voluntary, and data was kept confidential. According to Danish legislation, individual informed consent or additional permission from the National Committee on Health Research Ethics is not required for registry-based studies. The parents gave oral informed consent to the study participation. Separate written consent was not required by the ethical committee. Moreover, the written information given to the participants stated that the participation in the study is voluntary, that by answering the questionnaires consent is given and consent can be withdrawn at any time.

Data

The Copenhagen Child Cohort. The Copenhagen Child Cohort (CCC2000) is a Danish birth cohort that includes all children born in 16 municipalities in the former Copenhagen County in the year 2000. In total, the cohort comprises 6,090 children, and constitutes 9% of all children born in Denmark that year. The cohort is representative of the general Danish population born that year in regards to both gender and a number of key perinatal factors such as age of parents and perinatal illness. Skovgaard et al. [26]provide further information on the cohort at birth. The main purpose of CCC2000 is to study the presentation, developmental pathways and risk mechanisms of psychopathology longitudinally from birth. As part of the follow up studies of the cohort, the parents of the children were invited to complete the "Strengths and Difficulties Questionnaire" (SDQ) at different points in time, i.e. when the children are aged 5–7 and 11–12 years. The present study uses SDQ responses from when the

children were aged 11–12 years. In total, parents of 2,126 children (35%) completed SDQ at the 11–12 years follow-up.

The screening algorithm. We use SDQ to define whether or not a child screens positive for MHP. SDQ is widely used and validated as a tool to identify and assess children with MHP in both clinical samples and in general population-based samples [27–32]. SDQ contains 25 items, which cover five subscales relating to the children's emotional problems, peer problems, behavioural problems, hyperactivity and pro-social behaviour. Responses to the subscales on emotional problems, peer problems, behavioural problems, and hyperactivity can be used to calculate a total difficulties score. Each subscale score ranges from 0–10, implying that the total difficulties score ranges from 0 to 40 [33].

We use an extended version of SDQ, which also includes an impact assessment to evaluate how much the identified mental difficulties interfere with the child's everyday life. A functional impairment score is calculated from five items on whether the difficulties upsets or distresses the child and how much the difficulties interfere with home life, friendships, classroom learning, and leisure activities. Each item is scored on a scale from 0 to 2. To score 1 or 2, the interference from the difficulties in that domain must be assessed to either "quite a lot" or "a great deal" [28,34].

The present study defines MHP as having a functional impairment score of at least 1 in combination with at least one of the following scores: a total difficulties score of at least 14 and/or an emotional problem score of at least 5 and/or a behavioural problem score of at least 3. This definition is based on age-matched populations in Denmark, Germany and the UK [35–37], such that scoring above the cut-off implies that the children's mental health state is ranked among the 10% of the poorest in these populations. This definition of MHP is also used in an on-going Danish randomized, clinical trial of an early treatment intervention for children with MHP (Trial ID: NCT03535805).

Based on the above criterion, we define children as either having MHP (screening positive) or not (screening negative) at time of reporting, i.e. at age 11–12 years. Thus, we do not distinguish between different levels of MHP in this study.

School performance. In Denmark, it is compulsory to attend school for nine years, from the age of 7 years to 15 years. The majority of Danish children attend public schools, which are fully funded through taxes. Private schools are also heavily subsidised. We retrieve information on each child's attendance and grades for each final exam from their compulsory schooling from the Danish Education Registers [38]. This information includes test scores up until summer 2017, where the children were between the age of 16.5 years and 17.5 years and thus, expected to have finished compulsory schooling.

At the end of compulsory schooling (9^{th} grade), there are eight mandatory exams for all pupils distributed across four subjects: Danish, Mathematics, English and Physics/Chemistry. Following the Danish Ministry of Education that uses at least four exams to calculate the national Grade Point Average (GPA), we classify children who have not attended at least four of the eight mandatory exams as not having completed the 9^{th} grade exams. Each exam results in a grade based on the Danish 7-point grading scale. The 7-point scale consists of seven grades: −03, 00, 02, 4, 7, 10, 12. These are equivalent to the following ECTS grades: F, Fx, E, D, C, B, A [39]. We calculate the children's GPA from the 9^{th} grade mandatory exams (results from at least four exams) as a school performance outcome.

Health care utilisation. In Denmark the health care system is predominantly publicly financed by taxes, and information on individuals' health care utilisation and associated unit costs are listed in administrative registers. This study defines health care utilisation as the costs incurred for both somatic and mental health care and in both the primary and secondary health care sector. As there is no available information on privately funded health care, we

only include the costs of publicly funded care, with the exception of out-of-pocket spending on prescription medicine.

We analyse health care costs for the years 2012 to 2016 (equivalent to the five years post "screening"). We analyse the health care costs both at an aggregated level as total health care costs and in subsets based on sectors and specific types of mental health care, as outlined in section 2.4.1–2.4.3. We adjust all costs to 2016-prices using price- and wage-regulations of consumption prices from the Danish Regions and Statistics Denmark [40,41]. We convert the costs from Danish kroner (DKK) to Euros (€) using the currency rate DKK 100 = € 13.42.

We retrieve the costs of health care in the primary health care sector from the Danish National Health Service Register [42]. The costs are based on providers' activity-based remuneration. We differentiate between specialised and non-specialised mental health treatment in primary care. Specialised mental health treatment includes services from psychiatrists and psychologists, whereas non-specialised mental health treatment includes services provided by all remaining primary health care providers, e.g. general practitioners.

We retrieve costs of prescription medication from the Danish National Prescription Registry [43]. This database contains information on all purchase of prescription medications. We use the retail price at the date of purchase to measure costs. Depending on the annual consumption, between 60% and 100% of the cost of prescription medication is subsidised. Hence, we analyse the total costs of prescription medication for each child, implying that we include both the public funding and out-of-pocket spending. As there is no available information on purchase of over-the-counter medication, we do not include costs of this type of medication.

Costs of hospital care are retrieved from the National Patient Register [44]. We differentiate between somatic care and mental health care in hospitals, as well as inpatient care and outpatient care. We compute costs of somatic care using the Danish System of Diagnostic Related Groups (DRG-tariffs). DRG-tariffs are unavailable for mental health care in Denmark. Instead, we use the National Board of Health's per diem charge and ambulatory charge for mental health care to assess the costs of mental health care in hospitals.

Covariates. School performance and health care utilisation may be affected by factors that may also be associated with MHP. In an effort to isolate the effect of MHP status on subsequent outcomes, the analyses include several covariates derived from Statistics Denmark's national registers. For the individual child, this includes dummies for the following: the child's gender, whether the child is the first-born in the family, whether the child was born with low-birth-weight (<2,400 gram) (LBW) and/or was small-for-gestational-age (SGA), whether the child is second-generation immigrant (there are non-first generation immigrants in the cohort), and whether the child lived with both biological parents the 1st of January 2012.

For both biological parents of the child, we include dummies for the following: whether the parent received mental health care in a hospital during the years 1995 and 2011, whether the highest education in the beginning of 2012 was primary school (up to 9 years education), whether the parent was unemployed in 2011 (defined as unemployed for at least six months), and whether the parent was among the 25% with the highest total health care costs in the CCC2000 population (split by gender) from 2000 to 2011.

Children's IQ is potentially a confounding variable when estimating the association between school performance and MHP. The Block Design (BD) test is a subtest of Wechsler Intelligence Scale for Children (WISC-III). The BD test is known to be highly correlated with IQ on the full Wechsler test of intelligence [45]. We therefore proxy the children's IQ using results from the BD test, which was done face-to-face for those children in the cohort who accepted the invitation. The tests were done at approximately the same time as SDQ responses were completed by the parent.

Research design and data analysis

Study population. The focus of this study is to evaluate the potential of the screening process, we therefore exclude children who had already received specialised mental health care (SMHC) before the date of the parent completing SDQ. SMHC is defined as any contact with psychologists or psychiatrists in the primary health care sector or any contact to a hospital mental health care department. In total, 111 children (5.2%) had received SMHC at some point prior to the parents completing SDQ. We exclude these children from the study population, resulting in a study population of 2,015 children.

Statistical analyses. In an effort to isolate the effects of MHP on school performance and health care utilisation, we control for the covariates listed in Section 2.5. However, for most of these covariates, it remains unclear whether they confound an association between MHP, school performance and health care costs, or if the single covariate represents an intermediary step on a causal pathway. All analyses are therefore done using three different models: 1) a crude model, 2) a model that includes covariates for the child's characteristics, and 3) a model that includes covariates for both the child's and the parents' characteristics.

For analyses of school performance, we also include the BD test score (IQ-proxy) in separate models, yielding two additional models: 4) a model including the IQ-proxy and the additional covariates for the child's characteristics and 5) a "full model" including the child's IQ-proxy and additional covariates for both the child's and the parents' characteristics. A total of 1,537 children (76.3% of the study population) completed a face-to-face BD test. Hence, the models that include the IQ-proxy are based on this smaller study population.

We also analyse the association between MHP and the outcomes based on whether the children with MHP received SMHC during the five-year follow-up period. For the group of children with MHP that received SMHC, we analyse whether there is a "window of opportunity" for early treatment of the children.

The comprehensive number of covariates in some of the models increases the risk of "over controlling" and thereby, possibly underestimating the associations. We therefore expect the true estimate to be somewhere in the range of the estimates from the different models. By presenting results from all models, we aim to explore the robustness of any significant associations found.

We measure the children's school performance using two different outcome measures: 1) completion of the 9^{th} grade exams and 2) the GPA from the 9^{th} grade exams. We employ a logit regression to estimate the association of MHP with the completion of the exams, whereas we use an Ordinary Least Squares (OLS) regression to estimate the association between MHP and the GPA, which is measured on a continuous scale.

We measure the children's health care utilisation in different parts of the health care sector. These data are heavily right-skewed, and some outcomes have a high proportion of zeros (corresponding to no consumption). We therefore employ both Generalized Linear Models (GLM) and two-part models combining a logit model with a GLM to estimate the association between MHP and the different types of health care utilisation.

We specify the GLM models based on the performance of GLMs with normal, gamma, Gaussian, and inverse Gaussian distributions combined with linear, log, power (0.5), and power (-1) links. Following Deb et al. [46], we choose the link function and distribution using Akaike Information Criteria, link-test, and Park's modified test. For the analysis of the children's total health care costs and primary health care sector care costs (without psychologist/psychiatrist), a GLM with log-link and gamma distribution performs best. For the rest of the health care costs outcomes, two-part models combining a logit and GLM with log-link and gamma distribution performs best.

Results

Descriptive statistics

Table 1 describes the study population split by whether or not the child screened positive for MHP at age 11–12 years. We categorise 9.6% of the full study population as having MHP. Several covariates are statistically significantly associated with having MHP. This includes being the first-born child in the family (including only child), not living with both biological parents, the mother having had contact to mental health hospital services, the mother's highest educational attainment being primary school, the mother being unemployed, and either parent being among the 25% of parents with the highest health care costs in decade leading up to the screening. In the study population for which an IQ-proxy test score is available, the percentage of children with MHP is slightly lower than in the full study population.

Table 2 shows descriptive statistics for the outcome variables. A larger share of children with MHP do not manage to complete the 9th grade exams, and amongst those who do complete the exams, a lower GPA is attained. Furthermore, the average total health care costs from 2012 to 2016 is higher for children with MHP, and a higher share receive SMHC in the five-

Table 1. Characteristics of the children and their parents in the study population, by their mental health problems status.

	Children without mental health problems N = 1,821 (90.4%)	Children with mental health problems N = 194 (9.6%)
Girl	51.8%	50.5%
First-born in the family	49.3%	57.7% **
Born with low-birth-weight or small-for-gestational-age	5.9%	8.3%
Second-generation immigrant[a]	10.5%	10.8%
Child did not live with both biological parents 1. Jan 2012	23.1%	38.7% ***
Mother had contact to the mental health hospital services between 1995 and 2011	7.6%	13.9% ***
Father had contact to the mental health hospital services between 1995 and 2011	5.4%	7.7%
Mother's highest education primary school (up to 9 years)	11.8%	17.0% **
Father's highest education primary school (up to 9 years)	16.2%	19.6%
Mother was unemployed in 2011	5.3%	11.3% ***
Father was unemployed in 2011	4.5%	7.2% *
Mother among the 25% of the CCC2000 mothers with highest health care costs between 2000 and 2011	20.2%	30.4% ***
Father among the 25% of the CCC2000 fathers with highest health care costs between 2000 and 2011 in the study population	23.0%	29.9% **
N (study population with IQ-proxy)	**1,395 (90.8%)**	**142 (9.2%)**
IQ proxy (Mean Block Design score (SE))	9.78 (0.10)	9.18 (0.34) *

Note: The table describes the characteristics of children and their parents in the study population by their children's mental health problem status. Differences in the characteristics across the children's mental problem status is tested using a Chi-square test and two-sided t-test for Block Design score.

[a]There are no first generation immigrants in the cohort.

*p-level <0.1

**p-level <0.05

***p-level <0.01.

https://doi.org/10.1371/journal.pone.0223314.t001

Table 2. School performance and health care utilisation of the study population based on their mental health problems status.

	N	Children without mental health problems N = 1,821 (90.4%)	Children with mental health problems N = 194 (9.6%)
Have not completed the 9th grade exams	2,015	3.1%	9.3%
Mean grade point average (GPA) 9th grade exams	1,941	7.72	6.45
Mean total health care cost (€) 2012–2016	2,015	4,785	10,829
Received SMHC 2012–2016	2,015	5.6%	22.7%

Note: The table summarises the school performance and health care utilisation of the study population based on their mental health status. Having completed the 9th grade exams is defined by having attended at least four of the eight mandatory exams. Total health care costs include all publicly funded care in both the primary and secondary health care sector. Specialised mental health care (SMHC) is defined as a contact with a publicly funded psychiatrist or psychologist.

https://doi.org/10.1371/journal.pone.0223314.t002

year period. We present statistical tests results for associations between these outcomes and MHP in the following sections.

School performance

The association between MHP and school performance is estimated using two different school outcomes. The first school outcome measures completion of the 9th grade exams. Table 3

Table 3. The association between mental health problems at age 11–12 years and not completing the 9th grade exams.

	N	Mental health problems at age 11–12 years	
		OR	SE
Crude model	2,015	3.22 ***	0.91
Model including child's characteristics[a]	2,015	3.14 ***	0.91
Model including child's characteristics[a] and IQ-proxy	1,537	1.71	0.68
Model including child's and parent's characteristics[a,b]	2,015	3.00 ***	0.87
Model including child's, parent's characteristics[a,b], and IQ-proxy	1,537	1.57	0.65

Note: The table contains estimated Odds Ratios (OR) of having a mental health problem at age 11–12 years and not completing the 9th grade exams. Logistic regression models are used to estimate OR with standard error (SE). As not all children have an IQ-test score the model including this population has fewer observations. We define not having completed the 9th grade exams as not attending at least four of the eight mandatory exams. Estimates for all covariates are available in Table A in S1 Table.

[a] Gender, first-born child in the family, born small-for-gestational-age or with low-birth-weight, children are second generation immigrants, child did not live with both biological parents 1st January 2012.

[b] Mother and/or father had contact to mental health hospital services between 1995 and 2011, mother's and/or father's highest education on 1st January 2012 was primary school (up to 9 years schooling), mother and/or father was unemployed in 2011, mother and/or father was among the parents of the full CCC2000 cohort with the 25% highest health care cost between 2000 and 2011.

*p-level <0.1

**p-level <0.05

***p-level <0.01.

https://doi.org/10.1371/journal.pone.0223314.t003

shows the results. The three models without the IQ-proxy suggest that children with MHP have statistically significant higher odds of not completing the 9th grade exams. This difference is, however, statistically insignificant when controlling for the children's IQ.

The second school outcome is the GPA for the children, who have completed the 9th grade exams. Table 4 shows a statistically significant negative association between MHP and the GPA. Across models, the marginal effect ranges from -0.98 grade point to -1.27 grade point.

Health care utilisation

Table 5 presents the results of the analyses of the association between MHP and the health care utilisation in the subsequent five years after the SDQ-screening (2012 to 2016). The association between MHP and total health care utilisation is statistically significant and ranges from €5,858 to €7,336 across models. We split the total health care utilisation into seven subcategories. The analyses of the subcategories provide information on which type of health care service drives the differences in the total health care costs. The results show that the higher costs for children with MHP are primarily driven by higher inpatient and outpatient mental hospital care. However, costs related to prescription medicines and utilisation in the primary sector, both specialised mental care and other types of care, are also positively associated with MHP. There is no statistically significant association found between MHP and utilisation of somatic hospital care during the five-year period of analysis.

Table 5 also shows that the results are robust across all models. The only exception is the association between MHP and utilisation of mental inpatient hospital care, which is statistically insignificant when controlling for both children's and parents' characteristics. This

Table 4. The association between mental health problems at age 11–12 years and Grade Point Average at 9th grade exams.

	N	Mental health problems at age 11–12 years	
		ME	SE
Crude model	1,941	-1.27 ***	0.19
Model including child's characteristics[a]	1,941	-1.13 ***	0.18
Model including child's characteristics[a] and IQ-proxy	1,486	-1.03 ***	0.19
Model including child's and parent's characteristics[a,b]	1,941	-1.04 ***	0.18
Model including child's, parent's characteristics[a,b], and IQ-proxy	1,486	-0.98 ***	0.19

Note: The table contains estimated coefficients (in grade points) for the association between mental health problems at age 11–12 years and Grade Point Average (GPA) at 9th grade exams. The study population includes children who have done at least four of the 9th grade exams. As not all children have an IQ-test score the model including this population has fewer observations. Ordinary Least Square regression models are used to estimate Marginal effect (ME) with robust standard error (SE) of mental health problems on GPA. Estimations for all covariates are available in Table B in S1 Table.

[a] Gender, first-born child in the family, born small-for-gestational-age or with low-birth-weight, children are second generation immigrants, child did not live with both biological parents 1st January 2012.

[b] Mother and/or father had contact to mental health hospital services between 1995 and 2011, mother's and/or father's highest education on 1st January 2012 was primary school (up to 9 years schooling), mother and/or father was unemployed in 2011, mother and/or father was among the parents of the full CCC2000 cohort with the 25% highest health care cost between 2000 and 2011.

*p-level <0.1

**p-level <0.05

***p-level <0.01.

https://doi.org/10.1371/journal.pone.0223314.t004

Table 5. The associations between mental health problems at age 11–12 years and health care costs in the subsequent five years.

	Crude models		Models including child's characteristics[c]		Models including child's and parent's characteristics[c,d]	
	Mental health problems at age 11–12 years					
	ME	SE	ME	SE	ME	SE
Total health care costs[a]	6,044 ***	2,254	7,336 ***	2,301	5,858 ***	1,713
Primary sector care without psychologist/psychiatrist[a]	180 **	73	176 **	72	149 **	72
Psychologist/psychiatrist in primary care sector[a]	202 *	104	171 **	73	280*	161
Prescription medicine[b]	161 **	78	140 *	77	148 **	71
Mental outpatient hospital care[b]	717 ***	208	772 ***	223	696 ***	222
Mental inpatient hospital care[b]	4,395 **	1,965	4,755 **	1,940	12,413[e]	13,872 [e]
Somatic outpatient hospital care[b]	332	262	324	250	182	219
Somatic inpatient hospital care[b]	55	609	686	660	904	690

Note: The table contains estimated coefficients (in €) of the association between mental health problems at age 11–12 years and health care costs in the five year-period. The number of observations N = 2,015. Estimations for all covariates in the analyses of total health care costs are available in Table C in S1 Table.

[a] Marginal effect (ME) and robust standard error (SE) estimated with a GLM model.

[b] Marginal effect (ME) and robust standard error (SE) estimated with a two-part model combining a logit and a GLM model.

[c] Gender, first-born child in the family, born small-for-gestational-age or with low-birth-weight, children are second generation immigrants, child did not live with both biological parents 1st January 2012.

[d] Mother and/or father had contact to mental health hospital services between 1995 and 2011, mother's and/or father's highest education on 1st January 2012 was primary school (up to 9 years schooling), mother and/or father was unemployed in 2011, mother and/or father was among the parents of the full CCC2000 cohort with the 25% highest health care cost between 2000 and 2011.

[e] Several variables were omitted in the GLM part of the two-part model due to lack of observations.

*p-level <0.1

**p-level <0.05

***p-level <0.01.

https://doi.org/10.1371/journal.pone.0223314.t005

finding can, however, be explained by the relatively few children utilising mental inpatient hospital care compared to the large number of covariates in the model, which generates a large standard error.

Health care utilisation and specialised mental health care treatment. Table 5 shows that SMHC utilisation is the main driver of the higher health care utilisation of children with MHP. However, as reported in Table 2, only 22.7% of the children with MHP receive SMHC during the five-year follow-up period. To explore whether the children with MHP, who did not receive SMHC, receive more of other types of health care services, we split the analysis of the health care costs of children with MHP by whether or not they received SMHC. In these analyses, the reference group is still the "children with no MHP".

Table 6 shows the results of the crude models, i.e. without controlling for child and parental characteristics. We find that the higher total health care costs for the whole group of children with MHP is driven by the group of children with MHP who received SMHC. Besides generating high costs through their use of SMHC, this group of children also have higher costs due to utilisation of the primary health care sector and prescription medicine, and they tend to use somatic outpatient health care services more. The larger group of children with MHP who did not receive SMHC (77.3%) do not receive more health care services of any kind compared to

Table 6. The associations between mental health problems at age 11–12 years (with/without specialised mental health care) and health care cost in the subsequent five years.

	Crude models			
	Children with mental health problems who receive no specialised mental health care (n = 150)		Children with mental health problems who receive specialised mental health care (n = 44)	
	ME	SE	ME	SE
Total health care costs[a]	-918	930	29,778 ***	8,325
Primary sector care without psychologist/psychiatrist[a]	126	84	365 ***	139
Psychologist/psychiatrist in primary care sector[b]	NA	NA	1,004 **	435
Prescription medicine[b]	-18	62	775 ***	244
Mental outpatient hospital care[b]	NA	NA	3,853 ***	727
Mental inpatient hospital care[b]	NA	NA	22,043 ***	7,917
Somatic outpatient hospital care[b]	157	292	929 *	508
Somatic inpatient hospital care[b]	-166	673	809	828

Note: The table contains estimated coefficients (in €) of the association between mental health problems at age 11–12 years and health care costs in the five year-period. We run separate analysis based on whether or not the children with MHP receive specialised mental health care. The reference group is children with no MHP. The number of observations N = 2,015. Some estimates are non-applicable (NA) based on the definition of the group. Specialised mental health care (SMHC) is defined as contact with publicly funded psychiatrist or psychologist. The findings are robust to including controls for both child and parental characteristics, see Table D in S1 Table.

[a] Marginal effect (ME) and robust standard error (SE) estimated with a GLM model.

[b] Marginal effect (ME) and robust standard error (SE) estimated with a two-part model combining a logit and a GLM model.

*p-level <0.1

**p-level <0.05

***p-level <0.01.

https://doi.org/10.1371/journal.pone.0223314.t006

children with no MHP. The findings are robust to including controls for both child and parental characteristics, see Table D in S1 Table.

School performance and specialised mental health care treatment. The results of the analyses of health care utilisation show that 77.3% of the children with MHP do not receive any SMHC and do not utilise the health care system more than children without MHP during the five-year follow-up period. To investigate whether this subgroup of children with MHP have an unmet need for care, we use the children's later school performance as an indicator of the consequences of their MHP at the age of 11–12 years. Thus, finding a negative association between having MHP without receiving SMHC and later school performance, compared to children with no MHP, will indicate that there is an unmet need for care.

We compare the school performance between children without MHP and the two groups of children with MHP, who are divided by whether or not they received SMHC during the five-year follow-up period. Tables 7 and 8 show the results. The group of children with MHP who did not receive SMHC does not have higher odds of not completing the 9th grade exams. The models without IQ-proxy shows a tendency towards a higher risk, but this tendency is statistically insignificant when controlling for IQ. However, the group of children with MHP who received SMHC have statistically significant higher odds of not having completed the 9th grade exams. The estimated odds ratio is almost halved after controlling for IQ. The statistical significance level is, however, independent of the chosen model.

Table 7. The associations between mental health problems at age 11–12 years with/without specialised mental health care and not having completed the 9th grade exams.

	N	Children with mental health problems who receive no specialised mental health care (n = 150/109)[c]		Children with mental health problems who receive specialised mental health care (n = 44/33)[c]	
		OR	SE	OR	SE
Crude model	2,015	2.01 *	0.74	8.10 ***	3.22
Model including child's characteristics[a]	2,015	1.90 *	0.72	8.99 ***	3.61
Model including child's characteristics[a] and IQ-proxy	1,537	1.04	0.57	4.91 ***	2.66
Model including child's and parent's characteristics[a,b]	2,015	1.93 *	0.75	7.67 ***	3.10
Model including child's, parent's characteristics[a,b], and IQ-proxy	1,537	1.00	0.56	4.18 ***	2.26

Note: The table contains estimated Odds Ratios (OR) of having a mental health problem at age 11–12 years and not completing the 9th grade exam. Logistic regression models are used to estimate OR with standard error (SE) for not having completed the 9th grade exams. Not having completed the 9th grade exams is defined by not having done at least four of the eight mandatory exams. Specialised mental health care (SMHC) is defined as contact with publicly funded psychiatrist or psychologist.

[a] Gender, first-born child in the family, born small-for-gestational-age or with low-birth-weight, children are second generation immigrants, child did not live with both biological parents 1st January 2012.

[b] Mother and/or father had contact to mental health hospital services between 1995 and 2011, mother's and/or father's highest education on 1st January 2012 was primary school (up to 9 years schooling), mother and/or father was unemployed in 2011, mother and/or father was among the parents of the full CCC2000 cohort with the 25% highest health care cost between 2000 and 2011.

[c] The sample size is lower in models which include IQ-proxy.

*p-level <0.1

**p-level <0.05

***p-level <0.01.

https://doi.org/10.1371/journal.pone.0223314.t007

Table 8. The associations between mental health problems at age 11–12 years with/without specialised mental health care and Grade Point Average at 9th grade exams.

	N	Mental health problems and no specialised mental health care (n = 141/105)[c]		Mental health problems and specialised mental health care (n = 35/29)[c]	
		ME	SE	ME	SE
Crude model	1,941	-1.29 ***	0.21	-1.17 ***	0.45
Model including child's characteristics[a]	1,941	-1.07 ***	0.20	-1.35 ***	0.42
Model including child's characteristics[a] and IQ-proxy	1,486	-0.95 ***	0.21	-1.31 ***	0.42
Model including child's and parent's characteristics[a,b]	1,941	-0.98 ***	0.19	-1.30 ***	0.40
Model including child's, parent's characteristics[a,b], and IQ-proxy	1,486	-0.90 ***	0.20	-1.26 ***	0.41

Note: Ordinary Least Square regression models are used to estimate Marginal effect (ME) with robust standard error (SE) of mental health problems on Grade Point Average (GPA) at 9th grade exams. Specialised mental health care (SMHC) is defined as contact with publicly funded psychiatrist or psychologist.

[a] Gender, first-born child in the family, born small-for-gestational-age or with low-birth-weight, children are second generation immigrants, child did not live with both biological parents 1st January 2012.

[b] Mother and/or father had contact to mental health hospital services between 1995 and 2011, mother's and/or father's highest education on 1st January 2012 was primary school (up to 9 years schooling), mother and/or father was unemployed in 2011, mother and/or father was among the parents of the full CCC2000 cohort with the 25% highest health care cost between 2000 and 2011.

[c] The N is lower in models which include IQ-proxy.

*p-level <0.1

**p-level <0.05

***p-level <0.01.

https://doi.org/10.1371/journal.pone.0223314.t008

In the analyses of GPA, shown in Table 8, we find a statistically significant negative association between MHP and GPA independently of whether or not the child has received SMHC. Amongst the children who have MHP but do not receive SMHC, the GPA score is between -0.90 and -1.29 grade points lower than non-MHP children, while it is between -1.17 and -1.35 grade points lower for the children with MHP who do receive SMHC.

Window of opportunity and diagnoses. To analyse the potential for earlier treatment of the children with MHP who receive specialised care within the five-year follow-up period, we analyse whether they receive the care with some delay compared to the timing of screening. The results include the 44 children with mental health problems and minimum one contact with SMHC. We find that the mean number of days between the parents' completion of SDQ and their children's first contact with SMHC was 945 days, and the median was 1,000 days. The 25% and 75% percentiles were 426 days and 1,311 days, respectively. Thus, the results show that both the average and median time span was more than two and a half years from when the parent responded to SDQ to the first contact with any SMHC.

Diagnoses are available for the children in the study population that received SMHC at a hospital. In total 16.8% of the children with MHP received a psychiatric diagnosis during the five-year follow-up period compared to 3.6% of the children we classify as not having MHP at age 11–12 years.

Of the 44 children with MHP receiving SMHC during the five-year follow-up period, 33 children (75%) received SMHC at a hospital. Thus, diagnoses are available for these children. They receive a broad variety of diagnoses: Hyperactivity disorders, adjustment reactions, pervasive developmental disorders, depressive disorders, psychotic disorders, and anxiety disorders are among the most frequent diagnoses. To comply with the data protection regulation, we refrain from reporting all specific diagnosis for this relatively small sample of individuals.

Discussion

Using parents' responses to SDQ and the screening algorithm, we categorise 9.6% of the 11-12-year-old children in our study population as having MHP. We find that the children with MHP independent of potential confounders perform statistically significantly worse at the 9th grade exams, which are conducted about five years after the screening. This finding is in line with previous studies that have found an association between SDQ scores and school outcomes [21,22]. The results indicate that the screening algorithm succeed in identifying children that may benefit from an early intervention.

Children with MHP also exhibit a statistically significant higher utilisation of health care during the five-year follow-up period. Sub-group analyses show that the higher health care utilisation is entirely driven by the 22.7% of the children with MHP who receive SMHC during the follow-up period. The children we identified with MHP, who receive SMHC during the follow-up period, have their first contact with specialised care on average 2.5 years after the parent's completion of SDQ. This indicates that there for the majority of the group is a 'window-of-opportunity' for an early intervention.

The 77.3% of children with MHP who do not receive any SMHC during the follow-up period also obtain a statistically significantly worse GPA at their 9th grade exams than children with no MHP. However, these children do not receive any other additional health care services, which indicates that they have unmet need for care throughout the follow-up period.

We find that the children with MHP that receive SMHC have worse school outcome than the children with MHP that do not receive SMHC. This finding should, however, not necessarily be interpreted as the SMHC having a negative effect on the children's long-term outcome. It is more likely that the difference is due to confounding by indication, meaning that those

who are affected the most from the MHP also are more likely to get access to SMHC doing the follow-up.

The results show that the screening algorithm combined with parental-responses to SDQ enables us to identify of a group of children with MHP and a poor prognosis based on their later school performances. In Denmark and all other countries, to the best of our knowledge, evidence-based early treatment programs are not being offered systematically. The results of our study indicate that when no systematic screening is in place, the majority of the children with MHP receive no care despite of a poorer prognosis. Furthermore, those who get access to SMHC get it with a delay, potentially due to the MHP having progressed in severity thus making them eligible for SMCH. In a future implementation of a stepped-care model a screening procedure, like the one presented, could play an important first step in the 'matching' of intervention and severity of MHP.

Strengths and limitations

A key strength of this study is the unique combination of cohort data including the face-to-face IQ-proxy test combined with the full register data on the children's school performance and health care utilisation, as well as information on the broad range of socio-economic and health-related covariates related to both the children and their parents. Despite the small population groups in some analyses, the results remain robust in all analyses independent of the inclusion of potential confounders, which further strengthens our findings.

There are, however, also some limitations. Due to data limitations, we were unable to include costs of possible visits and treatments at privately funded psychiatrists and psychologist. As publicly funded counselling or preventive measures vary across municipalities and are not systematically reported, we were also unable to include this information. However, the municipalities do not provide any systematic evidence-based practices. Their services can be considered as a first step in a stepped-care model and do not have the characteristics of early treatment. The missing information will therefore not change any of our conclusions. As we miss the information on private and municipality costs, the cost estimates should, however, be interpreted from a public health care sector perspective.

The study population is part of the CCC2000, which is representative of the Danish population on several key characteristics. However, as our study population is limited to children whose parents' respond to SDQ, our sample is not representative of the general population. On average, the study's parents have a higher educational attainment, are less unemployed, and have fewer contacts to mental health care services than the general population (see Table E in S1 Table). We do, however, not see any reason for our findings to be less relevant for children with parents from lower socioeconomic groups. As these parents are likely to have fewer resources to cope with the challenges of MHP, the negative impact of MHP on children's school performance could possibly be even worse for this group. Our sample size does not allow for comprehensive subgroup analysis of possible effect modifications between socioeconomic factors and MHP.

Screening tool and timing

This study evaluates the potential for identifying children, who could benefit from an early intervention, using the SDQ and a new screening algorithm. This approach has several appealing properties. First SDQ is relatively brief and can easily be distributed and quickly answered by parents. Secondly, the scoring of SDQ allows an algorithm that combines symptoms scores with the impairment score such that we can identify children both with a certain load of symptoms and who are affected in their daily lives. Third, SDQ can be used to assess a broad range

of MHP. Thus, we identify a wide variety in types of diagnosis among the children that screened positive and received a mental disorder diagnosis doing the follow-up. It is, however, possible that other instruments and/or more sophisticated screening algorithms could be as good or better at identifying children in the need of early intervention for MHP.

This study bases its results on a screening of children at the age of 11–12 years. SDQ data are also available for the CCC2000 from when the children were aged 5–7 years (pre-school age). The choice of using the data from when the children were aged 11–12 years is based on the availability of the IQ-proxy test score. In an additional analysis, we estimate the association between MHP at aged 5–7 years (using the same SDQ cut-off, we identify 5.6% of the children with MHP) and later school performance without controlling for IQ and health care utilisation. Table F and G in S1 Table, report the results for children with MHP at age 5–7 years and their later outcomes.

We do not find any association between MHP identified at pre-school age and the children's school performance in 9th grade. For health care utilisation, we find results similar to the ones we report for the children identified at age 11–12 years. More specifically, MHP at the pre-school age is associated with higher health care costs in the subsequent 11 years. The estimates are, however, statistically significantly lower than the ones we find when screening the older age group. Although far from conclusive, these findings suggest that the age of screening is an important factor, and that our definition of MHP might not be as predictive among a pre-school population. Future studies should investigate the age factor further, and studies focusing on early treatment should consider this factor in the identification of children with MHP.

Conclusions

This study provides evidence that supports the potential for systematically early identification of MHP among children. Future studies should conduct a more comprehensive psychopathological assessment of the children identified with the presented SDQ algorithm, to give a further understanding of the full potential of systematically early identification of MHP and subsequently early treatment.

Supporting information

S1 Table. Contains seven additional tables with results (Table A-G).
(DOCX)

Acknowledgments

We thank Anja Munkholm, Anne Mette Skovgaard, Charlotte Ulrikka Rask, Else Marie Olsen, Hanne Elberling, Lars Clemmensen, and Martin K. Rimvall from the co-author group "The CCC2000 Study Group".

Author Contributions

Conceptualization: Rasmus Trap Wolf, Pia Jeppesen, Dorte Gyrd-Hansen, Anne Sophie Oxholm.

Data curation: Rasmus Trap Wolf, Anne Sophie Oxholm.

Formal analysis: Rasmus Trap Wolf.

Funding acquisition: Rasmus Trap Wolf, Pia Jeppesen.

Investigation: Rasmus Trap Wolf, Pia Jeppesen.

Methodology: Rasmus Trap Wolf, Pia Jeppesen.

Project administration: Rasmus Trap Wolf.

Supervision: Pia Jeppesen, Dorte Gyrd-Hansen, Anne Sophie Oxholm.

Validation: Rasmus Trap Wolf, Pia Jeppesen.

Writing – original draft: Rasmus Trap Wolf.

Writing – review & editing: Rasmus Trap Wolf, Pia Jeppesen, Dorte Gyrd-Hansen, Anne Sophie Oxholm.

References

1. Polanczyk GV, Salum GA, Sugaya LS, Caye A, Rohde LA. Annual Research Review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. J Child Psychol Psychiatry. 2015; 56: 345–365. https://doi.org/10.1111/jcpp.12381 PMID: 25649325
2. Sawyer MG, Hazell PL. Health-Related Quality of Life of Children and Adolescents With Mental Disorders. J AM ACAD CHILD ADOLESC PSYCHIATRY. 2002; 8.
3. Costello EJ, Copeland W, Angold A. Trends in psychopathology across the adolescent years: What changes when children become adolescents, and when adolescents become adults?: Trends in psychopathology across the adolescent years. J Child Psychol Psychiatry. 2011; 52: 1015–1025. https://doi.org/10.1111/j.1469-7610.2011.02446.x PMID: 21815892
4. Copeland WE, Adair CE, Smetanin P, Stiff D, Briante C, Colman I, et al. Diagnostic transitions from childhood to adolescence to early adulthood. J Child Psychol Psychiatry. 2013; 54: 791–799. https://doi.org/10.1111/jcpp.12062 PMID: 23451804
5. Reef J, van Meurs I, Verhulst FC. Children's Problems Predict Adults' DSM-IV Disorders Across 24 Years. Adolesc PSYCHIATRY. 2010; 49: 8.
6. Hill RM, Pettit JW, Lewinsohn PM, Seeley JR, Klein DN. Escalation to Major Depressive Disorder among Adolescents with Subthreshold Depressive Symptoms: Evidence of Distinct Subgroups at Risk. J Affect Disord. 2014; 158: 133–138. https://doi.org/10.1016/j.jad.2014.02.011 PMID: 24655777
7. Shankman SA, Lewinsohn PM, Klein DN, Small JW, Seeley JR, Altman SE. Subthreshold conditions as precursors for full syndrome disorders: A 15-year longitudinal study of multiple diagnostic classes. J Child Psychol Psychiatry. 2009; 50: 1485–1494. https://doi.org/10.1111/j.1469-7610.2009.02117.x PMID: 19573034
8. Patton GC, Coffey C, Romaniuk H, Mackinnon A, Carlin JB, Degenhardt L, et al. The prognosis of common mental disorders in adolescents: a 14-year prospective cohort study. The Lancet. 2014; 383: 1404–1411. https://doi.org/10.1016/S0140-6736(13)62116-9
9. Copeland WE, Wolke D, Shanahan L, Costello EJ. Adult Functional Outcomes of Common Childhood Psychiatric Problems: A Prospective, Longitudinal Study. JAMA Psychiatry. 2015; 72: 892. https://doi.org/10.1001/jamapsychiatry.2015.0730 PMID: 26176785
10. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime Prevalence and Age-of-Onset Distributions of DSM-IV Disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry. 2005; 62: 593. https://doi.org/10.1001/archpsyc.62.6.593 PMID: 15939837
11. Arango C, Díaz-Caneja CM, McGorry PD, Rapoport J, Sommer IE, Vorstman JA, et al. Preventive strategies for mental health. Lancet Psychiatry. 2018; 5: 591–604. https://doi.org/10.1016/S2215-0366(18)30057-9 PMID: 29773478
12. Weisz JR, Kuppens S, Ng MY, Eckshtain D, Ugueto AM, Vaughn-Coaxum R, et al. What five decades of research tells us about the effects of youth psychological therapy: A multilevel meta-analysis and implications for science and practice. Am Psychol. 2017; 72: 79–117. https://doi.org/10.1037/a0040360 PMID: 28221063
13. Cross S, Hickie I. Transdiagnostic stepped care in mental health. Public Health Res Pract. 2017; 27. https://doi.org/10.17061/phrp2721712 PMID: 28474049
14. Levitt JM, Saka N, Hunter Romanelli L, Hoagwood K. Early identification of mental health problems in schools: The status of instrumentation. J Sch Psychol. 2007; 45: 163–191. https://doi.org/10.1016/j.jsp.2006.11.005
15. Anderson JK, Ford T, Soneson E, Coon JT, Humphrey A, Rogers M, et al. A systematic review of effectiveness and cost-effectiveness of school-based identification of children and young people at risk of, or

currently experiencing mental health difficulties. Psychol Med. 2019; 49: 9–19. https://doi.org/10.1017/S0033291718002490 PMID: 30208985

16. Kataoka SH, Zhang L, Wells KB. Unmet Need for Mental Health Care Among U.S. Children: Variation by Ethnicity and Insurance Status. Am J Psychiatry. 2002; 159: 1548–1555. https://doi.org/10.1176/appi.ajp.159.9.1548 PMID: 12202276

17. Sturm R, Ringel JS, Andreyeva T. Geographic Disparities in Children's Mental Health Care. PEDIATRICS. 2003; 112: e308–e308. https://doi.org/10.1542/peds.112.4.e308 PMID: 14523217

18. Roberts RE, Fisher PW, Blake Turner J, Tang M. Estimating the burden of psychiatric disorders in adolescence: the impact of subthreshold disorders. Soc Psychiatry Psychiatr Epidemiol. 2015; 50: 397–406. https://doi.org/10.1007/s00127-014-0972-3 PMID: 25358512

19. Balázs J, Miklósi M, Keresztény Á, Hoven CW, Carli V, Wasserman C, et al. Adolescent subthreshold-depression and anxiety: psychopathology, functional impairment and increased suicide risk: Adolescent subthreshold-depression and anxiety. J Child Psychol Psychiatry. 2013; 54: 670–677. https://doi.org/10.1111/jcpp.12016 PMID: 23330982

20. Gustafsson J-E, Westling AM, Åkerman A, Eriksson B, Eriksson C, Fischbein L, et al. School, learning and mental health: a systematic review. Stockholm: The Health Committee, Royal Swedish Academy of Sciences; 2010.

21. Kristoffersen JHG, Obel C, Smith N. Gender differences in behavioral problems and school outcomes. J Econ Behav Organ. 2015; 115: 75–93. https://doi.org/10.1016/j.jebo.2014.10.006

22. Keilow M, Sievertsen HH, Niclasen J, Obel C. The Strengths and Difficulties Questionnaire and standardized academic tests: Reliability across respondent type and age. Goertz M, editor. PLOS ONE. 2019; 14: e0220193. https://doi.org/10.1371/journal.pone.0220193 PMID: 31344079

23. French MT, Homer JF, Robins PK. What You Do in High School Matters: East Econ J. 2015; 41: 370–386.

24. Gevrek D, Gevrek ZE, Guven C. Benefits of Education at the Intensive Margin: Childhood Academic Performance and Adult Outcomes among American Immigrants. East Econ J. 2015; 41: 298–328. https://doi.org/10.1057/eej.2015.6

25. Oreopoulos P, Salvanes KG. Priceless: The Nonpecuniary Benefits of Schooling. J Econ Perspect. 2011; 25: 159–184. https://doi.org/10.1257/jep.25.1.159

26. Skovgaard AM, Olsen EM, Houmann T, Christiansen E, Samberg V, Lichtenberg A, et al. The Copenhagen County child cohort: Design of a longitudinal study of child mental health. Scand J Public Health. 2005; 33: 197–202. https://doi.org/10.1080/14034940510005662 PMID: 16040460

27. Goodman A, Goodman R. Strengths and Difficulties Questionnaire as a Dimensional Measure of Child Mental Health. J Am Acad Child Adolesc Psychiatry. 2009; 48: 400–403. https://doi.org/10.1097/CHI.0b013e3181985068 PMID: 19242383

28. Goodman R. The Extended Version of the Strengths and Difficulties Questionnaire as a Guide to Child Psychiatric Caseness and Consequent Burden. J Child Psychol Psychiatry. 1999; 40: 791. PMID: 10433412

29. Goodman R, Meltzer H, Bailey V. The strengths and difficulties questionnaire: A pilot study on the validity of the self-report version. Adolesc Psychiatry. 1998; 7: 7.

30. Theunissen MHC, de Wolff MS, Reijneveld SA. The Strengths and Difficulties Questionnaire Self-Report: A Valid Instrument for the Identification of Emotional and Behavioral Problems. Acad Pediatr. 2019; 19: 471–476. https://doi.org/10.1016/j.acap.2018.12.008 PMID: 30639760

31. Nielsen LG, Rimvall MK, Clemmensen L, Munkholm A, Elberling H, Olsen EM, et al. The predictive validity of the Strengths and Difficulties Questionnaire in preschool age to identify mental disorders in preadolescence. Hashimoto K, editor. PLOS ONE. 2019; 14: e0217707. https://doi.org/10.1371/journal.pone.0217707 PMID: 31158249

32. Ortuño-Sierra J, Aritio-Solana R, Fonseca-Pedrero E. Mental health difficulties in children and adolescents: The study of the SDQ in the Spanish National Health Survey 2011–2012. Psychiatry Res. 2018; 259: 236–242. https://doi.org/10.1016/j.psychres.2017.10.025 PMID: 29091822

33. Goodman R. The Strengths and Difficulties Questionnaire: A Research Note. J Child Psychol Psychiatry. 1997; 38: 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x PMID: 9255702

34. Stringaris A, Goodman R. The Value of Measuring Impact Alongside Symptoms in Children and Adolescents: A Longitudinal Assessment in a Community Sample. J Abnorm Child Psychol. 2013; 41: 1109–1120. https://doi.org/10.1007/s10802-013-9744-x PMID: 23677767

35. Meltzer H, Gatward R, Goodman R, Ford T. The Mental Health of Children and Adolescents in Great Britain: (622732007–001) [Internet]. American Psychological Association; 2000.

36. Woerner W, Becker A, Rothenberger A. Normative data and scale properties of the German parent SDQ. Eur Child Adolesc Psychiatry. 2004; 13. https://doi.org/10.1007/s00787-004-2002-6 PMID: 15243780

37. Arnfred J, Svendsen K, Rask C, Jeppesen P, Fensbo L, Houmann T, et al. Danish norms for the Strengths and Difficulties Questionnaire. Dan Med J. 2019; 66.

38. Jensen VM, Rasmussen AW. Danish education registers. Scand J Public Health. 2011; 39: 91–94. https://doi.org/10.1177/1403494810394715 PMID: 21775362

39. Ministry of Education. 7-point grading scale [Internet]. 2017 [cited 12 Nov 2018]. http://eng.uvm.dk/general-overview/7-point-grading-scale

40. Danmarks Statistik. Forbrugerprisindeks [Internet]. 2018 [cited 3 Jul 2018]. https://dst.dk/da/Statistik/emner/priser-og-forbrug/forbrugerpriser/forbrugerprisindeks

41. Danske Regioner. Pris- og lønudvikling 1988–2018 fordelt på løn og priser [Internet]. 2018 [cited 3 Jul 2018]. http://regioner.dk/aftaler-og-oekonomi/oekonomisk-vejledning/oekonomisk-vejledning-2018

42. Sahl Andersen J, De Fine Olivarius N, Krasnik A. The Danish National Health Service Register. Scand J Public Health. 2011; 39: 34–37. https://doi.org/10.1177/1403494810394718 PMID: 21775348

43. Wallach Kildemoes H, Toft Sørensen H, Hallas J. The Danish National Prescription Registry. Scand J Public Health. 2011; 39: 38–41. https://doi.org/10.1177/1403494810394717 PMID: 21775349

44. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. Scand J Public Health. 2011; 39: 30–33. https://doi.org/10.1177/1403494811401482 PMID: 21775347

45. Wechsler David. WISC-IV : Technical and Interpretive Manual 4th edition. San Antonio, TX: The Psychological Corporation; 2003.

46. Deb P, Norton EC, Manning WG Jr.. Health Econometrics Using Stata. Stata Press; 2017.

Paper I Supplementary material

S1 Table

Table A: The association between mental health problems at age 11-12 years and not completing the 9th grade exams

	Crude model N=2,015	Model including child's characteristics N=2,015	Model including child's characteristics and IQ-proxy N=1,537	Model including child's and parent's characteristics N=2,015	Model including child's, parent's characteristics, and IQ-proxy N=1,537
Mental health problems	3.22 (0.91) ***	3.14 (0.91) ***	1.71 (0.68)	3.00 (0.87)***	1.57 (0.65)
Girl		0.53 (9.13) **	0.53 (0.16)**	0.54 (0.14)**	0.51 (0.15)**
First-born in the family		1.04 (0.25)	1.09 (0.31)	1.01 (0.25)	1.04 (0.31)
Born with low-birth-weight or small-for-gestational-age		0.39 (0.29)	0.27 (0.27)	0.34 (0.25)	0.24 (0.24)
Parents are first- or second-generation immigrants		1.84 (0.59)*	1.56 (0.62)	1.60 (0.60)	1.34 (0.62)
Child did not live with both biological parents 1. Jan 2012		1.23 (0.33)	1.31 (0.42)	1.05 (0.60)	1.11 (0.37)
Mother had contact to the mental health hospital services between 1995 and 2011				1.74 (0.63)	2.28 (1.01)*
Father had contact to the mental health hospital services between 1995 and 2011				0.69 (0.36)	0.91 (0.53)
Mother's highest education primary school (up to 9 years)				0.51 (0.20)*	0.60 (0.27)
Father's highest education primary				1.66 (0.50)*	1.05 (0.41)

school (up to 9 years)					
Mother was unemployed in 2011				1.08 (0.50)	1.72 (0.87)
Father was unemployed in 2011				2.02 (0.86)*	1.95 (0.98)
Mother among the 25% of the CCC2000 mothers with highest health care costs between 2000 and 2011				1.37 (0.38)	0.99 (0.36)
Father among the 25% of the CCC2000 fathers with highest health care costs between 2000 and 2011				1.59 (0.42)*	1.38 (0.44)
Block Design Score			0.91 (0.03)***		0.91 (0.03)**

Note: The table contains estimated Odds Ratios (OR) of having a mental health problem at age 11-12 years and not completing the 9th grade exams. Logistic regression models are used to estimate OR with standard error (SE). As not all children have an IQ-test score the model including this population has fewer observations. We define not having completed the 9th grade exams as not attending at least four of the eight mandatory exams.
*p-level <0.1 **p-level <0.05 ***p-level <0.01.

Table B: The association between mental health problems at age 11-12 years and Grade Point Average at 9th grade exams

	Crude model N=1,941	Model including child's characteristics N=1,941	Model including child's characteristics. and IQ-proxy N=1,486	Model including child's and parent's characteristics N=1,941	Model including child's, parent's characteristics, and IQ-proxy N=1,486
	ME (SE)				
Mental health problems	-1.27 (0.19)***	-1.13 (0.18)***	-1.03 (0.19)***	-1.04 (0.18)***	-0.98 (0.19)***
Girl		0.80 (0.10)***	0.67 (0.10)***	0.75 (0.10)***	0.63 (0.10)***
First-born in the family		0.09 (0.10)	0.23 (0.10)**	0.11 (0.10)	0.23 (0.10)**
Born with low-birth-weight or small-for-gestational-age		-0.61 (0.20)***	-0.30 (0.22)	-0.48 (0.20)**	-0.24 (0.23)
Parents are first- or second-generation immigrants		-2.20 (0.17)***	-1.59 (0.19)***	-1.19 (0.19)***	-0.84 (0.21)***
Child did not live with both biological parents 1. Jan 2012		-0.86 (0.11)***	-0.64 (0.13)***	-0.58 (0.12)***	-0.41 (0.13)***
Mother had contact to the mental health hospital services between 1995 and 2011				0.22 (0.19)	0.24 (0.22)
Father had contact to the mental health hospital services between 1995 and 2011				-0.27 (0.23)	-0.30 (0.24)
Mother's highest education primary school (up to 9 years)				-0.88 (0.16)***	-0.69 (0.19)***
Father's highest education primary school (up to 9 years)				-0.91 (0.15)***	-0.71 (0.16)***

Mother was unemployed in 2011				-1.01 (0.22)***	-0.76 (0.26)***
Father was unemployed in 2011				-0.42 (0.24)*	-0.40 (0.27)
Mother among the 25% of the CCC2000 mothers with highest health care costs between 2000 and 2011				0.11 (0.13)	0.12 (0.14)
Father among the 25% of the CCC2000 fathers with highest health care costs between 2000 and 2011				-0.17(0.12)	-0.27 (0.12)**
Block Design Score			0.22 (0.01)***		0.20 (0.01)***

Note: Note: The table contains estimated coefficients (in grade points) for the association between mental health problems at age 11-12 years and Grade Point Average (GPA) at 9th grade exams. The study population includes children who have done at least four of the 9th grade exams. As not all children have an IQ-test score the model including this population has fewer observations. Ordinary Least Square regression models are used to estimate Marginal effect (ME) with robust standard error (SE) of mental health problems on GPA. *p-level <0.1 **p-level <0.05 ***p-level <0.01.

Table C: The associations between mental health problems at age 11-12 years and total health care costs in the period 2012 to 2016

	Crude model	Model including child's characteristics	Model including child's and parent's characteristics
	ME (SE)		
Mental health problems	6,044 (2,254)***	7,336 (2,301)***	5,858 (1,713)***
Girl		2,436 (1,122)**	2,232 (1,142)*
First-born in the family		1,711 (1,111)	1,419 (1,142)
Born with low-birth-weight or small-for-gestational-age		-2,121 (817)***	-1.946 (765)**
Parents are first- or second-generation immigrants		-1,845 (751)**	-1,810 (763)**
Child did not live with both biological parents 1. Jan 2012		1,340 (1,582)	1,259 (1,782)
Mother had contact to the mental health hospital services between 1995 and 2011			1,248 (1,566)
Father had contact to the mental health hospital services between 1995 and 2011			313 (1,127)
Mother's highest education primary school (up to 9 years)			-441 (987)
Father's highest education primary school (up to 9 years)			2,469 (1,328)*
Mother was unemployed in 2011			-1.852 (723)**
Father was unemployed in 2011			-2,158 (721)***
Mother among the 25% of the CCC2000 mothers with highest health care costs between 2000 and 2011			1,676 (766)**
Father among the 25% of the CCC2000 fathers with highest health care costs between 2000 and 2011			1,441 (1000)

Note: The table contains estimated coefficients (in €) of the association between mental health problems at age 11-12 years and total health care costs in the following five year-period. The number of observations N=2,015. Marginal effect (ME) and robust standard error (SE) estimated with a GLM model.
*p-level <0.1 **p-level <0.05 ***p-level <0.01.

Table D: The associations between mental health problems at age 11-12 years with/without specialised mental health care and health care cost in the period 2012 to 2016

	Crude models		Model including child's characteristics[c]		Model including child's and parent's characteristics[c,d]	
	Mental health problems and no specialised mental health care (n=150)	Mental health problems and specialised mental health care (n=44)	Mental health problems and no specialised mental health care (n=150)	Mental health problems and specialised mental health care (n=44)	Mental health problems and no specialised mental health care (n=150)	Mental health problems and specialised mental health care (n=44)
	ME (SE)	ME (SE)	ME (SE)	ME (SE)	ME (SE)	ME (SE)
Total health care costs[a]	-918 (930)	29,778 (8,325)***	-90 (960)	32,568 (8,706)***	-71 (996)	27,781 (6,832)***
Primary sector care without psychologist/psychiatrist[a]	126 (84)	365 (139)***	118 (78)	373 (160)**	99 (78)	322 (158)**
Psychologist/psychiatrist in primary care sector[b]	NA	1,004 (435)**	NA	769 (262)***	NA	1,023 (522)*
Prescription medicine[b]	-18 (62)	775 (244)***	-20 (60)	714 (250)***	-18 (51)	879 (330)***
Mental outpatient hospital care[b]	NA	3,853 (727)***	NA	3,943 (769)***	NA	3,853 (822)***
Mental inpatient hospital care[b]	NA	22,043 (7,917)***	NA	22,592 (7,613)***	NA	57,833 (67,081)
Somatic outpatient hospital care[b]	157 (292)	929 (508)*	151 (275)	913 (502)*	37 (240)	685 (425)
Somatic inpatient hospital care[b]	-166 (673)	809 (828)	545 (818)	1,210 (948)	723 (856)	1,570 (1,079)

Note: The table contains estimated coefficients (in €) of the association between mental health problems at age 11-12 years and health care costs in the following five year-period. The number of observations N=2,015. Some estimates are non-applicable (NA) based on the definition of the group. Specialised mental health care (SMHC) is defined as contact with publicly funded psychiatrist or psychologist.

[a] Marginal effect (ME) and robust standard error (SE) estimated with a GLM model.

[b] Marginal effect (ME) and robust standard error (SE) estimated with a two-part model combining a logit and a GLM model.

[c] Gender, first-born child in the family, born small-for-gestational-age or with low-birth-weight, parents are second generation immigrants, child did not live with both biological parents 1st January 2012.
[d] Mother and/or father had contact to mental health hospital services between 1995 and 2011, mother's and/or father's highest education on 1st January 2012 was primary school (up to 9 years schooling), mother and/or father was unemployed in 2011, mother and/or father was among the parents of the full CCC2000 cohort with the 25% highest health care cost between 2000 and 2011.
*p-level <0.1 **p-level <0.05 ***p-level <0.01.

Table E: The full Copenhagen Child Cohort (CCC2000) population and the population of children with Strengths and Difficulties questionnaire (SDQ) data at age 11-12 years

	The full CCC2000 population	Children with SDQ data at age 11-12 years
Girl	48.7%	50.3% *
First-born in the family	51.8%	50.4%
Born with low-birth-weight or small-for-gestational-age	6.5%	6.1%
Parents are first- or second-generation immigrants	16.4%	10.4% ***
Child did not live with both biological parents 1. Jan 2012	21.5%	14.5% ***
Mother had contact to the mental health hospital services between 1995 and 2011	7.8%	6.7% **
Father had contact to the mental health hospital services between 1995 and 2011	4.2%	3.4% **
Mother's highest education primary school (up to 9 years)	23.2%	14.1% ***
Father's highest education primary school (up to 9 years)	25.2%	17.5% ***
Mother was unemployed in 2011	17.2%	9.6% ***
Father was unemployed in 2011	7.6%	4.4% ***
Mother among the 25 % with highest health care sector costs between 2000 and 2011 in the study population	25.0%	23.2% **
Father among the 25 % with highest health care sector costs between 2000 and 2011 in the study population	25.0%	24.7%
N	6,090	2,126

Note: The table describes the characteristics of children and their parents in the Copenhagen Child Cohort by whether they are part of the study population or not. Differences in the characteristics across the children's mental problem status is tested using a Chi-square test. *p-level <0.1 **p-level <0.05 ***p-level <0.01.

Table F: The association between mental health problems at age 5-7 and Grade Point Average at 9th grade exams

	Mental health problems at age 5-7 years	
	ME	SE
Crude model	-0.57***	0.21
Model including child's characteristics[a]	-0.25	0.21
Model including child's and parent's characteristics[a,b]	-0.17	0.19

Note: The table contains estimated coefficients (in grade points) for the association between mental health problems at age 11-12 years and Grade Point Average (GPA) at 9th grade exams. N=3,105. The study population includes children who have done at least four of the 9th grade exams. Ordinary Least Square regression models are used to estimate Marginal effect (ME) with robust standard error (SE) of mental health problems on GPA

[a] Gender, first-born child in the family, born small-for-gestational-age or with low-birth-weight, parents are second generation immigrants, child did not live with both biological parents 1st January 2006.

[b] Mother and/or father had contact to mental health hospital services between 1995 and 2005, mother's and/or father's highest education on 1st January 2006 was primary school (up to 9 years schooling), mother and/or father was unemployed in 2005, mother and/or father was among the parents of the full CCC2000 cohort with the 25% highest health care cost between 2000 and 2005.

*p-level <0.1 **p-level <0.05 ***p-level <0.01.

Table G: The associations between mental health problems at age 5-7 and health care costs in the period 2006 to 2016

	Crude models		Models including child's characteristics[c]		Models including child's and parent's characteristics[c,d]	
	Mental health problems at age 5-7 years					
	ME	SE	ME	SE	ME	SE
Total health care costs[a]	5,155 **	2,120	4,885 **	1,958	4,338 **	1,691
Primary sector care without psychologist/psychiatrist[a]	615 *	329	619*	334	467 **	191
Psychologist/psychiatrist in primary care sector[a]	277 **	156	257 **	111	339 **	155
Prescription medicine[b]	2,183 ***	543	2,035 ***	530	2,289 ***	618
Mental outpatient hospital care[b]	1,977 ***	398	1,934 ***	406	1,787 ***	387
Mental inpatient hospital care[b]	323	1,415	825	1,683	596	1,326
Somatic outpatient hospital care[b]	269	376	264	368	156	287
Somatic inpatient hospital care[b]	-489	606	-302	571	-547	505

Note: The table contains estimated coefficients (in €) of the association between mental health problems at age 5-7 years and health care costs in the following eleven year-period (2006 to 2016). The number of observations N=3,306.

[a] Marginal effect (ME) and robust standard error (SE) estimated with a GLM model.

[b] Marginal effect (ME) and robust standard error (SE) estimated with a two-part model combining a logit and a GLM model.

[c] Gender, first-born child in the family, born small-for-gestational-age or with low-birth-weight, parents are second generation immigrants, child did not live with both biological parents 1st January 2006.

[d] Mother and/or father had contact to mental health hospital services between 1995 and 2005, mother's and/or father's highest education on 1st January 2006 was primary school (up to 9 years schooling), mother and/or father was unemployed in 2005, mother and/or father was among the parents of the full CCC2000 cohort with the 25% highest health care cost between 2000 and 2005.

*p-level <0.1 **p-level <0.05 ***p-level <0.01.

Paper II

Systematic identification and stratification of help-seeking school-aged youth with mental health problems - A novel approach to stage-based stepped-care.

Submitted for publication.

Systematic identification and stratification of help-seeking school-aged youth with mental health problems - A novel approach to stage-based stepped-care

Rasmus Trap Wolf[1,3]*, Louise Berg Puggaard[1], Mette Maria Agner Pedersen[1], Anne Katrine Pagsberg[1,2], Wendy K. Silverman[12], Christoph U. Correll[13,14,15,16], Kerstin Jessica Plessen[1,4], Simon-Peter Neumer[10,11], Dorte Gyrd-Hansen[3], Mikael Thastum[8], Niels Bilenberg[5,6], Per Hove Thomsen[7,8], Pia Jeppesen[1,2]

[1]Child and Adolescent Mental Health Centre, Mental Health Services – Capital Region of Denmark, Copenhagen, Denmark.

[2]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark.

[3]Danish Centre for Health Economics, Department of Public Health, University of Southern Denmark, Denmark.

[4]Division of Child and Adolescent Psychiatry, Department of Psychiatry, Lausanne University Hospital CHUV, Switzerland.

[5]Department for Child and Adolescent Psychiatry, Mental Health Services in the Region of Southern Denmark.

[6]University of Southern Denmark, Odense, Denmark.

[7]Research Center at Department for Child- and Adolescent Psychiatry, Aarhus University Hospital, Denmark.

[8]Institute of Clinical Medicine, Faculty of Health, Aarhus University, Aarhus, Denmark.

[9]Centre for the Psychological Treatment of Children and Adolescents, Department of Psychology and Behavioural Sciences, Aarhus BSS, Aarhus University, Denmark.

[10]Centre for Child and Adolescent Mental Health, Oslo, Norway.

[11]Centre for Child and Youth Mental Health and Child Welfare, The Arctic University of Norway, North Norway (RKBU North), Tromsø, Norway.

[12]Anxiety and Mood Disorders Program, Yale Child Study Center, Yale University, School of Medicine, New Haven, CT, USA.

[13]Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Department of Psychiatry and Molecular Medicine, Hempstead, New York, USA.

[14]The Zucker Hillside Hospital, Department of Psychiatry, Glen Oaks, New York, USA.

[15]Feinstein Institute for Medical Research, Center for Psychiatric Neuroscience, Manhasset, New York, USA.

[16]Charité Universitätsmedizin, Department of Child and Adolescent Psychiatry, Berlin, Germany.

Abstract

We investigated whether a novel visitation model for school-aged youth with mental health problems based on a stage-based stepped-care approach facilitated a systematic identification and stratification process without problems with equity in access.

The visitation model was developed within the context of evaluating a new transdiagnostic early treatment for youth with anxiety, depressive symptoms, and/or behavioural problems. The model aimed to identify youth with mental health problems requiring an intervention, and to stratify the youth into three groups with increasing severity of problems. This was accomplished using a two-phase stratification process involving a web-based assessment and a semi-structured psychopathological interview of the youth and parents. To assess problems with inequity in access, individual-level socioeconomic data were obtained from national registers with data on both the youth participating in the visitation and the background population.

Altogether, 573 youth and their parents took part in the visitation process. Seventy-five (13%) youth had mental health problems below the intervention threshold, 396 (69%) were deemed eligible for the early treatment, and 52 (9%) had symptoms of severe mental health problems. Fifty (9%) youth were excluded for other reasons. Eighty percent of the 396 youth eligible for early treatment fulfilled criteria of a mental disorder. The severity of mental health problems highlights the urgent need for a systematic approach. Potential problems in reaching youth of less resourceful parents, and older youth were identified. These findings can help ensure that actions are taken to avoid equity problems in future mental health care implementations.

1. Introduction

A large proportion of children and adolescents worldwide suffer from mental disorders [1]. A recent nationwide study reported a cumulative treatment-incidence of 15% for any mental disorder before the age of 18 years in Denmark [2]. Child and Adolescent Mental Health Services (CAMHS) in Denmark and other countries worldwide are faced with a huge demand for diagnostic assessment procedures and effective treatments [3–5]. The urgent need for safe and effective interventions targeting child and adolescent (youth) emotional and behavioural problems is accentuated by the negative individual and societal consequences associated with early mental health problems. These include putting the youth on a trajectory for chronic mental disorders with the consequence of impaired daily functioning, educational and social difficulties, and absence from the job market in adulthood [6]. Such negative consequences are also present for subthreshold conditions, which are linked to many of the same negative long-term outcomes [6], including a higher risk of suicide [7]. With most mental disorders presenting first in childhood and adolescence [8], prevention and timely delivery of care are solutions to lowering the rates of adult mental health disorders and to meet the needs of the youth suffering from disorders and subthreshold conditions.

Several national health bodies thus recommend a stepped-care approach in child and adolescent mental health [9–11]. The stepped-care approach aims to offer interventions in increasing intensity levels by offering low intensity interventions as the first step, and subsequently interventions with higher intensity. The goal is to minimize intervention costs and time burden for youth and their parents are kept as minimal as possible, while still offering evidence-based interventions [12]. Stepped-care approaches have been criticized for potentially delaying appropriate treatment by having individuals go initially through insufficient low intensity interventions [9]. The approach has also been criticized for not fitting 'real-world' settings in that youth do not usually present with well-defined single disorders or conditions that are easily recognized and matched to a particular step of care [9, 13]. Thus, it is often unclear when and to whom specific interventions should be given in a stepped-care approach [14], as the diagnostic thresholds can be too arbitrary to define treatment and treatment intensity indication [9]. As an answer to this criticism, a stage-based stepped-care approach has been discussed and developed [9, 15]. Staging, as known from general medicine, is based on placing the problem or condition on a continuum of stages that are defined by the available interventions, clarifying the match of problems and interventions [16]. Furthermore, the stage-based stepped-care approach allows for the complexity of children's and adolescents' mental health problems by not necessarily focusing on specific diagnoses [15].

Despite the appeal of the stage-based stepped-care approach, it is still not widely implemented for school-aged children and adolescents. For such an approach to work in clinical practice, there is a need for a coherent process from early identification and systematic stratification to interventions of all intensities. In many countries, mental health care is divided so that prevention and general counselling is located outside the health care system in schools and community services under local governments, while diagnostic assessment and treatments are carried out in hospital or outpatient clinics [17]. With increasing rates of families seeking help for youth mental health problems and disorders, this dichotomy creates a gap in care for youth who need more help than community counselling, but less than specialized CAMHS treatment. For example, 27% (N=1,566) of all referrals to CAMHS in the Capital Region of Denmark were rejected in 2018 (personal communication). This may cause a vicious cycle in which youth with milder conditions must either wait for their conditions to deteriorate before they can get appropriate care, or they never receive care despite the

risk of negative long-term consequences [6, 7]. When applied to internal medicine, this current approach corresponds to not treating a person with high blood pressure before a cardiac infarct has manifested.

Evidence-based programs for anxiety, milder depression, and behavioural problems and disorders in children and adolescents exist [18–21], but are not systematically offered. This lack of implementation may be partly due to the high costs and logistic difficulties in training and implementation of several diagnosis specific programs. These are problems that can potentially be addressed by the development of effective transdiagnostic programs [22]. The Mind My Mind (MMM) intervention is a transdiagnostic and modular cognitive and behavioural intervention for school-aged youth who have clinically significant levels of emotional and/or behavioural problems that impact their daily lives and threaten to disrupt their development, but who do not qualify for treatment in CAMHS. The intervention aims at providing early treatment of anxiety, depressive symptoms and/or behavioural problems. The MMM trial is a pragmatic, multi-site, randomized, parallel-group, controlled trial of the MMM intervention versus management as usual (MAU).

For the purpose of identifying youth in need of an intervention for emotional and/or behavioural problems and to prepare for dissemination of evidence-based treatment, a standardized visitation model was developed and implemented in four municipalities across Denmark. The model aimed to stratify the help-seeking youth into three groups with increasing severity of mental health problems based on a stage-based stepped-care principle. Hence, the groups with increasing severity of mental health problems were defined based on the possible actions: Stage 1) self-help and general counselling in the municipalities, Stage 2) early treatment for mental health problems, in this cases, in the form of participation in the MMM trial, and Stage 3) referral to CAMHS or other health care professionals for more comprehensive assessment and treatment.

In this study, we had two primary aims: 1) To describe the help-seeking population and the visitation model based on psychopathological questionnaires, and 2) To analyse the representativeness of the population in the visitation compared to the background population in the four municipalities using unique individual-level data on socio economic factors from national registers. Hence, we investigated whether the visitation model, when operating in a naturalistic setting, facilitated a systematic identification and stratification process with equal access to care. We further examined the feasibility of the visitation process by analysing the duration and mean changes in psychopathology for the youth evaluated a second time in relation to the baseline of the MMM trial.

2. Method

2.1 Setting

The MMM trial was conducted in four Danish municipalities located in three of the five regions in Denmark. The procedures that led to inclusion in the trial took place from 7th September 2017 till 18th December 2018. The details of the intervention are described in the online study record (NCT03535805).

The trial was implemented in the context of Educational and Psychological Counselling Services (PPR) in the municipalities. PPR psychologists conducted the visitation and intervention whilst supervised by a senior consultant in child and adolescent psychiatry (author PJ).

2.2 Assessment of psychopathology

The *Strengths and Difficulties Questionnaire* (SDQ) is a widely used and well-validated tool to identify and assess children with mental health problems in both clinical samples and in general population-based samples [23–25]. SDQ is available in validated versions for parents and for youth ≥11 years old. The questionnaire contains 25 items, which cover five subscales relating to the children's Emotional problems, Peer problems, Behavioural problems, Hyperactivity and Pro-social behaviour. Responses to the first four subscales are summarized to calculate a total difficulties score. Each subscale score ranges from 0-10, so that the Total difficulties score ranges from 0-40 [26]. In this study, the extended version of SDQ was used. This version includes an impact assessment to evaluate how much and for how long the identified mental difficulties interfere with the child's everyday life. A functional impairment score is calculated based on five items on whether the difficulties upsets or distresses the child and how much the difficulties interfere with home life, friendships, classroom learning, and leisure activities. Each item is scored on a scale from 0-2. To score 1 or 2, the interference from the difficulties in that domain must be assessed to either "quite a lot" or "a great deal" [23, 27]. Hence, the Impact score ranges from 0-10.

All parents responded to the SDQ in the beginning of the visitation process. A second response to the SDQ was obtained as a baseline measure for the children who at the end of visitation process were enrolled into the MMM trial. Youth aged ≥11 years old also answered the SDQ themselves. As the enrolled school-aged population was aged 5-16 years, the focus in this study was on parental-reported SDQ. Child reported SDQ results are available in Table S2 Supplemental Material.

In addition to the second SDQ measurement, a more thorough psychopathological assessment was performed for the youth who were enrolled in the MMM trial using *The Developmental And Well-Being Assessment* (DAWBA). DAWBA is an online questionnaire and rating techniques designed to generate *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV/5*) psychiatric diagnosis in children and adolescents aged 5-17 years, covering the common emotional, behavioural and hyperactivity disorders [28, 29]. When scoring positive on screening questions, the interview opens for additional questions covering all the operationalized diagnostic criteria. Also, open-ended questions were administered to capture the respondent's own description of the problems. All information was reviewed by experienced child and adolescent psychiatrists trained in DAWBA to decide on DSM-IV/5 diagnoses. When in doubt, consensus was reached among two or more raters. The Fleiss kappa coefficient for inter-rater reliability of the main groups of disorders was 0.65. The diagnoses were solely used to describe the psychopathology of the included youth for research purposes.

The *Mood and Feelings Questionnaire* (MFQ) [30] and *Spence Children's Anxiety Scale* (SCAS) [31] were used in the visitation process to investigate the mental health problems. Both scales are specific psychopathological questionnaires focusing on depression and anxiety symptoms, respectively, unlike the SDQ, which has a general focus. In this study, we screened for different mental problems of the children, hence, we focused on the SDQ responses. The MFQ and SCAS responses are reported in Table S3 Supplemental Material together with Danish norms [31, 32].

2.3 Age, gender, and socioeconomic status

Information on age, sex, and socioeconomic factors were derived on an individual level from national registers. Socioeconomic factors included immigration status (being first- or second-generation immigrant), mother's and father's highest achieved education, number of children in the household, type of household, and household income. Individual level data of the background population was also used in this study. The background population was defined as all youth aged 6-16 years old in the four municipalities on January 1., 2018. All register data used in this study were based on data from the year 2018.

2.4 The visitation model

Referral to the visitation process was based on self-referral by help-seeking parents to ensure an easy and fast access to help. Information about the MMM trial and requirements for inclusion were published online in school intranets and municipality websites and given to local educational and health care professionals in the four municipalities including teachers, PPR psychologists, and general practitioners. No formal referral was required. The help-seeking parents were guided by professionals at the child's school or in the municipality to contact the local PPR via telephone to sign up the child for assessment. An administrative assistant asked the parents to provide contact information (telephone and e-mail of child and both parents/guardians) and to return a form with written informed consent to take part in the web-based, digital data collection.

The visitation model was implemented during the period of the MMM trial in the four municipalities as part of standard care. The aim of the visitation was to identify the target population for an early treatment, as offered with the MMM intervention, during the trial period and thereafter. This was achieved through a systematic stratification of the youth with help-seeking parents. The youth were stratified into three groups: Stage 1, Stage 2, and Stage 3. The stratification process is referred to as the visitation model.

The visitation model consisted of two phases.

2.4.1 Phase 1: Web-based initial assessments

First, the parents received a text message with link to answer the SDQ online. The SDQ responses and a screening cut-off algorithm were used to identify the youth deemed having mental health problems not severe enough to be offered an early treatment (Stage 1 group). The cut-off was based on age-matched populations in Denmark, Germany and the UK [27, 33, 34], so that scoring above the cut-off implied that the youth's mental health state is ranked among the 10% of the most affected in these populations [35]. A score above the cut-off required SDQ specific scores of ≥1 in the SDQ Impact score combined with at least one of the following scores: Total difficulties score of ≥14 and/or an Emotional problem score of ≥5 and/or a Behavioural problem score of at least 3. This cut-off was evaluated in a recent study as effective in identifying youth with mental health problems and a high risk of poor school outcomes years later [35]. Youth ≥11 years old also answered the SDQ initially; these answers were however not used in the cut-off process.

If the parents' responses to SDQ scored below the cut-off, they were automatically informed that the youth's current mental health problems were not of sufficient severity that an intervention was required, and standard low intensity offers like self-help and general counselling in the municipalities was considered. If the parents reported SDQ scores above the cut-off, they and their youth were automatically presented with

MFQ and SCAS online. After having responded to SDQ, MFQ, and SCAS, the parents and the youth were invited to a meeting with a local PPR psychologist.

2.4.2 Phase 2: Meeting with the psychologist and psychopathological interview

The PPR psychologist reviewed the responses from the parents and youth to the SDQ, MFQ and SCAS before the meeting. At the meeting, the psychologist conducted a structured interview concerning the youth's development, family and social situation, school attendance, learning problems, symptoms and functioning in daily life. The meeting also included a brief, semi-structured psychopathological interview with both youth and parent(s), inspired by the KSADS-PL [36] and ADIS-IV [37] but modified to screen for all the main groups of mental health disorders in childhood and adolescent. Based on the interview, an agreement was made between the psychologist, youth and parent(s) about the formulation of the youth's "top problem" which they would like to address.

If the assessment revealed signs of severe developmental or severe mental disorder, including intellectual functional impairment, autism spectrum disorder (ASD), attention deficit/hyperactivity disorder (ADHD), psychotic disorder, eating disorder, obsessive-compulsive disorder, repeated self-harm, abuse or dependence of alcohol or psychoactive drugs, the youth was excluded from the trial (Stage 3 group). If the remaining youth did not meet any of the other exclusion criteria (see Table S1 Supplemental Material) including previous mental disorder diagnosis, informed consent of participations in the trial was obtained and the youth were included in the trial and a baseline assessment including SDQ was conducted (Stage 2 group).

2.5 Data analysis

2.5.1 Stratification

This article focuses on the youth stratified to the three stage groups as defined earlier. The youth excluded due to other exclusion criteria (e.g. unable to participate in weekly session (see Table S1 Supplemental Material for full list of exclusions)) were not included in the analyses.

The parental responses to SDQ and the derived scores were used to characterize the help-seeking population and to analyse whether the visitation process succeeded in stratifying the youth into three categories with increasing severity of mental health problems. Danish norms from a population survey were used as reference values [34]. The youth-reported SDQ, as well as parental- and youth-reported MFQ and SCAS were analysed the same way and the results are available in Table S2 and S3 Supplemental Material.

The duration of the presented mental health problems was quantified as part of the SDQ with a question about the duration of the difficulties, where the parents could indicate if the difficulties have been present "less than a month", "1-5 months", "6-12 months", or "Over a year".

The youth that were included in the MMM trial (stage 2 group) were assessed with DAWBA. The diagnoses derived from this assessment are presented in the results section to explore the severity of mental health problems in relation to the diagnostic thresholds in this group of community treated youth.

2.5.2 Feasibility

To examine the feasibility of the visitation model, we analysed the duration of the visitation procedures and the mean changes in SDQ scores.

The time from the help-seeking parents responded to the questionnaires in the phase 1 online module until they had the phase 2 interview with the PPR psychologist analysed. The predetermined aim was a median time of two weeks for the process. This aim was decided based on a combination of being ambitious about providing fast assessment and being realistic about minimum of one parent being able to attend the interview during normal working hours.

For the youth included in the MMM trial, we analysed the differences in SDQ sub- and sum-scores between the beginning and end of the visitation process. This process was only possible for this specific group, as their parents answered SDQ again at the baseline of the trial after inclusion.

2.5.3 Representativeness

We compared characteristics between the three groups in visitation and the background population in the four municipalities to investigate possible inequity problems with the approach of the visitation model. The background population was defined as all youth aged 6-16 years old in the four municipalities on first of January 2018. Age, sex, and socioeconomic factors were compared. We expected a positive selection from the beginning of the visitation process towards youth from higher socioeconomic status groups, due to the need of parents to take the initiative of getting involved and answer questionnaires in the visitation process. Differences between the youth participating in the visitation and the background population are not necessarily due to the visitation model itself as it could also simply be due to the differences in prevalence of mental health problems among age-groups, sex, and socioeconomic groups. Thus, the findings will be interpreted based on the known variations in prevalence of mental health problems in two Danish cohorts [2, 35, 38]. If there was no inequity in access, we would expect to find that the population taking part in the visitation would have at least a similar proportion of youth with immigration background as the background population, that a larger proportion of the mothers in the visitation population would have a lower education, that a higher proportion of the youth in the visitation population would come from a household with a single parent, and that a lower proportion in the visitation population would come from a high income household. We expected to find a similar sex proportion in the visitation population as in the background population due to small sex differences in the rate of mental health problems, while we expected to observe a higher proportion of older youth in the visitation population due to increasing incidence rates of mental health problems with age.

2.5.4 Statistical analysis

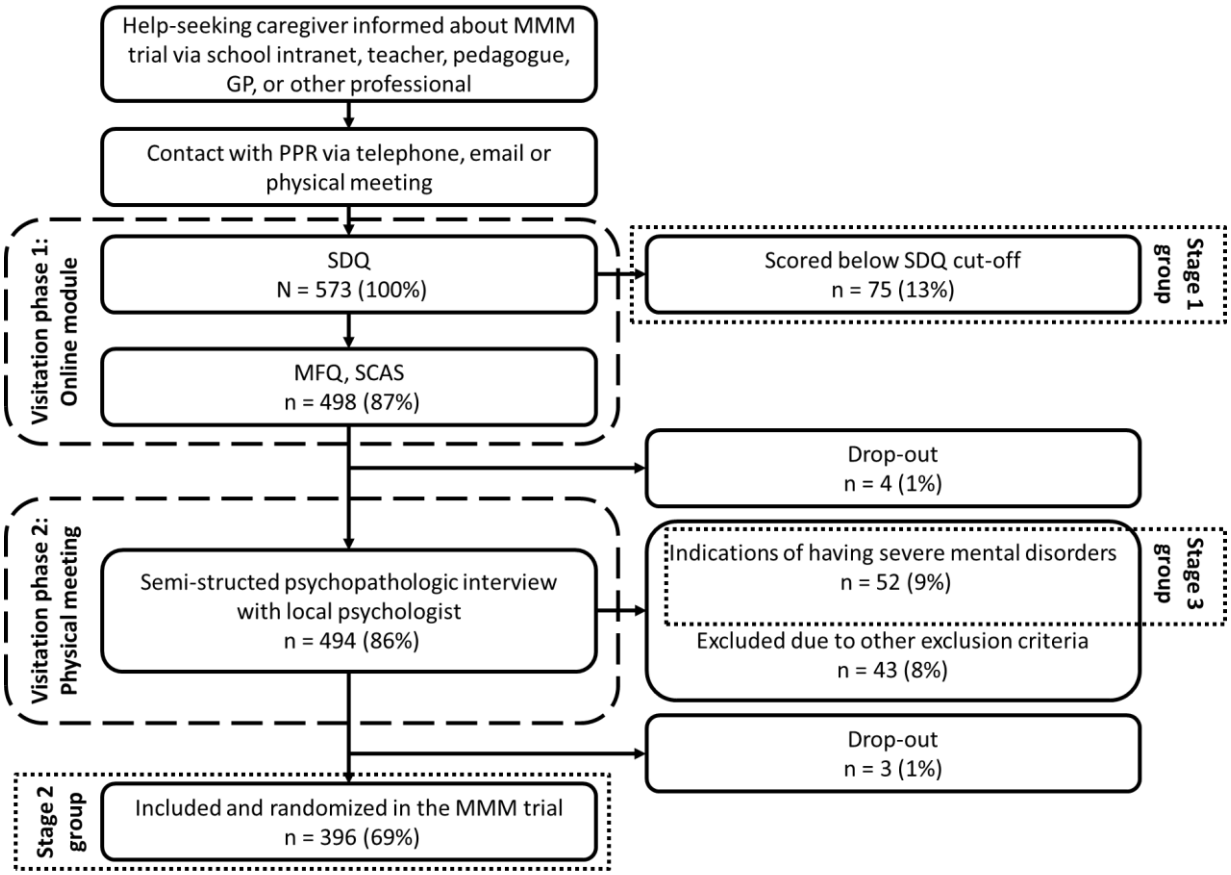
Pearson's chi-squared test and Student's t-tests were used to compare groups for categorical and continuous variables respectively. Shapiro-Francia test was used to test for normality. The sub-scores of the SDQ (not Total difficulties and Impact score) failed the normally test. As a robustness test, Mann-Whitney U test was utilized in the comparison of these scores. Similarly, when paired t-test was used, Wilcoxon matched-pairs signed-ranks test was used for robustness check. This was not possible when comparing with Danish norms, as these data were only available in aggregated numbers. All robustness checks did, however, find the results of the Student's t-tests robust. P-levels <0.05 are referred to as statistically significant throughout. To ensure

data confidentiality, Statistics Denmark does not allow output for cells containing data from <4 individuals, and we therefore do not report exact results in such cases, but merely report that the number of individuals is <4. All analyses were performed in STATA-15 [39].

3. Results

Altogether, 573 youth/their parents contacted their local PPR, gave consent and contact information, and subsequently entered the online module of phase 1 in the visitation process and responded to the SDQ. The flow of the visitation process and the delineation of the stratification groups are illustrated in Figure 1. Seventy-five (13%) of the youth scored below the cut-off. These youth constitute the stage 1 group in the stratification. In phase 2 of the visitation process, 52 (9%) youth were excluded due to indications of severe mental disorders, and these youth constitute the stage 3 group. Furthermore, 43 youth were excluded based on other exclusion criteria. Seven children dropped out for unknown reasons. Finally, 396 (69%) youth were included in the MMM trial, and these youth constitute the stage 2 group in the stratification.

Figure 1: Flowchart of visitation process



Note: PPR: Educational and psychological counselling. SDQ: The Strength and Difficulties Questionnaire. MFQ: The Mood and Feelings Questionnaire. SCAS: Spence Children's Anxiety Scale. MMM: Mind My Mind.

The youth in Stage group 1, 2, and 3 had a mean age of 10.2 years (median 10.0) years, and 47% were girls. For sex, age, immigration status, household type, and household income, we found no statistically significant

differences between the three groups (Table 1). For the mother's highest achieved education, we found a statistically significant difference between group 2 and 3. A higher proportion of youth in group 3 had mothers with a lower secondary education. Despite not being statistically significantly different, the same tendency was found for group 1 versus 3. For the father's highest achieved education, we found a statistically significant difference between group 1 and 3. This is due to a large proportion of fathers in group 1 having an upper secondary education. The findings do not seem to indicate any systematic bias in the selection of the groups, as they were not consistent between variables.

Table 1: Socio economic characteristics of the three staging groups

	Stage 1 group **n = 75 (74[a])**	**Stage 2 group** **n = 396**	**Stage 3 group** **n = 52**
Sex, female, n (%)	34 (45)	190 (48)	21 (40)
Age, years, mean (SD)	10.1 (0.30)	10.2 (0.12)	10.2 (0.40)
1st or 2nd generation immigrant, n (%)	NA	5 (1)	NA
Mother's highest education, n (%)		[+]	
Lower secondary	5 (7)	27 (7)	10 (19)
Upper secondary	30 (41)	152 (41)	18 (35)
Short cycle tertiary, Bachelor or equivalent	33 (45)	186 (48)	20 (38)
Master or equivalent	6 (8)	25 (6)	4 (8)
Father's highest education, n (%)	[*]		
Lower secondary	8 (13)	53 (15)	9 (19)
Upper secondary	41 (64)	178 (49)	19 (40)
Short cycle tertiary, Bachelor or equivalent	10 (15)	91 (25)	18 (38)
Master or equivalent	5 (8)	38 (11)	NA
No. of children in household, n (%)	[$]		
1	24 (32)	67 (17)	6 (12)
2	30 (41)	235 (59)	30 (58)
3+	20 (27)	94 (24)	16 (31)
Household type, n (%)			
Single woman	15 (20)	90 (23)	15 (29)
Couple	47 (64)	276 (70)	32 (62)
Other constellations (single male or more than one family)	12 (16)	30 (8)	5 (10)
Household income before tax, n (%)			
0-500,000 DKK	24 (32)	112 (28)	21 (40)
>500,000-1,000,000 DKK	40 (54)	235 (59)	23 (44)
>1,000,000 DKK	10 (14)	49 (12)	8 (15)

Note: [a]One child is not identifiable in the national register, hence only sex and age were available. NA: Not applicable due to the combination of low number and data confidentiality. [*]Group 1 is statistically significant different from group 3 on a p-level <0.05. [+]Group 2 is statistically significant different from group 3 on a p-level <0.05. [$]Group 1 is statistically significant different from group 2 and 3 on a p-level <0.05.

3.2 The stratification

When we analysed the parents' responses to the SDQ, we found that all three groups had a statistically significant worse mean Total difficulties score and Impact score compared to the Danish norms. For the sub-scores, this was also the case for Emotional problems, Behavioural problems, and Hyperactivity, while only group 2 and 3 had statistically significant worse scores than the Danish norm in the sub-scores Peer problems and Pro-social behaviour. All mean scores, except Emotional problems, followed the same pattern of worse mean scores from group 1 to group 2, and from group 2 to group 3. For Emotional problems, group 2 had a non-significant higher mean score compared to group 3. Group 1 had statistically significant better mean scores in all scoring categories compared to group 2 and 3. Compared to group 3, group 2 had significant better mean scores in Hyperactivity, Peer-problems, Total difficulties and Impact score (Table 2). The mean difference in impact score from group 1 to the group 2 was 3.1 (95% CI: 2.7-3.6), and 1.0 (95% CI: 0.5-1.6) from group 2 to group 3.

Table 2: Psychopathology of the youth at visitation

	Stage 1 group n = 75	**Stage 2 group n = 396**	**Stage 3 group n = 52**	**Danish norms[1] N = 3,146**
The Strength and Difficulties Questionnaire, mean (SD)				
Emotional problems	3.5 (2.0)[*]	7.0 (2.4)	6.4 (2.4)	2.3 (2.3)[$]
Behavioural problems	1.4 (1.3)[*]	2.8 (2.0)	3.3 (1.8)	0.9 (1.3)[$]
Hyperactivity	3.1 (2.1)[*]	5.0 (2.8)[+]	6.3 (2.3)	2.5 (2.5)[$]
Peer problems	1.6 (1.6)[*]	2.8 (2.1)[+]	3.7 (2.6)	1.3 (1.7)[£]
Pro-social behaviour	8.4 (1.9)[*]	7.6 (2.1)	7.0 (2.3)	8.7 (1.5)[£]
Total difficulties score	9.6 (4.1)[*]	17.5 (5.1)[+]	19.7 (5.1)	7.1 (5.8) [$]
Impact score	1.1 (1.6)[*]	4.2 (1.9)[+]	5.3 (2.2)	0.6 (1.6) [$]

Note: [1]Danish norms of children aged 6-17 from Arnfred et al. [34]. [*]Group 1 is statistically significantly different from group 2 and 3 on a p-level <0.05 tested with both t-test and Mann-Whitney U test. [+]Group 2 is statistically significantly different from group 3 on a p-level <0.05 tested with both t-test and Mann-Whitney U test. [$]Danish norms are statistically significantly different from group 1, 2, and 3 on a p-level <0.05 tested with t-test. [£]Danish norms are statistically significantly different from group 2 and 3 on a p-level <0.05 tested with t-test.

For the majority of the youth (82%), the mental health problems had lasted more than a year, while only 39 youth (7%) had experienced problems for less than five months. No statistically significant differences were found between the three groups, but there was a tendency to a shorter duration of problems in group 1.

At the DAWBA assessment of the 396 youth who were included in the MMM trial, 317 (80%) fulfilled the criteria of at least one DSM-IV/5 diagnosis and 26% fulfilled diagnostic criteria from more than one diagnostic group. More than half of the 396 youth (56%) suffered from one or more anxiety disorder, 97 (24%) suffered from a behavioural disorder, and 58 (15%) from a depressive disorder. The group of youth with neurodevelopmental disorders comprised individuals who fulfilled the criteria for ADHD (n=46), ASD (n=3),

and/or tics disorders (n=11). In total, 57 (14%) of the 396 youth had at least one of these neurodevelopmental disorders (Table 3).

3.3 Feasibility

For the 494 children who continued to phase 2 of the visitation process, the mean time in visitation from when the online module of phase 1 was finalized to the interview with the PPR psychologist constituting phase 2 was 18.7 (SD: 19.7) days. The median time was 15 days, and 95% had the interview within 36 days.

The 396 youth who were included in the MMM trial had a mean time between the two SDQ assessments of 30.9 days (SD: 20.7) (Table 3). For all, but Prosocial and Impact score, mean scores were significantly improved at the baseline responses compared to the time of visitation. The mean Impact score did not change between the two points in time, but the variance in the score increased. At baseline, 8% had the minimum Impact score of 0 and 5% had a score of 9 or 10, compared to 0% and 3%, respectively, at the time of visitation.

Table 3: Psychopathology of the youth included in the Mind My Mind

The **Strength and Difficulties Questionnaire, mean (SD)**			
	Visitation score	Baseline score	P-value[1]
Emotional problems	7.0 (2.4)	6.4 (2.5)	<0.001
Behavioural problems	2.8 (2.0)	2.6 (2.0)	<0.001
Hyperactivity	5.0 (2.8)	4.7 (3.0)	0.049
Peer problems	2.8 (2.1)	2.5 (2.1)	<0.001
Pro-social behaviour	7.6 (2.1)	7.6 (2.0)	0.566
Total difficulties score	17.5 (5.1)	16.3 (5.5)	<0.001
Impact score	4.2 (1.9)	4.2 (2.4)	0.550
Days between parent-reporting at visitation and baseline, mean (SD)			30.9 (20.7)
DSM-IV/5 Mental disorders based on the Development and Well Being Assessment (DAWBA), n (%)			
Anxiety disorder			220 (56)
Depressive disorder			58 (15)
Behaviour disorder			97 (24)
Neurodevelopmental disorder			57 (14)
Any disorder			317 (80)
Comorbidity, ≥ 2 disorders			102 (26)

Note: N = 396. [1]Paired t-test between visitation and baseline score. As a robustness test Wilcoxon matched-pairs signed-ranks tests were performed for all subscales which resulted in similar statistical significance levels.

3.4 Representativeness

Comparing the distribution of sex, age, and socioeconomic variables between the total population of the three groups in the visitation to the background population in the four municipalities, we found statistically significant differences for all variables except for sex and father's education level. For household number of children, type, and income, we found minor differences, and in the direction expected from the cohort studies [35, 38]. Age was significantly lower in the visitation population relative to the background population, due to a lower proportion of adolescents aged 14-16 years in the visitation population. This group represents 10% compared to 29% in the background population. The proportion of 1st and 2nd generation immigrants across the two populations differed significantly with few in the visitation population having an immigrant background. Moreover, the background population had a higher proportion of mothers with lower secondary as the highest education level (Table 4).

Table 4: Socioeconomic characteristics of the visitation population and the background population

	Total of group 1-3 **n = 523 (522[a])**	**Background population[1]** **n = 32,814**
Sex, female n (%)	245 (47)	15,989 (49)
Age, mean (SD)	10.2 (0.11)	11.2 (3.12)*
1st or 2nd generation immigrant, n (%)	NA	3,198 (10)*
Mother's highest education, n (%)		*
Lower secondary	42 (8)	4,900 (15)
Upper secondary	200 (39)	12,368 (39)
Short cycle tertiary, Bachelor or equivalent	239 (46)	11,678 (37)
Master or equivalent	35(7)	2,939 (9)
Father's highest education, n (%)		
Lower secondary	70 (15)	5,953 (19)
Upper secondary	238 (50)	15,022 (48)
Short cycle tertiary, Bachelor or equivalent	119 (25)	7,015 (23)
Master or equivalent	NA	3,029 (10)
Children in household, n (%)		*
1	97 (19)	5,152 (16)
2	295 (57)	16,030 (49)
3+	130 (25)	11,632 (35)
Household type based on adults, n (%)		*
Single woman	120 (23)	5,637 (17)
Couple	355 (68)	22,964 (70)
Other constellations	47 (9)	4,213 (13)
Household income before tax, n (%)		*
0-500,000 DKK	157 (30)	9,379 (29)
>500,000-1,000,000 DKK	298 (57)	16,186 (49)
>1,000,000 DKK	67 (13)	7,249 (22)

Note: [a]One child is not identifiable in the national register, hence only sex and age were available. NA: not applicable due to the combination of low number and data confidentiality. [*]The total population in the visitation is statistically significant different from the background population on a p-level <0.05. [1]All children aged 6-16 in the four participating municipalities.

4. Discussion

Using a stage-based stepped-care approach the visitation model succeeded in defining three groups of youth with increasing severity of mental health problems based on parental reported SDQ. The MFQ and SCAS showed the same tendencies in the results using both parent- and self-reported scores. The anxiety specific scores (SDQ emotional problems and SCAS scores) were not statistically different in group 2 and 3. This was, however, acceptable, as the MMM intervention was designed to handle high levels of anxiety if there is no profound comorbidity. We found that all three groups were more affected by mental health problems than the average Danish child. The majority of the youth who took part in the visitation were heavily affected by their mental health problems, with a total of 10.5% assessed as having signs of severe mental health disorders, and with further 61% of the youth fulfilling the criteria of a mental disorder diagnosis. These findings falls in line with findings of previous studies, in which young people seeking help in community-based services often have more severe mental health problems than first-line services are designed to address [40]. The present study highlights the urgent need for systematic identification of mental health issues in youth, so that interventions can be delivered to help this group of youth, who do not receive adequate or timely care today. Findings by Copeland et al. [6] and Wolf et al. [35] underline that without an intervention, these youth are likely to suffer from long-term adverse consequences.

The investigation of potential inequity in access with the visitation model highlighted three major concerns. Comparing the yotuh in the visitation model with the background population from the four municipalities, we found indications of a selection that disfavoured youth with immigration background as well as children of mothers with lower level of education. The difference in proportion of youth with immigration background cannot be explained by the exclusion criteria of no parent speaking Danish as only one child was excluded due to this (see Table S1 Supplemental Material), or by the prevalence, as 1st and 2nd generation immigrants have at least the same prevalence of mental health problems [35]. The visitation model seemed to be biased towards youth having parents with more resources, which could be due to the self-referral approach. Another issue identified in the comparison between the population that took part in the visitation and the background population was that the proportion of adolescents aged 14-16 years who took part in the visitation was low compared to the background population despite our expectation of a higher proportion due to increasing incidence rates of mental health problems during adolescence [38]. This finding points towards a limitation of the visitation model in recruiting youth with mental health problems in their early teenage years. This limitation could be due to two factors. First, internalizing mental health problems are more common in the older age group [38] and this type of problems might not be noticed by parents and/or teachers as often as other types of problems, especially externalizing ones. Furthermore, the visitation model was advertised with the MMM intervention. This material could perhaps have had a form that appealed more to parents and younger children, making the older age group less likely to take action themselves and professionals less likely to inform this group. Other models aiming specifically at adolescents typically take efforts to make themselves directly accessible for this age group [40]. In future implementation of the visitation model, careful information, nudging and other helpful initiatives to enrol the youth of mothers with

lower education, youth with immigration background, and youth in the age group 14-16 years into the visitation should be considered to avoid inequity in mental health care and to optimize the visitation model. The approach with open self-referral by help-seeking parents should, however, still be maintained as it clearly succeeded based on the findings of all three groups being more affected by mental health problems than the average Danish youth, and the majority were in need of an early intervention or further psychiatric assessment.

The presented visitation model had an acceptable duration from the online module to the meeting and assessment with the local psychologist. The median of 15 days does, however, show that there is room for improvement as half the youth experienced longer time than the pre-stated aim of two weeks. However, considering that more than nine out of ten youth had experienced their mental health problems for more than five months, the duration of the visitation process seems acceptable. The SDQ score change from visitation to baseline showed that the youth on average seemed to improve slightly during the visitation process.

Another important issue is the latency period from identification and assessment of problems to access to treatment [5]. An early systematic assessment and staging of the youth, as proposed with this model, can contribute to shorten wait times by ensuring that the correct target group reaches the different interventions in a timely way. With no systematic approach previously in place, it is likely, however, that more youth in need of an intervention will be identified initially, as many youth in need currently are not being identified and offered early treatment [35]. This means that the full effect of implementing the visitation model may likely only be achieved over time.

A limitation of the presented visitation model is that it only focuses on a segment of the full framework of mental health care [16]. Important areas, such as mental health literacy and universal prevention, are not addressed. Also, the actions and procedures in CAMHS are beyond the scope of this visitation model. An important aspect that was not addressed here is the collaboration across mental health care services within and beyond the health care system. A fast handover of youth with signs and symptoms of severe mental disorders is crucial in order, to assure effectiveness of the early identification and visitation model. Also, standard procedures for the youth who score below the SDQ cut-off should be further formalized to ensure that all help-seeking parents receive adequate support. This should include information on how to act if the mental health problem intensifies. Another limitation of the visitation model is that it does not directly address all types of mental health problems. If youth with these problems are identified in the face-to-face interview in the visitation process, the involved psychologist can, however, ensure referral to the appropriate professionals. A methodological limitation was that the diagnostic assessment with DAWBA was only carried out for stage 2-group, and not the others.

A strength of the visitation model is that its focus is not limited to a single type of disorder. The model allows for the link with for example, the transdiagnostic MMM intervention that focuses on the most prevalent childhood mental health problems: anxiety, depression, and behavioural problems. This feature makes the visitation model more feasible and sustainable for wide scale implementation. Even in smaller geographical areas, there will often be a sufficient number of youth suffering from these kind of mental health problems to render the systematic approach effective. Compared to having identification and stratification models for each specific area, the transdiagnostic approach is also likely to incur lower implementation and running costs.

In summary, this study provides evidence that the visitation model based on a stage-based stepped-care approach with systematic identification and stratification of help-seeking school-aged children and adolescents with mental health problems is possible to carry out successfully in a community setting. It appears realistic to reduce the observed inequity in access by strengthening the outreach to youth with immigration background, youth of mothers with lower level of education and youth aged 14-16 years before full, wide-scale implementation.

5. Declarations

5.1 Funding

The study was supported by unrestricted grants from TrygFonden and the Lundbeck Foundation. The funding agencies were not involved in the study design, data collection, analysis and interpretation, or in the writing of the manuscript and decision of submission for publication.

5.2 Conflicts of interest/Competing interests

RTW, LBP, MMAP, AKP, WKS, KJP, SPN, DGH, MT, NB, PHT and PJ declare that they have no conflict of interest. CUC has been a consultant and/or advisor to or has received honoraria from: Acadia, Alkermes, Allergan, Angelini, Axsome, Gedeon Richter, Gerson Lehrman Group, IntraCellular Therapies, Janssen/J&J, LB Pharma, Lundbeck, MedAvante-ProPhase, Medscape, Neurocrine, Noven, Otsuka, Pfizer, Recordati, Rovi, Sumitomo Dainippon, Sunovion, Supernus, Takeda, and Teva. He has provided expert testimony for Janssen and Otsuka. He served on a Data Safety Monitoring Board for Lundbeck, Rovi, Supernus, and Teva. He received royalties from UpToDate and grant support from Janssen and Takeda. He is also a stock option holder of LB Pharma.

5.3 Ethics approval

Ethical approval of the Mind My Mind study was provided by the Regional Committee on Health Research Ethics in the Capital Region of Denmark, Journal-nr.: H-17011408.

5.4 Consent to participate

Participation was voluntary, and data was kept confidential. The parents gave written consent to the study participation. According to Danish legislation, individual informed consent or additional permission from the National Committee on Health Research Ethics is not required for registry-based studies. The parents gave written consent to the study participation.

5.5 Consent for publication

The written consent to the study participation included consent for publication.

5.6 Availability of data and material

The data utilized in the current study are defined as sensitive personal data, and cannot be shared publicly due to existing data protection laws in Denmark, and imposed by the Danish Data Protection Agency. The collected data is uploaded and linked to data from the Danish National registers and analysed through Statistics Denmark (https://www.dst.dk/en).

5.7 Code availability

Not applicable.

References

1. Polanczyk GV, Salum GA, Sugaya LS, Caye A, Rohde LA (2015) Annual Research Review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. Journal of Child Psychology and Psychiatry 56:345–365

2. Dalsgaard S, Thorsteinsson E, Trabjerg BB, et al (2019) Incidence Rates and Cumulative Incidences of the Full Spectrum of Diagnosed Mental Disorders in Childhood and Adolescence. JAMA Psychiatry. https://doi.org/10.1001/jamapsychiatry.2019.3523

3. Sundhedsstyrelsen (2017) Prævalens, incidens og aktivitet i sundhedsvæsenet for børn og unge med angst eller depression, ADHD og spiseforstyrrelser. Sundhedsstyrelsen, København

4. NHS Benchmarking Network (2019) NHS Benchmarking network Work Programme Report 2018/2019.

5. O'Brien D, Harvey K, Howse J, Reardon T, Creswell C (2016) Barriers to managing child and adolescent mental health problems: a systematic review of primary care practitioners' perceptions. British Journal of General Practice 66:e693–e707

6. Copeland WE, Wolke D, Shanahan L, Costello EJ (2015) Adult Functional Outcomes of Common Childhood Psychiatric Problems: A Prospective, Longitudinal Study. JAMA Psychiatry 72:892

7. Balázs J, Miklósi M, Keresztény Á, et al (2013) Adolescent subthreshold-depression and anxiety: psychopathology, functional impairment and increased suicide risk: Adolescent subthreshold-depression and anxiety. Journal of Child Psychology and Psychiatry 54:670–677

8. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE (2005) Lifetime Prevalence and Age-of-Onset Distributions of DSM-IV Disorders in the National Comorbidity Survey Replication. Archives of General Psychiatry 62:593

9. Cross S, Hickie I (2017) Transdiagnostic stepped care in mental health. Public Health Research & Practice. https://doi.org/10.17061/phrp2721712

10. Sundhedsstyrelsen (2017) Forløbsprogram for børn og unge med angst og/eller depression.

11. National Institute for Health and Care Excellence (2011) Common mental health problems: identification and pathways to care - Clinical Guideline (CG123).

12. van Straten A, Hill J, Richards DA, Cuijpers P (2015) Stepped care treatment delivery for depression: a systematic review and meta-analysis. Psychological Medicine 45:231–246

13. McGorry P, van Os J (2013) Redeeming diagnosis in psychiatry: timing versus specificity. The Lancet 381:343–345

14. Silverman WK, Pettit JW, Lebowitz ER (2016) Stepping Toward Making Less More for Concerning Anxiety in Children and Adolescents. Clinical Psychology: Science and Practice 23:234–238

15. Cross SPM, Hermens DF, Scott EM, Ottavio A, McGorry PD, Hickie IB (2014) A Clinical Staging Model for Early Intervention Youth Mental Health Services. Psychiatric Services 65:939–943

16. McGorry PD, Hickie IB, Yung AR, Pantelis C, Jackson HJ (2006) Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. 8

17. Braddick F, Carral V, Jenkins R, Jané-Llopis E (2009) Child and Adolescent Mental Health in Europe: Infrastructures, Policy and Programmes. European Communities, Luxembourg

18. Arango C, Díaz-Caneja CM, McGorry PD, et al (2018) Preventive strategies for mental health. The Lancet Psychiatry 5:591–604

19. Weisz JR, Kuppens S, Ng MY, et al (2017) What five decades of research tells us about the effects of youth psychological therapy: A multilevel meta-analysis and implications for science and practice. American Psychologist 72:79–117

20. Kendall PC, Makover H, Swan A, Carper MM, Mercado R, Kagan E, Crawford E (2016) What Steps to Take? How to Approach Concerning Anxiety in Youth. Clinical Psychology: Science and Practice 23:211–229

21. Waddell C, Schwartz C, Andres C, Barican JL, Yung D (2018) Fifty years of preventing and treating childhood behaviour disorders: a systematic review to inform policy and practice. Evidence Based Mental Health 21:45–52

22. Weisz J, Bearman SK, Santucci LC, Jensen-Doss A (2017) Initial Test of a Principle-Guided Approach to Transdiagnostic Psychotherapy With Children and Adolescents. Journal of Clinical Child & Adolescent Psychology 46:44–58

23. Goodman R (1999) The Extended Version of the Strengths and Difficulties Questionnaire as a Guide to Child Psychiatric Caseness and Consequent Burden. J Child Psychol Psychiatry 40:791

24. Goodman A, Goodman R (2009) Strengths and Difficulties Questionnaire as a Dimensional Measure of Child Mental Health. Journal of the American Academy of Child & Adolescent Psychiatry 48:400–403

25. Nielsen LG, Rimvall MK, Clemmensen L, Munkholm A, Elberling H, Olsen EM, Rask CU, Skovgaard AM, Jeppesen P (2019) The predictive validity of the Strengths and Difficulties Questionnaire in preschool age to identify mental disorders in preadolescence. PLOS ONE 14:e0217707

26. Goodman R (1997) The Strengths and Difficulties Questionnaire: A Research Note. Journal of Child Psychology and Psychiatry 38:581–586

27. Meltzer H, Gatward R, Goodman R, Ford T (2000) The Mental Health of Children and Adolescents in Great Britain: (622732007-001). https://doi.org/10.1037/e622732007-001

28. Goodman A, Heiervang E, Collishaw S, Goodman R (2011) The 'DAWBA bands' as an ordered-categorical measure of child mental health: description and validation in British and Norwegian samples. Social Psychiatry and Psychiatric Epidemiology 46:521–532

29. Goodman R, Ford T, Richards H, Gatward R, Meltzer H (2000) The Development and Well-Being Assessment: Description and Initial Validation of an Integrated Assessment of Child and Adolescent Psychopathology. J Child Psychol Psychiatry 41:645–55

30. Daviss WB, Birmaher B, Melhem NA, Axelson DA, Michaels SM, Brent DA (2006) Criterion validity of the Mood and Feelings Questionnaire for depressive episodes in clinic and non-clinic subjects. Journal of Child Psychology and Psychiatry 47:927–934

31. Arendt K, Hougaard E, Thastum M (2014) Psychometric properties of the child and parent versions of Spence Children's Anxiety Scale in a Danish community and clinical sample. Journal of Anxiety Disorders 28:947–956

32. Eg J, Bilenberg N, Costello EJ, Wesselhoeft R (2018) Self- and parent-reported depressive symptoms rated by the mood and feelings questionnaire. Psychiatry Research 268:419–425

33. Woerner W, Becker A, Rothenberger A (2004) Normative data and scale properties of the German parent SDQ. European Child & Adolescent Psychiatry. https://doi.org/10.1007/s00787-004-2002-6

34. Arnfred J, Svendsen K, Rask C, Jeppesen P, Fensbo L, Houmann T, Obel C, Niclasen J, Bilenberg N (2019) Danish norms for the Strengths and Difficulties Questionnaire. Dan Med J 66:

35. Wolf RT, Jeppesen P, Gyrd-Hansen D, The CCC2000 Study Group, Oxholm AS (2019) Evaluation of a screening algorithm using the Strengths and Difficulties Questionnaire to identify children with mental health problems: A five-year register-based follow-up on school performance and healthcare use. PLOS ONE 14:e0223314

36. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, Williamson D, 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. Journal of the American Academy of Child & Adolescent Psychiatry 36:980–988

37. Silverman WK, Saavedra LM, Pina AA (2001) Test-Retest Reliability of Anxiety Symptoms and Diagnoses With the Anxiety Disorders Interview Schedule for DSM-IV: Child and Parent Versions. ournal of the American Academy of Child & Adolescent Psychiatry 40:937–944

38. Steinhausen H-C, Jakobsen H (2019) Incidence Rates of Treated Mental Disorders in Childhood and Adolescence in a Complete Nationwide Birth Cohort. The Journal of Clinical Psychiatry. https://doi.org/10.4088/JCP.17m12012

39. StataCorp (2017) Stata Statistical Software: Release 15. StataCorp LLC., College Station, TX

40. Settipani CA, Hawke LD, Cleverley K, Chaim G, Cheung A, Mehra K, Rice M, Szatmari P, Henderson J (2019) Key attributes of integrated community-based youth service hubs for mental health: a scoping review. International Journal of Mental Health Systems. https://doi.org/10.1186/s13033-019-0306-7

Paper II Supplementary material

S1: Exclusion criteria in phase 2

	n (%)
Not aged 6–16 years and in 0-9th grade	2 (2)
Indications that the child may have a severe mental disorder	52 (55)
Top problem not in the domains of anxiety, depressive symptoms or behavioral problems	51 (54)
No parents understand and speak Danish sufficiently to participate in the treatment	1 (1)
Indications of intellectual functional impairment, severe learning difficulties or other special needs	14 (15)
A prior diagnosis of any developmental or mental disorder	8 (8)
The child and/or parents are unable to participate in weekly sessions	8 (8)
Parents does not want their child to participate	30 (32)
Total	**95 (100)**

Note: One child can meet more than one exclusion criteria.

S2: The Strength and Difficulties Questionnaire (SDQ) self-reported

	Group 1 Below SDQ algorithm cut-off n = 9	Group 2 Included and randomized n = 178	Group 3 Indications of having severe mental disorders n = 20	Danish norms[1] n = 1583-1609
Psychopathology				
The Strength and Difficulties Questionnaire, mean (SD)				
Emotional problems	6.6 (2.9)	5.9 (2.6)	5.3 (2.7)	2.7 (2.3)[$]
Behavioural problems	1.8 (1.2)	2.4 (1.7)	2.9 (1.6)	1.4 (1.5)[£]
Hyperactivity	5.0 (3.3)	4.4 (2.7)[+]	5.8 (2.9)	3.5 (2.3)[£]
Peer problems	2.7 (1.7)	2.7 (2.0)	3.4 (2.0)	1.8 (1.7)[£]
Pro-social behaviour	7.8 (1.9)	7.7 (1.5)	7.3 (2.0)	7.6 (1.7)
Total difficulties score	16.0 (5.9)	15.3 (5.5)	17.3 (6.1)	9.4 (5.5)[$]
Impact score	1.6 (1.6)	2.4 (2.3)	2.7 (2.0)	0.4 (1.2)[$]

Note: [1]Danish norms of children aged 6-17 from Arnfred et al.[34]. *Group 1 is statistically significant different from group 2 and 3 on a p-level <0.05 tested with both t-test and Mann-Whitney U test. [+]Group 2 is statistically significant different from group 3 on a p-level <0.05 tested with both t-test and Mann-Whitney U test. [$]Danish norms are statistically significant different from group 1, 2, and 3 on a p-level <0.05 tested with t-test. [£]Danish norms are statistically significant different from group 2 and 3 on a p-level <0.05 tested with t-test.

S3: The Mood and Feelings Questionnaire (MFQ) and Spence Children's Anxiety Scale (SCAS) at visitation

	Group 2 Included and randomized	Group 3 Indications of having severe mental disorders	Danish norms[1]
MFQ[2]			
Parent reported, mean (SD)	21.9 (11.5)*	27.5 (13.2)	7.1 (7.8)$
n	395	52	703
Self-reported, mean (SD)	21.8 (12.2)*	29.7 (15.1)	13.2 (10.8)$
n	340	40	992
SCAS[3]			
Parent reported, mean (SD)	35.6 (16.5)	33.1 (17.9)	14.4 (10.3)$
n	396	51	537
Self-reported, mean (SD)	38.2 (17.6)	41.4 (22.2)	22.0 (14.3)$
n	339	40	972

Note: [1]Danish norms for MFQ is of children aged 9-17 from Eg et al.[32]. Danish norms for SCAS is of children aged 7-17 from Arendt et al. [31]. [2]MFQ parent reported has 34 items and a score range from 0 to 68. MFQ self-reported has 33 items and a score range from 0 to 66 [32]. [3]SCAS parent reported has 39 items and a score range from 0 to 117. SCAS self-reported has 44 items and a score range from 0 to 132 [31]. *Group 2 is statistically significant different from group 3 on a p-level <0.05 tested with both t-test and Mann-Whitney U test. $Danish norms are statistically significant different from group 2, and 3 on a p-level <0.05 tested with t-test.

Paper III

The longitudinal validity of the CHU9D – Evidence from a mental health trial.

Submitted for publication.

The longitudinal validity of CHU9D – Evidence from a mental health trial

Rasmus Trap Wolf, M.Sc.[1,2], Julie Ratcliffe, Ph.D.[3], Gang Chen, Ph.D.[4], Pia Jeppesen, M.D., Ph.D [2,5]

[1]Danish Centre for Health Economics, Department of Public Health, University of Southern Denmark
[2]Child and Adolescent Mental Health Centre, Mental Health Services - Capital Region of Denmark
[3]College of Nursing and Health Sciences, Flinders University, Australia
[4]Centre for Health Economics, Monash University, Australia
[5]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen

Acknowledgements: The private, noncommercial health science research project Mind My Mind was initiated by Psykiatrifonden. Birthe Wielandt Houe acted as project manager on behalf of Psykiatrifonden. The randomized controlled effectiveness trial was carried out through close collaboration between Psykiatrifonden, the principal investigator (Pia Jeppesen), the participating municipalities, and the regional child and adolescent mental health centers. Psykiatrifonden held the legal responsibility as the data controller for the Web-based data collection.

Abstract

Objectives: The Child Health Utility 9D (CHU9D) currently represents the only preference-based health-related quality-of-life instrument designed exclusively from its inception for application with children. The objective of this study was to examine the construct validity and responsiveness of CHU9D in a mental health setting using utility weights derived from an adult and adolescent population, respectively.

***Methods*:** The discriminant validity and convergent validity were examined using the mental health specific The Strengths and Difficulties Questionnaire (SDQ) and the generic KIDSCREEN-27. Responsiveness was assessed by examining the floor-ceiling effects, the magnitude of change over time, and the ability to differentiate between improvement and no improvement.

***Results*:** The study included 396 children with mental health problems. CHU9D showed good construct validity, with correlation coefficients ranging between 0.329-0.571 for SDQ Impact score and KIDSCREEN-27 Psychological Well-being. CHU9D was able to distinguish between groups of children with different levels of mental health problems ($p<0.001$). The absolute magnitudes of the group mean differences were larger using adolescent weights. No evidence of a floor/ceiling effect was found at the baseline. A standardized response mean of 0.634-0.654 was found for the children who experienced clinically significant improvements. CHU9D was able to discriminate between children who experienced positive and no health improvements ($p<0.001$).

***Conclusion*:** This study provides the first evidence on responsiveness for CHU9D. The findings demonstrate that CHU9D is an appropriate HRQOL measure for use in mental health trials. Furthermore, the results show that the preference weights generated from an adolescent population perform best in discriminating between groups.

1. Introduction

Evidence of the cost-effectiveness of new and emerging interventions is commonly used across health systems to assist policy makers in the allocation of scarce health care resources. In order to conduct cost-utility analysis, the most prevalent form of economic evaluation, a preference-based measure of health-related quality of life (HRQOL) is widely used to facilitate the calculation of quality adjusted life years (QALYs), a generic measure of effectiveness [1].

During the last decade there has been an increased focus on preference-based HRQOL instruments that aim to measure children's utility. The guidelines for estimating QALYs in youth populations are, however, still unclear [2]. A recent review, therefore, warrants further empirical evidence on the valuation of youth-specific preference-based measures [3]. A review identifies nine preference-based HRQOL instruments that have been used in paediatric populations [4]. Of these the Child Health Utility 9D (CHU9D) is the only one designed exclusively from its inception for application with young people [5]. The remaining instruments are not originally developed for children and adolescents, and others represent different forms of adoptions from measures designed for adults.

CHU9D has been demonstrated in several studies to have good content, face, and construct validity for young people in the age-group 7-17 years [6–11]. Furber & Segal [12] examined face validity, practicality, internal consistency, and convergent validity of CHU9D in a population of clients in South Australian child and adolescent mental health services. The authors concluded that their initial validation of CHU9D showed promising results, but that there was a need for further validation including a general need for validation of responsiveness, which their cross-sectional design was unable to capture. Responsiveness is critical for a preference-based HRQOL instrument, as its suitability for application in economic evaluation depends on its capacity to reliably detect changes in HRQOL due to the introduction of new interventions. To the best of our knowledge, the responsiveness of CHU9D has not yet been examined.

There exist two main scoring algorithms for deriving utilities from CHU9D, one based on an adult population [10] and one based on an adolescent population [13]. The choice of whose values to use in an economic evaluation can have important policy implications due to the potential impact this has on the QALY estimates and thus on the final incremental cost-effectiveness ratio. The choice could especially be important in interventions aiming to improve mental health, as Ratcliffe and colleagues found that adolescents placed more weight upon impairments in CHU9D dimensions related to mental health (sad, worried, annoyed) than adults [8].

The main objective of this study is to examine the construct validity and responsiveness of CHU9D in a mental health setting. This will be the first study to examine the validity of CHU9D in a longitudinal design, and thereby the first study to examine responsiveness of CHU9D. Furthermore, the examination of construct validity will add to the evidence from Furber & Segal [12], on the appropriateness of using CHU9D in a mental health setting by examining a larger population and having comparison with both a mental health specific measure (the Strengths and Difficulties Questionnaire, SDQ) and a generic HRQOL measure (KIDSCREEN-27). A second objective is to examine whether the utility weights derived from the adult population or the utility weights derived from the adolescent population demonstrate better validity and responsiveness in this context.

2. Method

2.1 Participants

Participants in this study were part of the Mind My Mind trial (Trial ID: NCT03535805). The details of the trial are described in the online study record [14]. Briefly the methods and study population are as follows.

The trial is designed to evaluate the effectiveness and cost-effectiveness of a new transdiagnostic modular cognitive and behavioural treatment versus treatment-as-usual for school-aged children with emotional and/or behavioural disturbances.

The trial was advertised for professionals and parents in the community by use of pamphlets and intra/internet, and the recruitment was based on parent's initiative to seek help in collaboration with professionals such as the schoolteacher, nurse or psychologist. The inclusion procedure included minimum scores derived from the SDQ as reported by the parents and a clinical interview with a psychologist. The minimum scores from the SDQ follows a screening algorithm designed to identify children with mental health problems in the need of an intervention, this is further described elsewhere [15]. To be included in the trial the child had to have a primary problem that falls within the domains of anxiety, depressive symptoms or behavioural problems, according to the classification by the psychologist conducting the interview. Children with prior mental disorder diagnosis and children with indication of severe mental disorders (e.g. signs of a full syndrome of ADHD and autism spectrum disorder) were excluded. Based on sample size calculation and a pilot trial a total of 396 children aged 6-15 from four Danish municipalities across the country were included and randomized on an individual level to the intervention or treatment-as-usual.

2.2 Measures

Data was collected via an online platform at both baseline and end-of-treatment 18 weeks later. Children and parents as proxies for the children completed the Danish versions of SDQ, KIDSCREEN-27 (KIDSCREEN) and CHU9D at baseline and end-of-treatment. This study solely focuses on parental responses since the SDQ was not reported by children younger than 11 years (which accounts for 48% of the study population). All three questionnaires were in Danish using validated translations.

The response rates were 100% for CHU9D and KIDSCREEN at baseline with only one participant not completing the SDQ. The response rate for CHU9D was 85.6% (n=339) at end-of-treatment, of these 100% also completed the SDQ and KIDSCREEN.

All data from the trial are still blinded and will remain so until after the last follow-up in the trial and analyses of outcomes is completed. Thus, it is beyond the scope of the present study to describe group differences in change. A cost-utility analysis of the intervention using CHU9D will be conducted later and published in a separate article.

2.2.1 The Strengths and Difficulties Questionnaire

The SDQ is a widely used and well validated questionnaire aiming to assess children's mental health problems in both clinical samples and in general population [16–19]. The SDQ contains 25 items, which cover five subscales relating to the children's emotional problems, peer problems, behavioural problems, hyperactivity

and pro-social behaviour. Responses to the subscales on emotional problems, peer problems, behavioural problems, and hyperactivity can be used to calculate a total difficulties score (SDQ-TD). Each subscale score ranges from 0-10, implying that the total difficulties score ranges from 0 to 40 [20]. An extended version of the SDQ includes an impact assessment to evaluate how much the identified mental difficulties interfere with the child's everyday life. An impact score (SDQ-I) is calculated from five items; whether the difficulties upsets or distresses the child and how much the difficulties interfere with home life, friendships, classroom learning, and leisure activities. Each item is scored on a scale from 0 to 2. To score 1 or 2, the interference from the difficulties in that domain must be assessed to either "quite a lot" or "a great deal" [18]. The impact score is the primary outcome of the Mind My Mind Trial. When completing the SDQ parents were asked to respond as a proxy for their child based upon the preceding six months at baseline, and the preceding month at end-of-treatment.

2.2.2 KIDSCREEN

KIDSCREEN is a 27-item generic measure of HRQOL and well-being. A total of 13 European countries were included in the cross-cultural harmonization and development of the measure. Several studies have found it to be valid and reliable in children with and without chronic health conditions, demonstrating adequate psychometric properties [21, 22]. KIDSCREEN measures HRQOL and well-being across five domains: Physical Well-Being, Psychological Well-Being, Autonomy & Parents, Peers & Social Support and School Environment. Item responses are based on a five-point Likert scale and T scores for each domain are computed with a mean of 50 and standard deviation of 10, whereby higher scores indicate better HRQOL [23]. KIDSCREEN domain scores do not allow for the calculation of a global HRQOL score. When completing the KIDSCREEN the parents were asked to respond as a proxy for their child based upon the last week.

2.2.3 Child Health Utility 9D

CHU9D is a generic preference-based HRQOL measure designed specifically for use in economic evaluation of health care interventions in children and adolescents. CHU9D has 9 items with five levels of severity representing nine dimensions of HRQOL: Worried, Sad, Pain, Tired, Annoyed, Schoolwork/homework, Sleep, Daily routine, and Activities. The responses to the nine items can be converted to utilities, on the 0–1 dead–full health QALY scale, using preference-based scoring algorithms. In this study two separate scoring algorithms were applied. The original algorithm is based on the standard gamble method of health state valuation and the preferences of an adult general population in the United Kingdom (N=300). This algorithm generates utility scores ranging from 0.3261 (pit-state) to 1 (perfect health) [10]. A newer algorithm is based on best-worst scaling methods and the preferences of adolescent Australians aged 11-17 from the general population (N=1,982) and a smaller sample (N=152) time-trade-off experiment with young adults to anchor the tariffs. This algorithm generates utility scores ranging from -0.1059 (pit-state) to 1 (perfect health) [13]. Danish-specific preference weights are not yet available. When completing CHU9D the parents were asked to respond as a proxy for their child based upon their HRQOL on the present day.

2.2.4 Conceptual overlaps

Table 1 provides a simplified overview of the conceptual overlaps between the three different measures. We categorized the items and the subscales of the three instruments into seven dimensions of quality of life based on direct comparisons of the content, even though the content and concepts are not likely to be

independent of each other. Table 1 provides information on which measures we hypothesize CHU9D to be closest related with. Thus, CHU9D has the largest conceptual overlap with SDQ-I followed by SDQ-TD and KIDSCREEN's Physical Well-being and Psychological Well-being. Contrary, there is no clear conceptual overlap between CHU9D and the KIDSCREEN Social Support & Peers measure.

Table 1: Simplified conceptual overlap between measures by quality of life dimensions

	Quality of life dimensions						
	Mental health	Physical health	Sleep	Parental relation	Social relations	School	Activities/Daily routine
Child Health Utility 9D	**X**	**X**	**X**			**X**	**X**
SDQ Total difficulties score	**X**				**X**		
SDQ Impact Score	**X**			**X**	**X**	**X**	**X**
KIDSCREEN Physical Well-being		**X**					**X**
KIDSCREEN Psychological Well-being	**X**						**X**
KIDSCREEN Autonomy & Parent Relation				**X**			
KIDSCREEN Social Support & Peers					**X**		
KIDSCREEN School Environment						**X**	

2.3 Psychometric Analyses

2.3.1 Construct validity

To assess construct validity, the baseline data was used, and the discriminant validity and convergent validity were examined.

Discriminant validity was assessed by testing whether CHU9D can discriminate between groups defined by the SDQ-TD, the SDQ-I and the KIDSCREEN Psychological well-being score. The entire sample in this study

exhibited some degree of mental health problems distributed on a continuum, and it was therefore not possible to define clearly distinguishable categories. Instead we assessed whether CHU9D could distinguish between groups of children with different levels of problems using percentiles as cut-off values on the SDQ-TD, SDQ-I, and KIDSCREEN Psychological Well-being score. We focused on these scores as mental health problems are expected to have the largest impact on HRQOL. The study sample is divided into three groups: the children with the 25% lowest scores (low), the 25% highest scores (high) and the 50% in between (medium). Statistical differences were tested using Kruskal-Wallis test due to non-normality of the utility distributions (tested using the Shapiro-Francia test), and the magnitude of mean difference was assessed based on a minimally important difference (MID) of 0.03 [24] as no formal MID is available for CHU9D. Regression analyses were conducted to control for potential confounding characteristics, including age, gender and family socioeconomic status (proxied by the highest achieved education level of the parent).

Convergent validity was assessed using Spearman rank correlation coefficients. Correlation between CHU9D, the SDQ-TD, the SDQ-I and the KIDSCREEN scores was assessed. Based on Table 1 we hypothesized moderate correlations between CHU9D and the SDQ-TD and SDQ-I scores, KIDSCREEN Psychological Well-being and Physical Well-being. For the other KIDSCREEN scores we hypothesized a low but positive correlation, as higher scores in these conceptually less overlapping dimensions of HRQOL would to some degree still be expected to correlate with higher CHU9D utility scores. A complete correlation matrix at the dimension/item level for CHU9D and each of the SDQ and KIDSCREEN scores and items is available in the supplementary material. Following established guidelines the following categories for Spearman rank correlations are used; ≥0.5, strong; ≥ 0.3 to <0.5, moderate; and <0.3, weak [25].

2.3.2 Responsiveness

To assess responsiveness the floor and ceiling effects were firstly examined; next, magnitude of change over time, and the ability to differentiate between improvement and no improvement were investigated.

Floor or ceiling effects (i.e. more than 15% respondents scored the lowest or highest possible score) affect the ability of the measure to detect deterioration or improvements in health respectively [26]. For CHU9D we hypothesized a low percentage at the floor and ceiling at baseline, but we expected a higher percentage at the ceiling at follow-up, given that an effective intervention should improve the mental health of the respondents randomized to intervention. The floor and ceiling effects on SDQ and KIDSCREEN scores are used as reference values for examining CHU9D.

The magnitude of change in scores from baseline to end-of-treatment was assessed using the standardised response mean (SRM) statistic. The following categories for SRM are used; <0.2 small; 0.5, moderate; and >0.8, large [25]. We firstly report the SRM for the whole sample. To study the responsiveness, we identified sub-groups of children whose mental health condition had improved according to the standardized measures SDQ and KIDSCREEN. Children who had improved at least 1 point on the SDQ-I was examined, as this is considered a minimum clinical important difference [27]. For these groups of children with improved mental health we hypothesized that CHU9D demonstrate a change in the same direction as SDQ and KIDSCREEN Psychological Well-being scale. Given that the latter two scales are more specific to the intervention, it was expected that larger effects would be found relative to CHU9D. The SRMs from the SDQ and KIDSCREEN scales are presented as reference values.

The mean change in CHU9D score for the children with improved mental health was estimated and compared with the mean changes for the children whose condition did not improve or got worse. Due to non-normality of the utility distributions, statistical differences were tested using Mann-Whitney test. The interpretation of the magnitude of mean difference was again based on a MID of 0.03. Regression analyses were conducted to control for potential confounding factors.

3. Results

The characteristics of the participants and the baseline scores are presented in Table 2. Using the Mann-Whitney test/Kruskal-Wallis test no associations ($p>0.1$) between age, gender or parent's education and CHU9D score are found regardless of the scoring algorithm applied.

Table 2: Characteristics and baseline scores for participants

	n	Mean (SD) / %
Age, mean (standard deviation)	396	10.7 (2.36)
Gender, %		
Boy	206	48%
Girl	190	52%
Mothers' highest achieved education, %		
Basic school (0-10 years)	26	6.6%
High school / vocational education	112	28.3%
Bachelor / Diploma degree	216	54.5%
Master / PhD	42	10.6%
Fathers' highest achieved education, %		
Basic school (0-10 years)	60	15.3%
High school / vocational education	166	42.2%
Bachelor / Diploma degree	120	30.5%
Master / PhD	47	12.0%
Baseline scores, mean (standard deviation)		
CHU9D adult weights	396	0.804 (0.121)
CHU9D adolescent weights	396	0.629 (0.214)
SDQ Total difficulties score (SDQ-TD)	395	16.06 (5.36)
SDQ Impact Score (SDQ-I)	395	4.16 (2.39)
KIDSCREEN Physical Well-being	396	44.40 (10.2)
KIDSCREEN Psychological Well-being	396	43.95 (9.12)
KIDSCREEN Autonomy & Parent Relation	394	46.93 (9.05)
KIDSCREEN Social Support & Peers	395	43.75 (10.8)
KIDSCREEN School Environment	394	46.16 (9.02)

Note: CHU9D: Child Health Utility 9D. SDQ: The Strength and Difficulties Questionnaire. KIDSCREEN: KIDSCREEN-27.

3.1 Construct validity

3.1.1 Discriminant validity

CHU9D performed well in discriminating between the groups scoring low, medium, and high on the three mental health related scales regardless of the scoring algorithm applied. Table 3 shows that the group mean differences on CHU9D utility scores were all statistically significant and larger than the commonly used MID value of 0.03 [24]. The absolute magnitudes on the group mean differences were larger when using the adolescent weights than the adult weights.

In controlling for potential confounding characteristics, a series of regression analyses were conducted; each time one of the three severity scores (included as two dummy variables), age, gender, and parents highest achieved education were used as independent variables, whilst CHU9D utility was used as the dependent variable. These regression analyses (not reported but available upon request from the corresponding author) support the conclusions from Table 3.

Table 3: Discriminant validity

		CHU9D adult weights, Mean (SD)	**CHU9D adolescent weights, Mean (SD)**
SDQ Total difficulties score (SDQ-TS)	Low (0-12) n=106	0.857 (0.095)	0.725 (0.175)
	Medium (13-19) n=184	0.796 (0.119)	0.615 (0.212)
	High (20-40) n=105	0.765 (0.129)	0.558 (0.220)
	P-level*	<0.001	<0.001
SDQ Impact score (SDQ-I)	Low (0-2) n=100	0.858 (0.093)	0.729 (0.172)
	Medium (3-5) n=186	0.807 (0.114)	0.631 (0.203)
	High (6-10) n=109	0.750 (0.133)	0.535 (0.225)
	P-level*	<0.001	<0.001
KIDSCREEN Psychological Well-being	High (44.4-100) n=95	0.890 (0.076)	0.789 (0.134)
	Medium (34-44.3) n=198	0.812 (0.099)	0.640 (0.184)
	Low (0-33) n=103	0.710 (0.128)	0.460 (0.204)
	P-level*	<0.001	<0.001

Note: *Kruskal-Wallis test. CHU9D: Child Health Utility 9D. SDQ: The Strength and Difficulties Questionnaire. KIDSCREEN: KIDSCREEN-27. SD: Standard Deviation

3.1.2 Convergent validity

In assessing the convergent validity, the correlations between each of the SDQ and the KIDSCREEN scores on one hand and CHU9D score on the other hand were calculated; the correlation coefficients were similar regardless of CHU9D weights. Moderate correlations between CHU9D score and the SDQ-I and the KIDSCREEN Physical Well-being, respectively, were found as hypothesized. For KIDSCREEN Psychological Well-being a strong correlation was evident with CHU9D, whilst a weak (but close to moderate) correlation

was found for SDQ-TD with CHU9D. For the remaining KIDSCREEN scores weak correlations were found in the hypothesized direction, as higher KIDSREEN score associated with higher CHU9D utility score. Consistent with the overview of overlapping concepts presented in Table 1, a stronger correlation with CHU9D was found for KIDSCRREN School Environment compared to Autonomy & Parent Relation and Support & Peers.

Table 4: Spearman rank correlation coefficients for SDQ, KIDSCREEN and CHU9D utility scores.

	CHU9D adult weights	**CHU9D adolescent weights**
SDQ Total difficulties score (SDQ-TD)	-0.285	-0.294
SDQ Impact Score (SDQ-I)	-0.330	-0.329
KIDSCREEN Physical Well-being	0.438	0.408
KIDSCREEN Psychological Well-being	0.565	0.571
KIDSCREEN Autonomy & Parent Relation	0.197	0.200
KIDSCREEN Social Support & Peers	0.217	0.187
KIDSCREEN School Environment	0.266	0.260
All P-value <0.001		

Note: CHU9D: Child Health Utility 9D. SDQ: The Strength and Difficulties Questionnaire. KIDSCREEN: KIDSCREEN-27.

Despite not measuring the exact same concepts and the different time recall periods for each instrument, the specific CHU9D dimensions/items corelated with specific items of the SDQ or one of the KIDSCREEN scales, in accordance with our hypotheses based on item content. E.g. moderate correlations were found between CHU9D item “sad” and SDQ item “Often unhappy” and CHU9D item “worried” and KIDSCREEN Psychological Well-being. A complete correlation matrix on item- and scale-level between CHU9D and the SDQ and KIDSCREEN is presented in the supplementary material.

3.2 Responsiveness

3.2.1 Floor and ceiling effect

When examining CHU9D utility scores at baseline and end-of-treatment, no evidence of a floor effect was found. None of the children in our study sample were reported as being in the worst health state at either time points. A ceiling effect was evident for CHU9D at end-of-treatment since 54 children (16%) reported as having no impairments (full health) in all nine dimensions. There was no ceiling effect at baseline since only 12 children (3%) reported to be in full health. In comparison the same tendency was found for the SDQ-I, where 30 children (8%) had the lowest possible score (no impact) at baseline and 105 children (31%) had it at end-of-treatment. In contrast, both the SDQ-TD and the KIDSCREEN scores exhibited neither floor nor ceiling effects (between 0 and 4%).

3.2.2 Standard response mean

For CHU9D using adult weights a mean improvement of 0.055 with a standard deviation of 0.121 was found resulting in a SRM of 0.452. In contrast, for CHU9D the adolescent weights generated a mean improvement

of 0.102 with a standard deviation of 0.220, equivalent to a SRM of 0.462. In comparison SRMs of between 0.458 and 0.767 were found for the mental health specific scores; SDQ-TD, SDQ-I and KIDSCREEN Psychological Well-being. For the non-mental health KIDSCREEN scores we found SRMs between 0.131 and 0.323.

A total of 233 children obtained a clinical improvement of more than 1 point on the SDQ-I (which is deemed the minimum important difference [27]). This group exhibited a moderate SRM of 0.634 (adult weights) and 0.654 (adolescent weights).

3.2.3 Discriminant validity

CHU9D was able to discriminate between the group of children that experienced an improvement and the group of children that did not on both the SDQ-TD, SDQ-I and KIDSCREEN Well-being score. The differences in mean change were statistically significant and larger than the MID of 0.03. As seen in Table 5 the mean difference between the groups was considerable larger when using the adolescent weights compared to when using the adult weights.

Regression analyses were conducted to handle potential confounding variables in the same manner as described previously in Section 3.1.1. The regression analyses support the conclusions from Table 5.

Table 5: Mean change in CHU9D utility scores for those who improved, or did not improve, post intervention and the mean difference between these two groups

Change in:	Change in CHU9D adult weights, mean (n)	Difference, mean (95% CI)	Change in CHU9D adolescent weights, mean (n)	Difference, mean (95% CI)
SDQ Total difficulties score (SDQ-TD) ≥0 n=105 <0 (improved score) n=233	 0.017 (105) 0.072 (233)	 0.055*** (0.028;0.083)	 0.030 (105) 0.135 (233)	 0.105*** (0.055;0.155)
SDQ Impact Score (SDQ-I) ≥0 n=105 <0 (improved score) n=233	 0.015 (105) 0.074 (233)	 0.059*** (0.032;0.086)	 0.026 (105) 0.137 (233)	 0.111*** (0.061;0.160)
KIDSCREEN Physical Well-being <0 n=86 ≥0 (improved score) n=253	 -0.018 (86) 0.080 (253)	 0.098*** (0.070;0.126)	 -0.023 (86) 0.144 (253)	 0.168*** (0.117;0.219)
Mann-Whitney test *P >0.05 **P >0.01 ***P >0.001				

Note: CHU9D: Child Health Utility 9D. SDQ: The Strength and Difficulties Questionnaire. KIDSCREEN: KIDSCREEN-27. CI: Confidence Interval

4. Discussion

In examining the construct validity of CHU9D in a mental health setting this study has demonstrated that CHU9D is capable of discriminating between groups with different severity of mental health problems. In all cases, the mean difference between the groups was higher than the MID of 0.03. These results were found

to be robust when controlling for age, gender and parental education level. The utilities derived using the adolescent scoring algorithm did, however, discriminate substantially better between the groups. The average mean difference between the low-medium-high groups was 0.115 across the three measures using adolescent weights, while it was 0.063 when using the utilities derived from the adult weights. When used in a cost-utility analysis this difference is likely to have substantial impact on the incremental cost-effectiveness ratio (ICER).

For convergent validity similar results were evident regardless of the weights used to derive CHU9D utilities. CHU9D shows the hypothesized correlations with all measures except with the SDQ-TD. Here a correlation just below 0.3 was found and thereby categorized as weak. The weak correlation could be due to the differences in the scope of CHU9D and SDQ-TD. CHU9D aims to capture the impact of mental and other health related problems of the child, and the SDQ-TD aims to measure the symptoms of the mental health problems. To further analyse the correlation we compared our item correlations with those found by Furber & Segal [12]. In their study they highlighted five correlations at the dimension/item-level which they argued have a clear conceptual overlap. Moderate correlations were found for three of them and weak correlations for two. In comparison this study found moderate correlations for four of them and a weak correlation for one (correlations are marked in supplementary material Table B). Furber & Segal [12] furthermore found correlations between CHU9D utility (adult weights) and SDQ items above 0.2 for 11 of the 20 items, that makes the Total difficulties score, whereas this study found correlations of 0.1 or lower for the same correlations (correlation are marked in supplementary material Table B). In examining these 11 items in our cohort, on average only 12% of respondents indicated that their child was in the worst category. For six of the 11 items; *well behaved, one good friend, often fights and bullies, often lies/cheats, picked on/bullied,* and *steals* we find that there were less than 10% that responded in the worst category. These findings could point towards SDQ-TD includes items that are less relevant for this specific population. These findings, and the convergent validity of the other measures, make the weak correlation between CHU9D and SDQ-TD less of a concern in relation to convergent validity.

There was neither floor nor ceiling effects of CHU9D. Although 16% of the children reported full health in the follow-up, it should reflect the fact that after intervention, their HRQOL improved.

The analyses of SRM shows acceptable responsiveness of CHU9D regardless of the weights used to derive utilities. In the analyses of change in mean utility CHU9D was capable of distinguishing between the group of children whose mental health improved and those who did not. Using CHU9D adolescent scoring algorithm we found that the magnitude of the mean difference is considerable larger compared to when using the adult weights. These results again point towards the choice of utility weights is likely to have great impact on the ICER in a cost-utility analysis.

The difference in recall time adopted in different questionnaires may influence validation analysis of CHU9D. However, all questionnaires were completed online at the same time, this could possibly minimize the impact of the time perspective differences.

This study provides a broad validation for the use of CHU9D in mental health settings as the participants consist of children with a broad range of mental health problems, ranging from internalizing to externalizing problems and combinations. The results are however limited in their generalizability due to the lack of participants with severe clinical mental disorders. E.g. commonly used preference-based HRQOL instruments

in adult populations have been shown to be less appropriate in trials with schizophrenia patients [28]. The cross-sectional findings by Furber & Segal [12] in a population that includes severe mental disorders does, however, indicate CHU9D is also appropriate for use in such populations.

5. Conclusions

The findings from this study demonstrate that CHU9D is an appropriate preference-based HRQOL measure for use in mental health trials. The inclusion of CHU9D will enable cost-utility analysis of interventions aiming to improve child and adolescent mental health, and thereby provide valuable evidence for health care resource allocation and decision-making.

The results show that the preference weights generated from an adolescent population performs best in both discriminating between groups with different severity of mental health problems, and between the children that measured with SDQ and KIDSCREEN improve their mental health and those who do not. This finding likely is reflective of the relatively stronger weight attached to mental health impairments in the adolescent scoring algorithm in comparison with the adult scoring algorithm and demonstrates that the choice of scoring algorithm will impact upon the findings from economic evaluations.

Declarations

Funding: This study was supported by unrestricted grants from TrygFonden and the Lundbeck Foundation. The funding agencies were not involved in the study design, data collection, analysis and interpretation, the writing of the manuscript, or the decision on submission for publication.

Ethics approval: The trial was approved by Danish Ethics Committee: Journal-nr.: 17011408, Date of approval: June 20, 2017.

Consent to participate and for publication: Written informed consent signed by the legal guardians was provided for all the participants.

Conflicts of interest/Competing interests: The authors have declared that no competing interests exist.

Availability of data and material: The data utilized in the current study are defined as sensitive personal data, and cannot be shared publicly due to existing data protection laws in Denmark, and imposed by the Danish Data Protection Agency.

Code availability: Not applicable.

References

1. Romeo, R., Byford, S., & Knapp, M. (2005). Economic evaluations of child and adolescent mental health interventions: a systematic review. *Journal of Child Psychology and Psychiatry*, *46*(9), 919–930. https://doi.org/10.1111/j.1469-7610.2005.00407.x
2. Hill, H., Rowen, D., Pennington, B., Wong, R., & Wailoo, A. (2020). A Review of the Methods Used to Generate Utility Values in NICE Technology Assessments for Children and Adolescents. *Value in Health*. https://doi.org/10.1016/j.jval.2020.02.011

3. Rowen, D., Rivero-Arias, O., Devlin, N., & Ratcliffe, J. (2020). Review of Valuation Methods of Preference-Based Measures of Health for Economic Evaluation in Child and Adolescent Populations: Where are We Now and Where are We Going? *PharmacoEconomics*, *38*(4), 325–340. https://doi.org/10.1007/s40273-019-00873-7
4. Chen, G., & Ratcliffe, J. (2015). A Review of the Development and Application of Generic Multi-Attribute Utility Instruments for Paediatric Populations. *PharmacoEconomics*, *33*(10), 1013–1028. https://doi.org/10.1007/s40273-015-0286-7
5. Stevens, K. (2009). Developing a descriptive system for a new preference-based measure of health-related quality of life for children. *Quality of Life Research*, *18*(8), 1105–1113. https://doi.org/10.1007/s11136-009-9524-9
6. Chen, G., Flynn, T., Stevens, K., Brazier, J., Huynh, E., Sawyer, M., … Ratcliffe, J. (2015). Assessing the Health-Related Quality of Life of Australian Adolescents: An Empirical Comparison of the Child Health Utility 9D and EQ-5D-Y Instruments. *Value in Health*, *18*(4), 432–438. https://doi.org/10.1016/j.jval.2015.02.014
7. Petersen, K. D., Chen, G., Mpundu-Kaambwa, C., Stevens, K., Brazier, J., & Ratcliffe, J. (2018). Measuring Health-Related Quality of Life in Adolescent Populations: An Empirical Comparison of the CHU9D and the PedsQLTM 4.0 Short Form 15. *The Patient - Patient-Centered Outcomes Research*, *11*(1), 29–37. https://doi.org/10.1007/s40271-017-0265-5
8. Ratcliffe, J., Huynh, E., Stevens, K., Brazier, J., Sawyer, M., & Flynn, T. (2016). Nothing About Us Without Us? A Comparison of Adolescent and Adult Health-State Values for the Child Health Utility-9D Using Profile Case Best-Worst Scaling: ADOLESCENT VERSUS ADULT VALUES FOR THE CHU-9D. *Health Economics*, *25*(4), 486–496. https://doi.org/10.1002/hec.3165
9. Ratcliffe, J., Stevens, K., Flynn, T., Brazier, J., & Sawyer, M. (2012). An assessment of the construct validity of the CHU9D in the Australian adolescent general population. *Quality of Life Research*, *21*(4), 717–725. https://doi.org/10.1007/s11136-011-9971-y
10. Stevens, K., & Ratcliffe, J. (2012). Measuring and Valuing Health Benefits for Economic Evaluation in Adolescence: An Assessment of the Practicality and Validity of the Child Health Utility 9D in the Australian Adolescent Population. *Value in Health*, *15*(8), 1092–1099. https://doi.org/10.1016/j.jval.2012.07.011
11. Petersen, K. D., Ratcliffe, J., Chen, G., Serles, D., Frøsig, C. S., & Olesen, A. V. (2019). The construct validity of the Child Health Utility 9D-DK instrument. *Health and Quality of Life Outcomes*, *17*(1). https://doi.org/10.1186/s12955-019-1256-0
12. Furber, G., & Segal, L. (2015). The validity of the Child Health Utility instrument (CHU9D) as a routine outcome measure for use in child and adolescent mental health services. *Health and Quality of Life Outcomes*, *13*(1), 22. https://doi.org/10.1186/s12955-015-0218-4
13. Ratcliffe, J., Huynh, E., Chen, G., Stevens, K., Swait, J., Brazier, J., … Flynn, T. (2016). Valuing the Child Health Utility 9D: Using profile case best worst scaling methods to develop a new adolescent specific scoring algorithm. *Social Science & Medicine*, *157*, 48–59. https://doi.org/10.1016/j.socscimed.2016.03.042
14. Transdiagnostic, Cognitive and Behavioral Intervention for in School-aged Children With Emotional and Behavioral Disturbances - Full Text View - ClinicalTrials.gov. (n.d.). Retrieved August 14, 2019, from https://clinicaltrials.gov/ct2/show/NCT03535805

15. Wolf, R. T., Jeppesen, P., Gyrd-Hansen, D., The CCC2000 Study Group, & Oxholm, A. S. (2019). Evaluation of a screening algorithm using the Strengths and Difficulties Questionnaire to identify children with mental health problems: A five-year register-based follow-up on school performance and healthcare use. *PLOS ONE*, *14*(10). https://doi.org/10.1371/journal.pone.0223314
16. Goodman, A., & Goodman, R. (2009). Strengths and Difficulties Questionnaire as a Dimensional Measure of Child Mental Health. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(4), 400–403. https://doi.org/10.1097/CHI.0b013e3181985068
17. Goodman, R., Meltzer, H., & Bailey, V. (1998). The strengths and difficulties questionnaire: A pilot study on the validity of the self-report version. *Adolescent Psychiatry*, *7*(3), 7.
18. Goodman, R. (1999). The Extended Version of the Strengths and Difficulties Questionnaire as a Guide to Child Psychiatric Caseness and Consequent Burden. *J Child Psychol Psychiatry*, *40*(5), 791.
19. Niclasen, J., Teasdale, T. W., Andersen, A.-M. N., Skovgaard, A. M., Elberling, H., & Obel, C. (2012). Psychometric Properties of the Danish Strength and Difficulties Questionnaire: The SDQ Assessed for More than 70,000 Raters in Four Different Cohorts. *PLoS ONE*, *7*(2), e32025. https://doi.org/10.1371/journal.pone.0032025
20. Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x
21. Ravens-Sieberer, U., Auquier, P., Erhart, M., Gosch, A., Rajmil, L., Bruil, J., … the European KIDSCREEN Group. (2007). The KIDSCREEN-27 quality of life measure for children and adolescents: psychometric results from a cross-cultural survey in 13 European countries. *Quality of Life Research*, *16*(8), 1347–1356. https://doi.org/10.1007/s11136-007-9240-2
22. Robitail, S., Ravens-Sieberer, U., Simeoni, M.-C., Rajmil, L., Bruil, J., Power, M., … the KIDSCREEN Group. (2007). Testing the structural and cross-cultural validity of the KIDSCREEN-27 quality of life questionnaire. *Quality of Life Research*, *16*(8), 1335–1345. https://doi.org/10.1007/s11136-007-9241-1
23. Ravens-Sieberer, U., & Kidscreen Group Europe (Eds.). (2016). *The Kidscreen questionnaires: quality of life questionnaires for children and adolescents: handbook* (3rd edition.). Lengerich: Pabst Science Publishers.
24. Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford University Press.
25. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, N.J: L. Erlbaum Associates.
26. McHorney, C. A., & Tarlov, A. R. (1995). Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Quality of Life Research*, *4*(4), 293–307. https://doi.org/10.1007/BF01593882
27. Stringaris, A., & Goodman, R. (2013). The Value of Measuring Impact Alongside Symptoms in Children and Adolescents: A Longitudinal Assessment in a Community Sample. *Journal of Abnormal Child Psychology*, *41*(7), 1109–1120. https://doi.org/10.1007/s10802-013-9744-x
28. Mulhern, B., Mukuria, C., Barkham, M., Knapp, M., Byford, S., Soeteman, D., & Brazier, J. (2014). Using generic preference-based measures in mental health: psychometric validity of the EQ-5D and SF-6D. *British Journal of Psychiatry*, *205*(03), 236–243. https://doi.org/10.1192/bjp.bp.112.122283

Paper III Supplementary material

Table A: Spearman rank correlation coefficients for CHU9D items, SDQ Total difficulties score, SDQ Impact score and KIDSCREEN-27

	Worried	Sad	Pain	Tired	Annoyed	Schoolwork / homework	Sleep	Daily Routine	Activities
SDQ Total difficulties score	0.053	0.256***	0.083	0.036	0.195***	0.243***	0.018	0.251***	0.230***
SDQ Impact Score	0.183***	0.308***	0.074	0.113*	0.160**	0.248***	0.020	0.282***	0.346***
KIDSCREEN Physical Well-being	-0.266***	-0.329***	-0.251***	-0.283***	-0.103*	-0.172***	-0.157**	-0.257***	-0.436***
KIDSCREEN Psychological Well-being	-0.439***	-0.541***	-0.186***	-0.309***	-0.222***	-0.265***	-0.220***	-0.334***	-0.430***
KIDSCREEN Autonomy & Parent Relation	-0.004	-0.103*	-0.052	0.049	-0.177***	-0.153**	-0.079	-0.227***	-0.079
KIDSCREEN Social Support & Peers	-0.150**	-0.173***	-0.079	-0.056	-0.028	-0.121*	-0.046	-0.117*	-0.316***
KIDSCREEN School Environment	-0.114*	-168***	0.039	-0.012	-0.118*	-0.458***	-0.080	-0.199***	-0.297***

Note: *P<0.05 **P<0.01 ***P<0.001

Table B: Spearman rank correlation coefficients for CHU9D items, SDQ scales and SDQ items.

	Worried	Sad	Pain	Tired	Annoyed	Schoolwork /homework	Sleep	Daily Routine	Activities	**CHU9D adult weights**	**CHU9D adolescent weights**
Total difficulties	0.053	0.256***	0.083	0.036	0.195***	0.243***	0.018	0.251***	0.230***	-0.285***	-0.294***
Emotion scale	0.435***	0.342***	0.248***	0.134**	0.008	0.051	0.237***	0.143**	0.271***	-0.348***	-0.365***
Conduct scale	-0.265***	0.044	-0.046	-0.036	0.321***	0.074	-0.074	0.236***	0.036	-0.105*	-0.082
Hyper scale	-0.099*	0.074	-0.022	-0.026	0.170***	0.265***	-0.041	0.209***	0.096*	-0.131**	-0.149**
Peer scale	0.049	0.115*	-0.003	0.001	-0.037	0.152**	-0.070	-0.006	0.125*	-0.068	-0.070
Prosocial scale #	0.160**	-0.005	0.108*	0.091	-0.219***	-0.105*	0.109*	-0.132**	-0.057	0.036	0.032
Considerate #	0.222***	-0.038	0.078	0.064	-0.246***	-0.061	0.064	-0.084	-0.023	0.041	0.016
Restless	-0.051	0.017	-0.021	-0.017	0.079	0.108*	0.019	0.166***	0.001	*-0.045*	-0.078
Sickness	0.286***	0.248***	0.282***	0.135**	0.046	0.062	0.214***	0.192***	0.226***	-0.311***	-0.311***
Shares #	0.041	-0.056	0.104*	0.025	-0.156**	-0.061	0.081	-0.068	-0.008	0.010	0.0381
Loses temper	-0.169***	0.074	-0.010	0.017	**0.394*** **	0.039	-0.037	0.212***	0.067	-0.147**	-0.134**
Pref solitary	0.125*	0.089	-0.006	0.031	-0.076	0.155**	-0.032	0.006	0.194***	*-0.071*	-0.69
Well behaved #	0.250***	-0.041	0.029	0.035	-0.221***	-0.086	0.024	-0.242***	-0.008	*0.101**	0.061
Many worries	**0.399*** **	0.220***	0.175***	0.109*	-0.079	-0.002	0.229***	0.015	0.169***	-0.217***	-0.236***
Help others if hurt #	0.064	0.037	0.100*	0.106*	-0.087	-0.059	0.059	-0.131**	-0.050	0.009	0.016
Fidgeting	-0.075	0.038	0.012	0.001	0.092	0.117*	-0.036	0.136**	0.036	*-0.063*	-0.083
One good friend #	0.022	-0.079	0.026	0.007	-0.015	-0.060	-0.024	-0.008	-0.067	*0.056*	0.043
Often fights or bullies	-0.132**	-0.009	-0.099	-0.05	0.138	0.044	-0.059	0.033	-0.019	*0.027*	0.017
Often unhappy	0.255***	**0.400*** **	0.154**	0.177***	0.108*	0.054	0.140**	0.168**	0.210***	-0.319***	-0.321***

Generally liked by others #	0.045	-0.054	-0.000	0.063	-0.023	-0.059	0.057	-0.007	-0.032	0.022	0.019
Easily distracted	-0.062	0.076	-0.017	0.017	0.142**	0.260***	-0.063	0.142**	0.106*	-0.126*	-0.127*
Low confidence	**0.125***	0.068	0.055	-0.015	-0.024	0.025	0.039	0.033	0.144**	-0.083	-0.083
Kind to younger children #	0.065	0.015	0.028	0.007	-0.082	-0.068	0.050	-0.108*	-0.047	0.049	0.045
Often lies / cheats	-0.202***	-0.021	-0.062	-0.033	0.049	0.023	-0.132**	0.131**	0.010	0.001	0.014
Picked on or bullied	-0.030	0.111*	0.055	-0.023	0.013	0.071	-0.046	0.012	-0.032	*-0.029*	-0.044
Volunteers to help #	0.134**	0.022	0.031	0.082	-0.156**	-0.074	0.085	-0.086	-0.049	0.024	0.011
Thinks before acting #	0.138**	-0.035	0.067	0.103	-0.193***	-0.113*	0.055	-0.160**	-0.065	*0.072*	0.072
Steals	-0.126*	-0.054	-0.021	-0.031	0.004	0.007	-0.098	-0.014	-0.044	*0.053*	0.063
Better with adult than peers	0.121*	0.096	0.011	0.061	-0.017	0.116*	-0.044	-0.066	0.117*	*-0.083*	-0.087
Many fears	**0.340*****	0.199***	0.135**	0.049	-0.028	0.021	0.167***	0.019	0.133**	-0.201***	-0.218***
Good attention span / follow through #	0.057	-0.118*	0.009	-0.007	-0.169***	-0.446***	0.014	-0.227***	-0.164**	0.202***	0.221***

Notes: # Higher value is better. *P<0.05 **P<0.01 ***P<0.001. Bolded figures are those where CHU9D and SDQ have clear conceptual overlap according to Furber & Segal[12]. Italic figures in CHU9D adult weights column are the correlations where Furber & Segal[12] find correlations larger than 0.2 and where we find correlations of 0.1 or lower.

Table C: Spearman rank correlation coefficients for CHU9D items and SDQ Impact items

	Worried	Sad	Pain	Tired	Annoyed	Schoolwork /homework	Sleep	Daily Routine	Activities	**CHU9D adult weights**	**CHU9D adolescent weights**
Impact score	0.183***	0.308***	0.074	0.113*	0.160**	0.248***	0.020	0.282***	0.346***	-0.330***	-0.329***
The difficulties upsets or distresses the child	0.237***	0.395***	0.149	0.182***	0.111*	0.024	0.087	0.176***	0.206***	-0.300***	-0.293***
Home life	0.049	0.217***	0.027	0.105*	0.248***	0.023	0.061	0.269***	0.072	-0.220***	-0.215***
Friendships	0.164**	0.201***	0.116	0.084	0.051	0.133**	0.050	0.113*	0.288***	-0.228***	-0.219***
Classroom learning	0.033	0.066	-0.063	-0.049	0.031	0.407***	-0.098	0.082	0.176***	-0.081	-0.110*
Leisure activities	0.106*	0.172***	0.0867	0.142**	0.108*	0.176***	0.063	0.181***	0.324***	-0.258***	-0.242***

Notes: *P<0.05 **P<0.01 ***P<0.001

Table D: Spearman rank correlation coefficients for CHU9D items and KIDSCREEN-27 scales and items.

	Worried	Sad	Pain	Tired	Annoyed	Schoolwork /homework	Sleep	Daily Routine	Activities	**CHU9D adult weights**	**CHU9D adolescent weights**
Physical Well-being	-0.266***	-0.329***	-0.251***	-0.283***	-0.103*	-0.172***	-0.157**	-0.257***	-0.436***	0.438***	0.408***
General health status #	0.305***	0.318***	0.303***	0.120***	0.073	0.118*	0.146**	0.191***	0.334***	-0.373***	-0.324***
Feel fit and well	-0.347***	-0.361***	-0.287***	-0.267***	-0.120	-0.172***	-0.169***	-0.263***	-0.364***	0.450***	0.440***
Physically active	-0.138**	-0.199***	-0.156**	-0.191***	-0.034	-0.119*	-0.085	-0.176***	-0.352***	0.274***	0.245***
Run well	-0.125*	-0.212***	-0.187***	-0.180***	-0.037	-0.101*	-0.129**	-0.163***	-0.325***	0.303***	0.265***
Feelt energetic	-0.270***	-0.300***	-0.157**	-0.323***	-0.154**	-0.229***	-0.145**	-0.258***	-0.356***	0.410***	0.400***
Psychological Well-being	-0.439***	-0.541***	-0.186***	-0.309***	-0.222***	-0.265***	-0.220***	-0.334***	-0.430***	0.565***	0.571***
Life been enjoyable	-0.333***	-0.409***	-0.175***	-0.257***	-0.223***	-0.253***	-0.174***	-0.236***	-0.368***	0.443***	0.445***
In good mood	-0.341***	-0.397***	-0.196***	-0.254***	-0.251***	-0.167***	-0.187***	-0.260***	-0.331***	0.465***	0.453***
Have had fun	-0.292***	-0.384***	-0.104*	-0.184***	-0.183**	-0.203***	-0.129*	-0.252***	-0.360***	0.410***	0.404***
Felt sad #	0.0435***	**0.595****	0.177***	0.238***	0.245***	0.205***	0.203***	0.280***	0.285***	-0.509***	-0.541***
Felt too sad and didn't want to do anything #	0.346***	0.411***	0.162**	0.288***	0.163**	0.245***	0.246***	0.304***	0.398***	-0.484***	-0.481***
Felt lonely #	0.313***	0.378***	0.087-0.066	0.193***	0.062	0.221***	0.155**	0.185***	0.302***	-0.367***	-0.380***
Been happy with the way he/she is	-0.288***	-0.330***	-0.057	-0.201***	-0.063	-0.178***	-0.090*	-0.238***	-0.253***	0.316***	0.328***
Autonomy & Parent Relation	-0.004	-0.103*	-0.052	0.049	-0.177***	-0.153**	-0.079	-0.227***	-0.079	0.197***	0.200***
Child had enough time for	-0.085	-0.062	0.007	-0.042	-0.063	-0.025	-0.114*	-0.119**	0.028	0.106*	0.134**

him/herself											
Child been able to do the things that he/she wants to do in his/her free time	-0.093	-0.084	-0.055	-0.032	-0.057	-0.127*	-0.064	-0.153***	-0.093	0.151**	0.163**
Child felt that his/her parent(s) had enough time for him/her	-0.040	-0.147**	-0.047	-0.015	-0.152**	-0.067	-0.091	-0.125*	-0.047	0.178***	0.161**
Child felt that his/her parent(s) treated him/her fairly	0.056	-0.137**	-0.066	0.029	-0.319***	-0.035	-0.083	-0.239***	-0.031	0.189***	0.184***
Child been able to talk to his/her parent(s) when he/she wanted to	0.051	-0.102*	0.005	0.036	-0.159**	-0.073	-0.072	-0.192***	-0.076	0.176***	0.159**
Child had enough money to do the same things as his/her friends	0.057	0.040	-0.026	0.121*	0.007	-0.105*	0.022	-0.062	-0.029	0.016	0.022
Child felt that he/she had enough money for his/her expenses	0.087	0.050	-0.096	0.081	-0.026	-0.129*	0.036	-0.074	-0.033	0.018	0.019

Social Support & Peers	-0.150**	-0.173***	-0.079	-0.056	-0.028	-0.121*	-0.046	-0.117*	-0.316***	0.217***	0.187***
Child spent time with his/her Friends	-0.110*	-0.150**	-0.082	-0.128*	-0.033	-0.145**	-0.028	-0.120*	-0.334***	0.218***	0.184***
Child had fun with his/her friends	-0.188**	-0.164**	-0.031	-0.077	-0.014	-0.106*	-0.055	-0.108*	-0.321***	0.218***	0.184***
Child and his/her friends helped each other	-0.106*	-0.119*	-0.036	0.017	-0.020	-0.104*	-0.017	-0.063	-0.250***	0.146**	0.125*
Child been able to rely on his/her friends	-0.120*	-0.138**	-0.054	-0.010	-0.013	-0.076	-0.087	-0.094	-0.210***	0.161**	0.144**
School Environment	-0.114*	-168***	0.039	-0.012	-0.118*	-0.458***	-0.080	-0.199***	-0.297***	0.266***	0.260***
Child been happy at school	-0.301***	-0.263***	-0.054	-0.126*	-0.081	-0.347***	-0.145	-0.155**	-0.373***	0.328***	0.323***
Child got on well at school	-0.083	-0.092	0.062	0.035	-0.063	-0.497***	0.006	-0.187***	-0.269***	0.201***	0.208***
Child been able to pay attention	-0.056	-0.163**	0.027	-0.102*	-0.147**	-0.388***	-0.097	-0.188***	-0.252***	0.249***	0.244***
Child got along well with his/her teachers	0.054	-0.038	0.089	0.130**	-0.100*	-0.224***	-0.049	-0.084	-0.073	0.079	0.061

Notes: # Higher value is better. *P<0.05 **P<0.01 ***P<0.001

Paper IV

Cost-effectiveness of a transdiagnostic psychotherapy program for youth with common mental health problems.

Prepared for submission.

Cost-effectiveness of a transdiagnostic psychotherapy program for youth with common mental health problems

Rasmus Trap Wolf, M.Sc.[1,2], Pia Jeppesen M.D., Ph.D.[1,3], Mette Maria Agner Pedersen M.Sc.[1], Louise Berg Puggaard M.Sc., Ph.D.[1], Mikael Thastum M.Sc., Ph.D.[4], Niels Bilenberg M.D., Ph.D.[5,6], Per Hove Thomsen M.D., Dr.Med.[7,8], Wendy K. Silverman M.Sc., Ph.D.[9], Kerstin Jessica Plessen M.D., Ph.D.[1,10], Simon-Peter Neumer M.Sc., Ph.D.[11,12], Christoph U. Correll M.D.[13,14,15,16], Anne Katrine Pagsberg M.D., Ph.D.[1,3], Dorte Gyrd-Hansen M.Sc., Ph.D.[2]

[1]Child and Adolescent Mental Health Center, Mental Health Services – Capital Region of Denmark, Copenhagen, Denmark

[2]Danish Centre for Health Economics, Department of Public Health, University of Southern Denmark, Odense, Denmark

[3]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[4]Centre for the Psychological Treatment of Children and Adolescents, Department of Psychology and Behavioural Sciences, Aarhus BSS, Aarhus University, Aarhus, Denmark

[5]Department for Child and Adolescent Psychiatry, Mental Health Services in the Region of Southern Denmark, Odense, Denmark

[6]Institute of Clinical Medicine, University of Southern Denmark, Odense, Denmark

[7]Research Center at the Department for Child- and Adolescent Psychiatry, Aarhus University Hospital, Skejby, Denmark

[8]Institute of Clinical Medicine, Aarhus University, Aarhus, Denmark

[9]Anxiety and Mood Disorders Program, Yale Child Study Center, Yale School of Medicine, New Haven, Connecticut, United States

[10]Division of Child and Adolescent Psychiatry, Department of Psychiatry, Lausanne University Hospital CHUV, Lausanne, Switzerland

[11]Centre for Child and Adolescent Mental Health, Oslo, Norway.

[12]The Arctic University of Norway, Centre for Child and Youth Mental Health and Child Welfare, North Norway (RKBU North), Tromsø, Norway.

[13]Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Department of Psychiatry and Molecular Medicine, Hempstead, New York, United States

[14]The Zucker Hillside Hospital, Department of Psychiatry, Glen Oaks, New York, United States

[15]Feinstein Institute for Medical Research, Center for Psychiatric Neuroscience, Manhasset, New York, United States

[16]Charité Universitätsmedizin, Department of Child and Adolescent Psychiatry, Berlin, Germany

Funding: This study was supported by unrestricted grants from TrygFonden and the Lundbeck Foundation. The funding agencies were not involved in the study design, data collection, analysis and interpretation, the writing of the manuscript, or the decision on submission for publication.

Acknowledgements: The private, noncommercial health science research project Mind My Mind was initiated by Psykiatrifonden. Birthe Wielandt Houe acted as project manager on behalf of Psykiatrifonden. The randomized controlled effectiveness trial was carried out through close collaboration between Psykiatrifonden, the principal investigator (Pia Jeppesen), the participating municipalities, and the regional child and adolescent mental health centers. Psykiatrifonden held the legal responsibility as the data controller for the Web-based data collection.

Abstract

Objectives: Our objective was to evaluate the cost-effectiveness of the transdiagnostic psychotherapy program Mind My Mind (MMM) for youth with common mental health problems using a cost-utility analysis (CUA) framework and data from a randomized controlled trial. Furthermore, we analyzed the impact of choice of informant with respect to both quality-of-life reporting and preference weights on the Incremental Cost Effectiveness Ratio (ICER).

Methods: A total of 396 school-aged youth took part in the trial. CUAs were carried out for the trial-period and for four one-year extrapolation scenarios. Costs were based on a combination of budget and self-reported costs. Youths and parents were asked to report on the youth's quality-of-life three times during the trial using the Child Health Utility 9D (CHU9D). Parental-reported CHU9D was used in the base case together with preference weights of a youth population. Analyses using self-reported CHU9D and preference weights of an adult population were also carried out.

Results: The analysis of the trial period resulted in an ICER of €170,465. This analysis does, however, inherently underestimate the cost-effectiveness of MMM. The analyses of the one-year scenarios resulted in ICERs between €23,653 and €50,480. The ICER increased by 24% and 71% when using self-reported CHU9D and adult preference weights, respectively.

Conclusion: The MMM intervention has the potential to be cost-effective, but the ICER is dependent on the duration of the treatment effects. Results varied significantly with the choice of respondent and the choice of preference weights indicating that both factors should be considered when assessing CUA involving youth.

1. Introduction

Mental health disorders affect a large proportion of children and adolescents (youth)[1–3]. The most common mental health disorders among youth; mild to moderate symptoms of anxiety, depression, and disruptive behavior disorders have an increased risk of adverse adult outcomes[4]. Several studies have found promising effects of cognitive-behavioral therapy (CBT) programs for these disorders[5–12]. Still, there is limited access to evidence-based programs. Mind My Mind (MMM) is a transdiagnostic CBT-program for indicated prevention and early treatment of anxiety, depression, and behavioral disturbances below threshold for psychiatric referral. It was designed to address the problem of limited access by providing a program that is feasible for large scale implementation. MMM has recently demonstrated superiority over Management As Usual (MAU) in a pragmatic, multi-site, and randomized controlled trial (RCT)[13].

Despite the abundance of studies investigating interventions for common mental health disorders among youth, evidence on cost-effectiveness of indicated prevention and interventions is still limited[14–20]. CBT-programs typically involve a considerable cost due to the amount of time delivered by the therapists. The budget impact is likely to be part of the explanation for the limited access to evidence-based programs. This calls for cost-utility analyses (CUA) of evidence-based interventions to inform decision makers.

Analyzing the cost-utility of interventions for youth presents several challenging aspects. While the methodology and the guidelines for estimating Quality Adjusted Life Years (QALY) in adult populations are well developed, there is less clear guidance when it comes to youth populations[21]. A recent review, therefore, warrants further empirical evidence on the valuation of youth-specific preference-based measures[22]. In a mental health setting, the choice of the respondent is especially important, as experiences across youths and parents are likely to vary[23]. Similarly, the preferences of the different quality-of-life health states vary in the mental dimensions when comparing preferences weights derived from youth and adult populations[24].

In this study, we analyzed the cost-effectiveness of the MMM intervention using a CUA framework. We also present how the CUA results were impacted by the choice of respondent (patients or parents) and the choice of preference weights (tariffs) used in the QALY estimation. The CUA was based on data collected in the RCT of MMM versus MAU[13]. MMM aims at treating current symptoms and preventing the development of severe mental disorders. In order to derive a true estimate of the long-term effect of an intervention it is recommended that health effects and costs are extrapolated for as long as they are assumed to differ between the compared interventions. This is often done using decision analytic modeling that incorporate data from external sources[25]. However, there is limited data to inform modeling in this case. Few studies of CBT-programs have long term follow-up, still beneficial and statistical significant effects are found up to 12 months after end of treatment[5–12]. Several cohort studies have shown associations between psychopathological outcomes and later risk of severe mental disorder[26,27]. The associations have, however, not been investigated in an experimental design, which makes intervention-based improvements uncertain. The limited data and uncertainty about the long-term effects lead us to analyze different possible scenarios of a one-year extrapolation period only. This means that we primarily analyzed the treatment effect of MMM, and not the potential preventive effects.

2. Methods

2.1 Trial design

Full details of the trial design are described in separate papers[13,28]. In brief, the MMM trial was a pragmatic randomized controlled trial carried out in a local municipality setting. Participants were randomly allocated on an individual level 1:1 via independent, blinded, computer-generated allocation sequences with variable and unknown block sizes, stratified by geographical region, age-group (6–10 years or 11–16 years) and their top-problem (anxiety, depressive symptoms, behavioral problems). Data used in the present analyses were collected at baseline, end-of-treatment (week 18), and at follow-up (week 26). The trial was approved by the local scientific ethics committee (Journal nr.:H-17011408). The study was registered at ClinicalTrials.gov (Identifier NCT03535805).

2.1.1 Participants

A total of 396 (197 randomized to MMM and 199 to MAU) youth aged 6-16 years were enrolled in the trial. A two-stage standardized visitation was implemented in the Educational-psychological advisory service in the four participating municipalities to identify study participants[28]. Eligibility criteria were: 1) Age between 6 to 16 years, 2) being in 0-9th grade, 3) having a parent-reported score on The Strength and Difficulties Questionnaire (SDQ) above a cut-off based on top 10% percent of mental health problems in the general age-matched Danish population[29], and 4) having anxiety, depressive symptoms, or behavioral problems as the top-problem based on the standardized assessments in the visitation procedure. Youth with prior diagnosis of any developmental or mental disorder, including autism spectrum disorder, attention-deficit/hyperactivity disorder, psychotic disorder, eating disorder, severe obsessive-compulsive disorder, repeated self-harm, alcohol or psychoactive drug abuse, or with signs of intellectual disability were not eligible for the trial. Further description of the participants is available elsewhere[13,28].

2.2 Intervention

The MMM intervention is a transdiagnostic, manualized, modular CBT-program comprising 9-13 individual sessions plus a booster-session targeting anxiety, depression and/or behavioral problems. All treatment was delivered by trained local psychologist from the Educational-psychological advisory service in the four municipalities. The treatment was supervised by psychologists from the regional Child and Adolescent Mental Health Services (CAMHS). For further description of the training see Table SM 1 in the supplementary material. Youth randomized to MMM were not eligible for other therapeutic interventions offered by the municipalities during the treatment and follow-up period. They could seek help at other health care professionals, e.g. general practitioners (GP). If the mental health condition of the youth progressed to the level at which specialized treatment was indicated according to standard guidelines, they were referred to CAMHS at any point during the trial[13].

2.3 Comparator

Youth and parents in the MAU group were offered two care-coordination visits to help coordinate usual care in the municipality. These visits were not a standard part of usual care in the municipalities but were introduced in the trial period for ethical concerns. The treatment offered in MAU varied considerably from no intervention to individual and group therapy, and parental training. Psychologists trained in MMM did not

conduct care-coordination visits or provide any kind of therapy to the MAU group. The youth in the MAU group could also seek help at other health care professionals or be referred to CAMHS.

2.4 Perspective and time horizon

The CUAs was conducted from an extended health sector perspective and included the costs of all interventions directed toward the individual youth's mental health problems no matter the provider of services.

The base case analysis had a 26-week time horizon consistent with the trial length. Four one-year extrapolations scenarios were also analyzed: 1) Temporary or catch-up effect, 2) No development, 3) Back to baseline, 4) Continuing trend from end-of-treatment to follow-up. In scenario 1) we assumed that the effect of MMM was either temporary or there was a catch-up effect from the worst performing group so that the MMM and the MAU group would end up having the same average health state utility value (HSUV) one year after the last observed time point (week 26). The incremental QALY-gain is the same independently of which group is assumed to be steady and which group is assumed to change in mean HSUV, thus the scenarios were analyzed as one. In scenario 2) we assumed that both groups would have the same HSUV one year after follow-up as they had at follow-up. In scenario 3) we assumed that the trial period only had a temporary effect, so that one year after follow-up all youth would have the same HSUV as they had at baseline. In scenario 4) we assumed that the trend in the period from end-of-treatment (week 18) to follow-up in which only a booster-session was given, would continue for one year. For scenario 4) we also assumed the highest and lowest HSUV at the extrapolation timepoint to be the highest and lowest observed HSUV. We assumed no group difference in costs during the one-year extrapolation in all analyses of the scenarios.

2.5 Costs

All costs were estimated at the individual participant level. Table SM 1 in the supplementary material provides information on the specific cost components, assumptions, sources of unit costs, and cost estimates. All costs are presented in Euros (€). We converted the costs from Danish kroner (DKK) using the currency rate DKK100 = €13.38.

2.5.1 Mind My Mind

The intervention costs comprise the cost of training the psychologists and the supervisors, the costs of materials for the intervention, and the salaries of the psychologists and supervisors. The training costs were divided equally among youth to estimate individual level costs. As the training of the psychologists can be considered an investment, the full return of the training will not be achieved within the trial period. This was addressed in a sensitivity analysis.

2.5.2 Management as usual

The MAU costs comprise cost of the care-coordination visits and costs of individual and group therapy, and parent psychoeducation provided by the municipality.

2.5.3 Other health care use

The other costs included in the analyses comprise health care use related to mental health problems including visits at the GP, pediatrician, CAMHS, and private-sector psychologists.

2.6 Effect

The health effects are expressed in QALY. To calculate QALYs the youth's health states must be determined and each state must be given a utility. In the base case, health states were defined using parent's answers to The Child Health Utility 9-Dimension (CHU9D), and the utilities associated with each health state were derived from tariffs based on preferences of youth aged 11-17 years.

CHU9D is a generic preference-based health-related quality of life (HRQOL) measure constructed for use in youth. CHU9D has nine items with five levels of severity representing nine dimensions of HRQOL: Worried, Sad, Pain, Tired, Annoyed, Schoolwork/homework, Sleep, Daily routine, and Activities[30]. CHU9D is validated for use in mental health settings[31,32]. Parents and youth both completed the CHU9D questionnaire online at baseline, end-of-treatment, and a follow-up. For consistency with the primary clinical outcome of the trial[13], and due to the lower proportion of missing data, parents' responses were used in the base case.

A validation of CHU9D in a mental health setting has been done using blinded data from the MMM trial[32], but there are no Danish tariffs for CHU9D. In the study, two sets of tariffs were investigated: 1) preferences of Australians aged 11-17 years from the general population (N=1,982)[33], 2) preferences of an adult general population in the United Kingdom (N=300)[34]. The validation study showed that the tariffs based on the youth's preferences were better in distinguishing between groups with different severity of mental health problems[32]. Thus, the tariffs of Australians aged 11-17 years were used in the present base case.

2.7 Missing data

Descriptive analyses were carried out to determine how to handle missing outcome data following a guide for handling missing data in cost-effectiveness analyses conducted within randomized trials[35]. There were no missing baseline CHU9D values, there were 57 (14%) missing at end-of-treatment, and 63 (16%) at follow-up. For the self-reported cost components there was 59 (15%) missing at end-of-treatment and 63 (16%) at follow-up. In total, 74 (19%) had a missing value at any timepoint. There was a higher proportion of missing values in the MAU group compared to the MMM group for both utilities and costs (74 (24%) versus 48 (18%) missing at any time point). Logistic regression models found treatment group to be the only statistically significant predictor of missing values when investigating baseline utility, sex, age-group, region, and top problem (anxiety, depression, or behavioral difficulties). Based on the findings, it was assumed that the probability of data being missing was independent of unobserved characteristics, hence the data was treated as being missing at random within each group.

Multiple-imputation using chained-equations was performed on utility score level and cost-component level. Due to the nature of utility scores and costs, predictive mean matching was used in the imputation. Twenty imputations were performed based on the proportion of missing data[36]. Imputations were performed separately for each trial group. To be consistent with the analysis of the clinical outcomes[13], the imputation models included baseline utility, sex, age-group, region, and top problem. After imputation, incremental QALYs and costs were calculated as the mean of the incremental QALY estimates and incremental costs estimates, respectively, generated in each imputed dataset following Rubin's rule[37]. The distributions of the observed and imputed values confirmed the validity of the imputation procedure. A complete case analysis was conducted and confirmed the robustness of the imputations. The analysis is available in the supplementary material Table SM 2.

2.8 Data analysis and results

The RCT found that the MMM intervention lead to significantly better clinical outcomes than MAU[13]. We thus expected positive incremental QALY results. The results of the CUAs are therefore presented as the Incremental Cost-Effectiveness Ratio (ICER):

$$ICER = \frac{\Delta \text{Cost}}{\Delta \text{Effect}} = \frac{\text{Cost}_{\text{MMM}} - \text{Cost}_{\text{MAU}}}{\text{QALY}_{\text{MMM}} - \text{QALY}_{\text{MAU}}}$$

Estimates of costs and effects were calculated using Seemingly Unrelated Regression models[38]. Adjustment for baseline utility were included in the estimation of QALYs to account for the imbalance between treatment groups[39]. P-values below 0.05 were considered statistically significant. All analyses were performed in STATA-15[40].

2.9 Uncertainty

2.9.1 Sample uncertainty

Sample uncertainty was examined using non-parametric bootstrapping with 10,000 iterations. Bootstrap samples of 396 were drawn from each of the 20 multiple imputed datasets and difference in net benefit between the groups calculated for each bootstrap sample given different thresholds[41]. The proportion of bootstrap samples in which the net benefit is positive represents the probability that the treatment is cost-effective for each imputed dataset. The probability is then averaged across all the imputed datasets to construct cost-effectiveness acceptability curves.

2.9.2 Sensitivity analyses

The training of the psychologist can be considered an investment in which the full return is based on the number of youths each psychologist treated. The restricted trial-period makes the number of youths each psychologist treated lower than the full potential. We therefore conducted sensitivity analyses, in which we assumed the number of treated youths per psychologist to be the highest number one psychologist treated within the trial-period.

The impact of the choice of informant who evaluates the youth's health state (the parent or the youth) and the tariffs that are subsequently applied (youth's or adult's) to estimate QALYs were explored by conducting a sensitivity analysis using youth's responses to CHU9D, and another sensitivity analysis using the tariffs derived from an adult general population.

Furthermore, sensitivity analyses of the one-year extrapolation period were performed by estimating ICERs for alternative time periods (six months and 18 months) in the four extrapolation scenarios presented in section 2.4.

3. Results

3.1 Cost

MMM was estimated to cost €3,471 on average per youth. The cost of the different components of MMM, including training of psychologists, is available in the supplementary material table SM 1. The cost of training a psychologist was €9,178 and €12,119 for training a supervisor. The cost of the training was divided equally among the youth randomized to MMM. This made the number of treated youths per psychologist in the trial

period the most determinant factor of the total cost of the intervention. On average each psychologist treated 8.2 youth, with the highest number being 20 youth for one psychologist. If all trained psychologist treated 20 youth the average cost would be €2,476.

The costs of MMM, MAU, and other health care utilization are presented in table 1. The 199 youths allocated to MAU received an average of 1.6 coordinating visits. 198 (95%) received at least one coordinating visit. The mean incremental costs for MMM were €2,981. MAU and private-sector psychologist were the primary sources of costs difference besides MMM. There were no other statistically significant differences in other health care utilization.

Table 1: Costs

	Mind My Mind (n=197)		Management as usual (n=199)	
	Units, Mean ±SE	Cost €, Mean ±SE	Units, Mean ±SE	Cost €, Mean ±SE
Mind My Mind	1	3,471 ±19		
Management as usual				
Care-coordination visits			1.6 ±0.0	111±3
Individual session with psychologist			1.7 ±0.3	117 ±20
Group therapy with psychologist			0.8 ±0.3	9 ±3
Parent psychoeducation			0.1 ±0.1	9 ±5
MAU Total				247 ±22
Other health care				
General practitioner	0.5 ±0.1	35 ±11	0.7 ±0.3	48 ±25
Pediatrician	0.4 ±0.2	27 ±11	0.3 ±0.1	17 ±8
Child and Adolescent Mental Health Services	0.3 ±0.1	82 ±30	0.4 ±0.1	99 ±32
Private-sector psychologist	0.9 ±0.3	107 ±34	2.8 ±0.7	329 ±85
Other health care total		252 ±63		493 ±101
Total costs		3,722 ±64		741 ±105
Incremental costs				2,981 (95% CI: 2,731-3,251)

Note: Costs are presented as mean (standard error (SE)) in Euros and are based on the intention-to-treat population from 20 imputed datasets. 95% Confidence interval of total costs is derived from 10,000 bootstrap replications from 20 imputed datasets using the standard normal method.

3.2 Effects

MMM had a non-statistically significant higher baseline utility and statistically significant higher utility at both end-of-treatment and follow-up. MMM gained 0.360 QALY in the period and MAU 0.337 QALY. After adjusting for baseline utility, the incremental QALYs were estimated to 0.017 for MMM. Table 2 presents the utilities for the observed mean values.

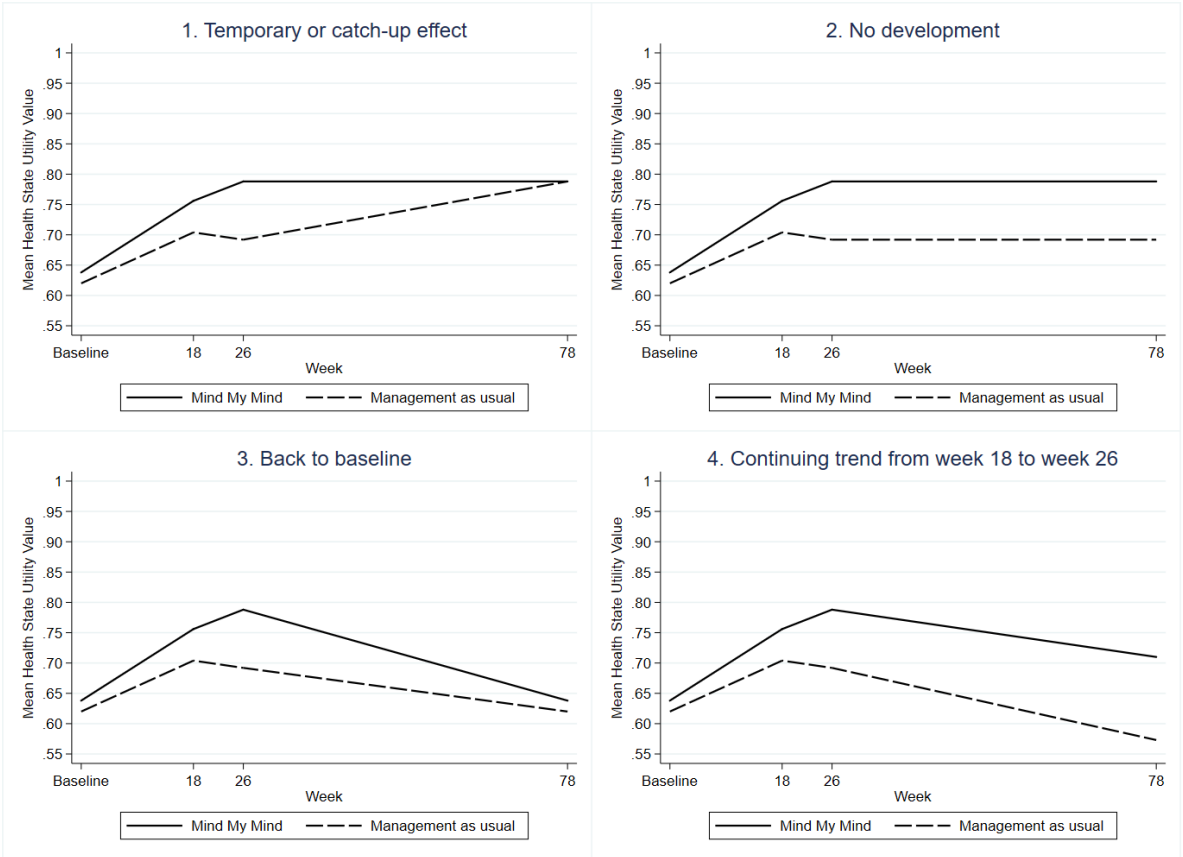
Table 2: Health-related quality of life and quality adjusted life years

	MMM (n=197) Mean ±SE	MAU (n=199) Mean ±SE
Health-related quality of life		
Baseline	0.638 ±0.015	0.620 ±0.015
End-of-treatment (week 18)	0.756 ±0.015	0.704 ±0.017
Follow-up (week 26)	0.788 ±0.014	0.692 ±0.018
QALY	0.360 ±0.006	0.337 ±0.007
Incremental QALYs		
Unadjusted	0.024 ±0.009	
Adjusted for baseline	0.017 (95% CI: 0.006-0.029)	

Note: Values presented are mean (standard error (se)) of the intention-to-treat population from 20 imputed datasets. Adjusted incremental Quality Adjusted Life years (QALY) are adjusted for baseline utility. Confidence interval (CI) is derived from 10,000 bootstrap replications from 20 imputed datasets using the standard normal method.

The four one-year extrapolation scenarios are illustrated in figure 1 with the mean HSUV for the two groups at the three observed time points and at the extrapolated time point (week 78) one year after the last observed time point (week 26). The counter intuitive development in scenario 4 with decreasing mean HSUV was due to a higher number of youths reaching the HSUV ceiling level of 1 during the extrapolation period compared to the number reaching the floor level of 0.053.

Fig 1: Mean health state utility for the four scenario extrapolations



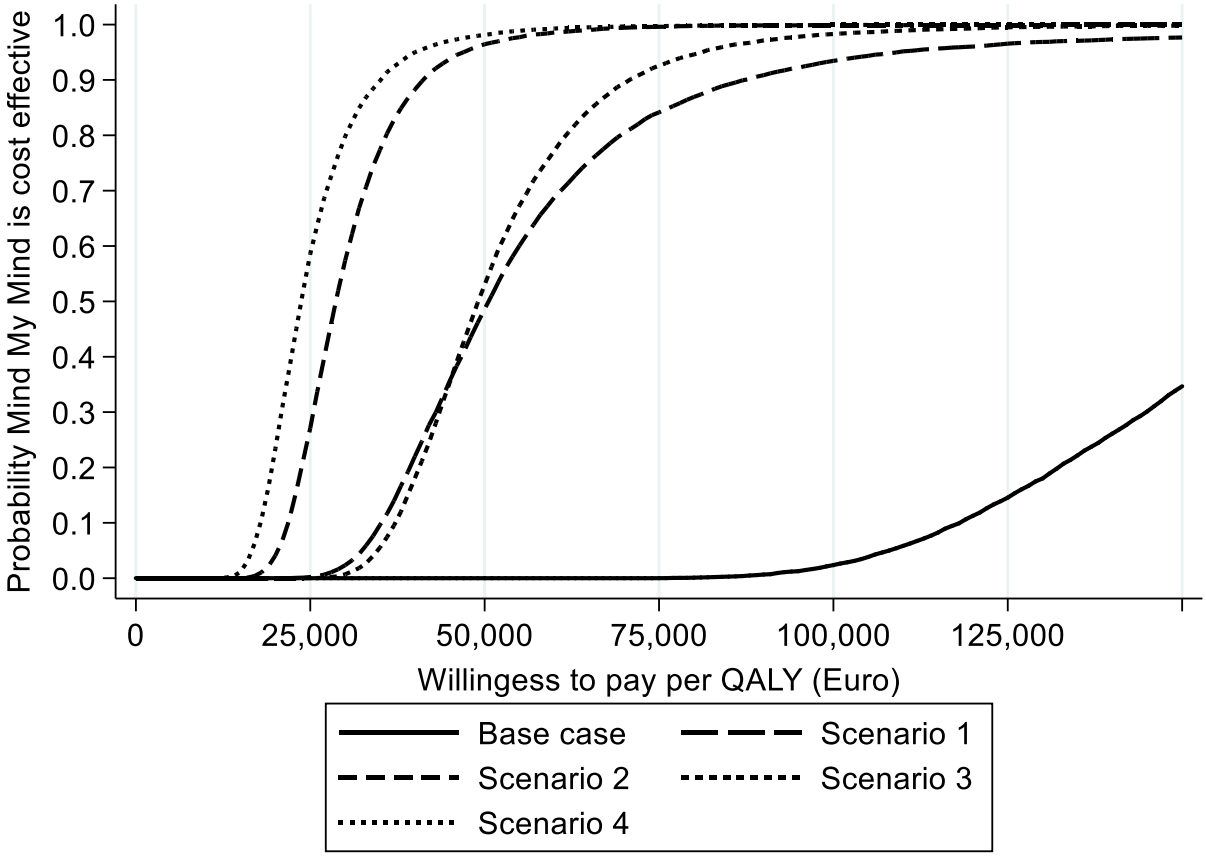
All four scenarios resulted in statistically significant higher incremental QALYs compared with the base case. Adjusted for baseline utility the mean incremental QALY were 0.059 (95% CI: 0.023-0.101) for scenario 1, 0.104 (95% CI: 0.056-0.151) for scenario 2, 0.061 (95% CI:0.032-0.089) for scenario 3, and 0.126 (95% CI: 0.065-0.186) for scenario 4.

3.3 Cost-effectiveness

The ICER was estimated to €170,465 per QALY gained for the restricted base case analysis.

The analyses of the four one-year scenario extrapolations resulted in the following ICERs: €50,480 per QALY gained for scenario 1, €28,659 per QALY gained for scenario 2, €49,069 per QALY gained for scenario 3, and €23,653 per QALY gained for scenario 4. The cost-effectiveness acceptability curves (CEAC) are shown in figure 2.

Fig 2: Cost-effectiveness acceptability curves



3.4 Sensitivity analyses

The ICER was €114,713 when assuming each psychologist would be able to treat 20 youth each instead of the average of 8.2 that was observed in the trial-period.

To analyze the sensitivity of the duration of the scenario extrapolations on results, the ICERs were calculated using the same assumptions but changing the period from one year to six months and 18 months,

respectively. For the six months scenarios the ICERs range from €41,245 to €77,894. For the 18 months scenarios the ICERs range from €16,659 to €37,339. All ICERs and CEACs are available in the supplementary materials Table SM3 and Figure SM1.

The base case ICER is based on QALYs estimated using parent-reported CHU9D and tariffs based on preferences of a youth population. Table 3 shows the results for two separate analyses. One in which the self-reported CHU9D is used in the QALY estimation instead of the parent-reported, and another analysis in which the tariffs are based on preferences of an adult population rather than on a youth population. The two analyses both resulted in non-statistically significantly lower incremental QALYs compared to the base case. The ICER was €40,292 (24%) higher than the base case when using self-reported CHU9D and €121,451 (71%) higher when using the tariffs based on an adult population. CEACs are available in the supplementary materials Figure SM2.

Table 3: Analyses using self-reported CHU9D and tariffs from an adult population

	Self-reported CHU9D Tariffs based on preferences of a youth population		Parent-reported CHU9D Tariffs based on preferences of an adult population	
	MMM (n=197) mean ±se	MAU (n=199) mean ±se	MMM (n=197) mean ±se	MAU (n=199) mean ±se
Health-related quality of life				
Baseline	0.642 ±0.016	0.631 ±0.016	0.806 ±0.009	0.802 ±0.008
End-of-treatment (week 18)	0.728 ±0.016	0.684 ±0.018	0.873 ±0.008	0.844 ±0.009
Follow-up (week 26)	0.769 ±0.015	0.708 ±0.186	0.886 ±0.008	0.839 ±0.009
QALY	0.352 ±0.007	0.335 ±0.007	0.426 ±0.003	0.414 ±0.004
Incremental QALYs				
Unadjusted	0.017 ±0.009		0.012 ±0.005	
Adjusted for baseline	0.014 (95% CI: 0.002-0.027)		0.010 (95% CI: 0.004-0.016)	
ICER	€210,757 /QALY gained		€291,916 /QALY gained	

Note: Values presented are mean (standard error (se)) of the intention-to-treat population from 20 imputed datasets. Adjusted incremental quality adjusted life years (QALY) are adjusted for baseline utility. Confidence intervals (CI) are derived from 10,000 bootstrap replications from 20 imputed datasets using the standard normal method. CHU9D: Child Health Utility 9D. ICER: Incremental Cost Effectiveness Ratio.

4. Discussion

The ICER found in the base case with the time horizon restricted to the trial time period was €170,465, which is considerably higher than standard willingness to pay per QALY. For the scenarios with a continuing effect (scenario 2 and 4) the ICERs were below €30,000, and the MMM intervention had a 90% probability of being cost-effective at a willingness to pay per QALY gained of €41,000 (see figure 2). For scenarios 1 and 3 assuming temporary effects, the ICERs were both estimated around €50,000. The sensitivity analyses addressing the assumption of a one-year extrapolation period showed that all scenarios resulted in ICERs below €40,000, when an 18-month extrapolation period was assumed, while only scenario 2 and 4 had ICERs below €50,000, when a six-month extrapolation period was assumed.

In light of the large proportion of youth suffering from common mental health disorders and the promising effects of CBT-programs, it is surprising that there is limited access to evidence-based programs. MAU was chosen as comparator in the trial as no evidence-based programs are widely accessible for youths with common mental health problems. Most evidence-based CBT-programs are problem-specific and only relevant for a part of the population with common mental health problems. Instead, the transdiagnostic and modular approach in MMM made it possible to offer treatment to a large and heterogenous population of youth with emotional and behavioral problems. A considerable part of the costs of delivering MMM is associated with the training of the psychologists and supervisors. The high training costs support the choice of offering the transdiagnostic MMM instead of traditional CBT-programs as this avoids parallel training costs for three single-disorder programs to target the same group of youth. The sensitivity analysis showed that the cost estimate decreased by 33% when assuming each psychologist would treat 20 youth instead of the average of 8.2 in the trial-period. The high costs were partly due to the training being held at a hotel including the costs of stay. If the same level of training can be delivered in a different setup, for example online, there could potentially be significant savings.

Due to the short observation period, we only considered the immediate effects of the MMM intervention. It is hypothesized that MMM will generate long-term preventive effects by lowering the risk of future severe mental disorders[13]. If this is the case, MMM may be considerably more cost-effective than our results suggest, both due to the health effects but also due to the potential cost savings, as contacts to CAMHS are associated with very large health care costs[29].

In this study we took a health care perspective. A previous review has, however, shown that the majority of costs of mental disorders among youth are in other sectors, primarily the education and social sectors[42]. Furthermore, we have not included the parent's costs and benefits in this study. Positive spillover effects for parents are often observed in interventions that benefit their offspring[43]. We lack data on these broader costs and benefits but given the positive effects of MMM compared to MAU, it is likely that a cost-effectiveness analysis with a societal perspective would generate lower ICERs than observed in this study.

The HRQOL at baseline in our trial was similar irrespective of whether reported by the youth or the parent. However, at end-of-treatment and follow-up the youth rated their HRQOL lower than their parents did in both groups. Using the self-reported CHU9D resulted in a 24% higher ICER. We could only identify one other CUA that reports CHU9D for both respondents[16]. Here, surprisingly, almost no difference in the QALY-gain was found when comparing across respondent types. Our findings align better with previous findings from a review of other preference-based HRQOL instruments used in youth, that concludes that youth and parental responses are not interchangeable[44].

The impact of using the adult tariffs from the UK (rather than the youth tariffs from Australia) was even greater making the ICER 71% higher. This difference may be driven by the youth population placing greater weight on the mental health dimension of CHU9D compared to adults[24]. There are, however, other methodological differences between the two tariffs besides the age and nationality of the samples they are based on, which also could explain some of the difference. The youth tariffs are derived using discrete-choice experiments and a smaller sample time-trade-off experiment with young adults to anchor the tariffs[33], while the adult tariffs are derived using a standard gamble experiment[34].

For health technology assessment agencies, that aim for comparability across cost-effectiveness studies, our finding that QALY estimates are highly sensitive to the choice of informant with respect to both HRQOL

reporting and tariffs, warrants attention. Since these choices are clearly not trivial, they should be further scrutinized in the development of guidelines for CUA of interventions for youth.

5. Conclusion

Whether the MMM intervention is cost-effective depends on the duration of treatment effects, and the intervention's preventive impact. Restricting the period of analysis to the trial period resulted in a cost per QALY-gained well above standard willingness to pay. Applying four different one-year extrapolation scenarios resulted in ICERs within a range that is likely to be considered cost-effective. An additional important source of variation in ICER estimates was the choice of informant for HRQOL and tariffs. This is a methodological aspect that warrants further attention.

Decisions on introducing MMM should take the many uncertainties into consideration. It may be difficult to generate more precise information on long-term effects of MMM via RCTs. However, future register-based studies can contribute to advancing understanding of the preventive effects of MMM.

Reference

1. Patel V, Saxena S, Lund C, et al. The Lancet Commission on global mental health and sustainable development. *The Lancet*. 2018;392(10157):1553-1598. doi:10.1016/S0140-6736(18)31612-X

2. Polanczyk GV, Salum GA, Sugaya LS, Caye A, Rohde LA. Annual Research Review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *J Child Psychol Psychiatry*. 2015;56(3):345-365. doi:10.1111/jcpp.12381

3. Dalsgaard S, Thorsteinsson E, Trabjerg BB, et al. Incidence Rates and Cumulative Incidences of the Full Spectrum of Diagnosed Mental Disorders in Childhood and Adolescence. *JAMA Psychiatry*. 2020;77(2):155. doi:10.1001/jamapsychiatry.2019.3523

4. Copeland WE, Wolke D, Shanahan L, Costello EJ. Adult Functional Outcomes of Common Childhood Psychiatric Problems: A Prospective, Longitudinal Study. *JAMA Psychiatry*. 2015;72(9):892. doi:10.1001/jamapsychiatry.2015.0730

5. Schwartz C, Barican JL, Yung D, Zheng Y, Waddell C. Six decades of preventing and treating childhood anxiety disorders: a systematic review and meta-analysis to inform policy and practice. *Evid Based Ment Health*. 2019;22(3):103-110. doi:10.1136/ebmental-2019-300096

6. Zhou X, Zhang Y, Furukawa TA, et al. Different Types and Acceptability of Psychotherapies for Acute Anxiety Disorders in Children and Adolescents. *JAMA Psychiatry*. 2019;76(1):41-50. doi:10.1001/jamapsychiatry.2018.3070

7. Zhou X, Hetrick SE, Cuijpers P, et al. Comparative efficacy and acceptability of psychotherapies for depression in children and adolescents: A systematic review and network meta-analysis. *World Psychiatry*. 2015;14(2):207-222. doi:10.1002/wps.20217

8. Hetrick SE, Cox GR, Witt KG, Bir JJ, Merry SN. Cognitive behavioural therapy (CBT), third-wave CBT and interpersonal therapy (IPT) based interventions for preventing depression in children and

adolescents. Cochrane Common Mental Disorders Group, ed. *Cochrane Database Syst Rev*. Published online August 9, 2016. doi:10.1002/14651858.CD003380.pub4

9. Eckshtain D, Kuppens S, Ugueto A, et al. Meta-Analysis: 13-Year Follow-up of Psychotherapy Effects on Youth Depression. *J Am Acad Child Adolesc Psychiatry*. 2020;59(1):45-63. doi:10.1016/j.jaac.2019.04.002

10. Waddell C, Schwartz C, Andres C, Barican JL, Yung D. Fifty years of preventing and treating childhood behaviour disorders: a systematic review to inform policy and practice. *Evid Based Ment Health*. 2018;21(2):45-52. doi:10.1136/eb-2017-102862

11. Furlong M, McGilloway S, Bywater T, Hutchings J, Smith SM, Donnelly M. Cochrane Review: Behavioural and cognitive-behavioural group-based parenting programmes for early-onset conduct problems in children aged 3 to 12 years (Review). *Evid-Based Child Health Cochrane Rev J*. 2013;8(2):318-692. doi:10.1002/ebch.1905

12. James AC, James G, Cowdrey FA, Soler A, Choke A. Cognitive behavioural therapy for anxiety disorders in children and adolescents. Cochrane Common Mental Disorders Group, ed. *Cochrane Database Syst Rev*. Published online February 18, 2015. doi:10.1002/14651858.CD004690.pub4

13. Jeppesen P, Wolf RT, Nielsen SM, et al. Transdiagnostic Psychotherapy for Youth with Common Mental Health Problems. *YET Publ*.

14. Hamilton MP, Hetrick SE, Mihalopoulos C, et al. Identifying attributes of care that may improve cost-effectiveness in the youth mental health service system. *Med J Aust*. 2017;207(S10). doi:10.5694/mja17.00972

15. Creswell C, Violato M, Cruddace S, et al. A randomised controlled trial of treatments of childhood anxiety disorder in the context of maternal anxiety disorder: clinical and cost-effectiveness outcomes. *J Child Psychol Psychiatry*. 2020;61(1):62-76. doi:10.1111/jcpp.13089

16. Creswell C, Violato M, Fairbanks H, et al. Clinical outcomes and cost-effectiveness of brief guided parent-delivered cognitive behavioural therapy and solution-focused brief therapy for treatment of childhood anxiety disorders: a randomised controlled trial. *Lancet Psychiatry*. 2017;4(7):529-539. doi:10.1016/S2215-0366(17)30149-9

17. Lynch FL, Dickerson JF, Clarke G, et al. Cost-Effectiveness of Preventing Depression in At-Risk Youth: Post-Intervention and Two-Year Follow-up. *Psychiatr Serv Wash DC*. 2019;70(4):279-286. doi:10.1176/appi.ps.201800144

18. Tie H, Krebs G, Lang K, et al. Cost-effectiveness analysis of telephone cognitive–behaviour therapy for adolescents with obsessive–compulsive disorder. *BJPsych Open*. 2019;5(1). doi:10.1192/bjo.2018.73

19. Jolstedt M, Wahlund T, Lenhard F, et al. Efficacy and cost-effectiveness of therapist-guided internet cognitive behavioural therapy for paediatric anxiety disorders: a single-centre, single-blind,

randomised controlled trial. *Lancet Child Adolesc Health*. 2018;2(11):792-801. doi:10.1016/S2352-4642(18)30275-X

20. Dickerson JF, Lynch FL, Leo MC, DeBar LL, Pearson J, Clarke GN. Cost-effectiveness of Cognitive Behavioral Therapy for Depressed Youth Declining Antidepressants. *Pediatrics*. 2018;141(2). doi:10.1542/peds.2017-1969

21. Hill H, Rowen D, Pennington B, Wong R, Wailoo A. *A Review of the Methods Used to Estimate and Model Utility Values in NICE Technology Appraisals for Paediatric Populations*. NICE DSU Report; 2019.

22. Rowen D, Rivero-Arias O, Devlin N, Ratcliffe J. Review of Valuation Methods of Preference-Based Measures of Health for Economic Evaluation in Child and Adolescent Populations: Where are We Now and Where are We Going? *PharmacoEconomics*. 2020;38(4):325-340. doi:10.1007/s40273-019-00873-7

23. De Los Reyes A, Cook CR, Gresham FM, Makol BA, Wang M. Informant discrepancies in assessments of psychosocial functioning in school-based services and research: Review and directions for future research. *J Sch Psychol*. 2019;74:74-89. doi:10.1016/j.jsp.2019.05.005

24. Ratcliffe J, Huynh E, Stevens K, Brazier J, Sawyer M, Flynn T. Nothing About Us Without Us? A Comparison of Adolescent and Adult Health-State Values for the Child Health Utility-9D Using Profile Case Best-Worst Scaling: ADOLESCENT VERSUS ADULT VALUES FOR THE CHU-9D. *Health Econ*. 2016;25(4):486-496. doi:10.1002/hec.3165

25. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford University Press; 2015.

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

27. Nielsen LG, Rimvall MK, Clemmensen L, et al. The predictive validity of the Strengths and Difficulties Questionnaire in preschool age to identify mental disorders in preadolescence. Hashimoto K, ed. *PLOS ONE*. 2019;14(6):e0217707. doi:10.1371/journal.pone.0217707

28. Wolf RT, Puggaard LB, Pedersen MMA, et al. Systematic identification and stratification of help-seeking school-aged youth with mental health problems - A novel approach to stage-based stepped-care. *Rev*.

29. Wolf RT, Jeppesen P, Gyrd-Hansen D, The CCC2000 Study Group, Oxholm AS. Evaluation of a screening algorithm using the Strengths and Difficulties Questionnaire to identify children with mental health problems: A five-year register-based follow-up on school performance and healthcare use. Fonseca-Pedrero E, ed. *PLOS ONE*. 2019;14(10):e0223314. doi:10.1371/journal.pone.0223314

30. Stevens K. Developing a descriptive system for a new preference-based measure of health-related quality of life for children. *Qual Life Res*. 2009;18(8):1105-1113. doi:10.1007/s11136-009-9524-9

31. Furber G, Segal L. The validity of the Child Health Utility instrument (CHU9D) as a routine outcome measure for use in child and adolescent mental health services. *Health Qual Life Outcomes*. 2015;13(1):22. doi:10.1186/s12955-015-0218-4

32. Wolf RT, Ratcliffe J, Chen G, Jeppesen P. The longitudinal validity of the CHU9D – Evidence from a mental health trial using preference weights from both adult and adolescent populations. *YET Publ*.

33. Ratcliffe J, Huynh E, Chen G, et al. Valuing the Child Health Utility 9D: Using profile case best worst scaling methods to develop a new adolescent specific scoring algorithm. *Soc Sci Med*. 2016;157:48-59. doi:10.1016/j.socscimed.2016.03.042

34. Stevens K. Valuation of the Child Health Utility 9D Index. *PharmacoEconomics*. 2012;30(8):729-747. doi:10.2165/11599120-000000000-00000.

35. Faria R, Gomes M, Epstein D, White IR. A Guide to Handling Missing Data in Cost-Effectiveness Analysis Conducted Within Randomised Controlled Trials. *PharmacoEconomics*. 2014;32(12):1157-1170. doi:10.1007/s40273-014-0193-3

36. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*. 2011;30(4):377-399. doi:10.1002/sim.4067

37. Rubin D, Schenker N. Multiple imputation in health-care databases: An overview and some applications. *Stat Med*. 1991;10(4):585–598.

38. Willan AR, Briggs AH, Hoch JS. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. *Health Econ*. 2004;13(5):461-475. doi:10.1002/hec.843

39. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ*. 2005;14(5):487-496. doi:10.1002/hec.944

40. StataCorp. *Stata Statistical Software: Release 15.* StataCorp LLC; 2017.

41. Efron B. Missing Data, Imputation, and the Bootstrap. *J Am Stat Assoc*. 89(426):14.

42. Snell T, Knapp M, Healey A, et al. Economic impact of childhood psychiatric disorder on public sector services in Britain: estimates from national survey data. *J Child Psychol Psychiatry*. 2013;54(9):977-985. doi:10.1111/jcpp.12055

43. Tubeuf S, Saloniki E-C, Cottrell D. Parental Health Spillover in Cost-Effectiveness Analysis: Evidence from Self-Harming Adolescents in England. *PharmacoEconomics*. 2019;37(4):513-530. doi:10.1007/s40273-018-0722-6

44. Khadka J, Kwon J, Petrou S, Lancsar E, Ratcliffe J. Mind the (inter-rater) gap. An investigation of self-reported versus proxy-reported assessments in the derivation of childhood utility values for economic

evaluation: A systematic review. *Soc Sci Med*. 2019;240:112543. doi:10.1016/j.socscimed.2019.112543

45. Office for Municipal and Regional Wage Data. Published 2019. Accessed January 30, 2020. https://www.krl.dk/#/sirka

46. The Danish association of general practitioners. Honorartabel. Published 2019. Accessed January 30, 2020. https://www.laeger.dk/sites/default/files/honorartabel_2019_oktober-1.pdf

47. The Danish Health Data Authority. Sygesikringsregisteret (SSR) - Sundhedsdatastyrelsen. Accessed January 30, 2020. https://sundhedsdatastyrelsen.dk/da/registre-og-services/om-de-nationale-sundhedsregistre/sundhedsoekonomi-og-finansiering/sygesikringsregisteret

48. The Danish Health Data Authority. Afregning og finansiering (DRG) - Sundhedsdatastyrelsen. Accessed January 30, 2020. https://sundhedsdatastyrelsen.dk/da/afregning-og-finansiering

49. Psykologbehandling - Lægehåndbogen på sundhed.dk. Accessed January 30, 2020. https://www.sundhed.dk/sundhedsfaglig/laegehaandbogen/socialmedicin/behandling/psykologbehandling/

Paper IV Supplementary material

Table SM1: Cost components, assumptions, and sources of unit costs

Cost component	Assumptions	Source of unit costs	Cost estimate
Salaries	The hourly salary of psychologist from the psychological school services and psychologist from the Child and Adolescent Mental Health Services (CAMHS) were separately based on the average salary of psychologists employed in the Danish municipalities and psychologists employed in the Danish regions. An effective working load of 1,600 hours a year were assumed.	Office for Municipal and Regional Wage Data [45]	€47 per hour for psychologist from Educational-psychological advisory service. €47 per hour for supervisors from CAMHS.
Mind My Mind			
Training of psychologist from the psychological school services	It was assumed that all psychologist received the full training package. This was not the case as some psychologist did not participate in the whole trial period. The costs are based on the primary training course where the majority of the psychologists got their training. The cost of training one psychologist was calculated as the average costs of the following plus the salary of the psychologist to participate in the training. • A one-week training course: Four-night stay at a hotel plus salary of teachers. • Three two-day training courses: One-night stay at a hotel plus salary of teachers. • Three one-day methodology courses: Seven participants had local courses. Costs only comprised the salary of teachers.	Cost from the trial budget + salaries	€9,178 per trained psychologist.
Training of supervisor psychologist from the Child and Adolescent Mental Health Services	The same assumptions made for the training of the psychologist from the psychological school services apply for the supervisors, as the supervisor underwent the same training together with psychologist. The supervisors further had the following:	Cost from the trial budget + salaries	€12,119 per trained supervisor.

	• Three one-day courses: Five participants at a hotel setting plus salary of teachers.		
Mind My Mind sessions	The actual number of MMM sessions the youth received were recorded on individual-level. The psychologists were allocated three hours per session for the first three youth they treated, hereafter they were allocated two hours per session.	(see salaries)	
Materials	Materials included the printed Mind My Mind Manual, small games and accessories for the sessions, and a tablet for recording the sessions. It is assumed the materials will last for the treatment of 100 youth.	Estimated by trial manager	€1,338
Supervision	The number of hours of supervision was based on the trial budget. 75% was to be delivered one-to-one online and 25% in groups in each of the four participating municipalities. As it was not possible to decide which youth were discussed when in the supervision, the total number of supervision hours were divided equally among the 197 youth randomized to Mind My Mind. That resulted in an average of 6.1 hours of individual supervision and 4.1 hours of group supervision per youth.	(see salaries)	€792 per youth randomized to Mind My Mind.
Management as usual			
Care-coordination visits	The actual number of care-coordination visits was recorded on individual-level. An hour consultation and half an hour preparation for each meeting was assumed.	(see salaries)	
Individual therapy	The number of sessions with individual therapy provided by the municipality was based on self-reporting by the parents at end-of-treatment (week 18) and at follow-up (week 26). Each session was assumed to have a duration of one hour and require half an hour preparation.	(see salaries)	
Group therapy	The number of sessions with group therapy provided by the municipality was based on self-reporting by the parents at end-of-treatment (week 18) and at follow-up (week 26). Each session was assumed to have a	(see salaries)	

	duration of one hour and require half an hour preparation. It was assumed that six youth take part in group therapy.		
Parental training	The number of sessions with parental training provided by the municipality was based on self-reporting by the parents at end-of-treatment (week 18) and at follow-up (week 26). Each session was assumed to have a duration of one hour and require half an hour preparation.	(see salaries)	
Other health care use			
General practitioner (GP)	For visits at the GP it was assumed that therapy was provided by the GP.	The collective agreement between the Danish Regions and the Danish association of general practitioners [46].	€74 per visit.
Pediatrician	All visits to pedestrians were assumed to be standard consultations.	The Danish Health Data Authority [47].	€66 per visit.
Child and Adolescent Mental Health Services (CAMHS)	The cost of a visit to CAMHS a based on the standard tariff of an out-patient visit to CAMHS.	The Danish Health Data Authority [48].	€247 per visit.
Private-sector psychologist	The prices at private psychologist are not regulated. Prices for subsidized private psychologist are, however, publicly available. The costs of private-sector psychologist were therefore based on these prices.	The Danish Regions [49].	€117 per visit.

Table SM 2: Complete cases analysis

Incremental costs	€2,978 ±140
Incremental QALY adjusted for baseline	0.018 ±0.006
Incremental Cost Effectiveness Ratio	€164,416 /QALY gained

Note: Costs and QALY presented as mean ± standard error. Mind My Mind n=170, Management as usual n=150.

Table SM 3: Sensitivity scenarios

	Incremental Cost Effectiveness Ratio
6 months extrapolation	
Scenario 1	€77,894
Scenario 2	€49,069
Scenario 3	€76,202
Scenario 4	€41,245
18 months extrapolation	
Scenario 1	€37,339
Scenario 2	€20,240
Scenario 3	€36,184
Scenario 4	€16,659

Note: Values presented are based on the intention-to-treat population from 20 imputed datasets. Incremental costs are the same as in base case.

Fig SM1: Cost Effectiveness Acceptability Curves for six- and 18-months scenarios

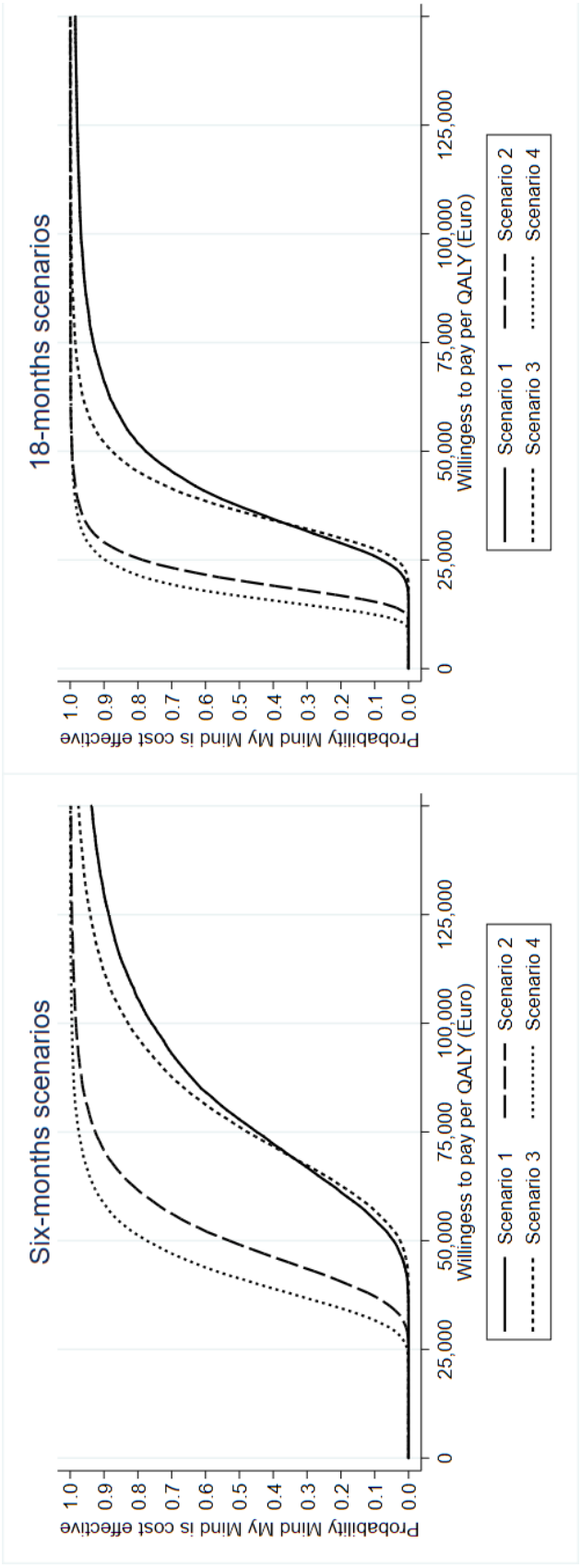


Fig SM2: Cost Effectiveness Acceptability Curves for sensitivity analyses results with self-reported CHU9D and adult tariffs